SCALE-INVARIANT CASCADES IN TURBULENCE AND EVOLUTION

BY

NICHOLAS RYAN GUTTENBERG

B.S., McGill University, 2004

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Physics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2009

Urbana, Illinois

Doctoral Committee:

Professor Karin Dahmen, Chair
Professor Nigel Goldenfeld, Director of Research
Professor Alfred Hubler
Professor Susan Kieffer

Abstract

In this dissertation, I present work addressing three systems which are traditionally considered to be unrelated: turbulence, evolution, and social organization. The commonality between these systems is that in each case, microscopic interaction rules give rise to an emergent behavior that in some way makes contact with the macroscopic scale of the problem. The open-ended evolution of complexity in evolving systems is analogous to the scale-free structure established in turbulent flows through local transportation of energy. In both cases, an invariance is required for the cascading behavior to occur, and in both cases the scale-free structure is built up from some initial scale from which the behavior is fed.

In turbulence, I examine the case of two-dimensional turbulence in order to support the hypothesis that the friction factor and velocity profile of turbulent pipe flows depend on the turbulent energy spectrum in a way unpredicted by the classic Prandtl theory. By simulating two-dimensional flows in controlled geometries, either an inverse energy cascade or forward enstrophy cascade can be produced. The friction factor scaling of the flow changes depending on which cascade is present, in a way consistent with momentum transfer theory and roughness-induced criticality.

In the problem of evolution, I show that open-ended growth of complexity can be obtained by ensuring that the evolutionary dynamics are invariant with respect to changes in complexity. Finite system size, finite point mutation rate, and fixed points in the fitness landscape can all interrupt this cascade behavior, producing an analogue to the integral scale of turbulence. This complexity cascade can exist both for competing and for symbiotic sets of organisms. Extending this picture to the qualitatively-different levels of organization of real lifeforms (viruses, unicellular, biofilms, multicellular) requires an understanding of how the processes of evolution themselves evolve. I show that a separation of spatial or temporal scales can enhance selection pressure on parameters that only matter several generations down the line. Because of this, I conclude that the prime candidates for the emergence of novel evolutionary mechanisms are biofilms and things living in oscillating environments.

Finally, in the problem of social organization, I show that different types of control hierarchies - leaders or communal decision making - can emerge depending on the relationship between the environment in which members of the social group act and the development and exchange of information.

To Gabe Hodziewich

Acknowledgments

I've always found it difficult to pin down my interests into a nicely packaged field of study. The world is filled with interesting things, and saying 'this is what I will dedicate myself to, and only this' was something that I never really felt I could do. Luckily for me, I took to heart the recommendations of Martin Grant when looking for places for my graduate studies and discovered Nigel Goldenfeld's group at the University of Illinois at Urbana-Champaign, and so I would like to thank him both for getting me started on research and for helping me find such an ideal place for my graduate studies. Nigel Goldenfeld's interests are incredibly broad, and even rarer, he is willing to allow - even mandates - his students to pursue a wide variety of research. Despite early reservations that some of my work on the evolution of complexity would be of interest only to Nigel, myself, and a scattered handful of computer science people around the world, Nigel allowed me to take what was originally a project I was doing in my spare time and turn it into a significant part of my thesis.

The atmosphere of Nigel's group was such that we all worked on each other's projects just a little. Group meetings were an opportunity to try to collectively get around stumbling blocks in our research. With the entire group on instant messenger, these meetings would spin off into various discussions on the side at odd times of day and night. For tolerating such distractions, and for many helpful conversations without which much of this work would have been ill-formed, undeveloped, or unconceived, I thank all of the members of the group while I was here: John Veysey, Hector Garcia Martin, Kalin Vetsigian, Nicholas Chia, David Reynolds, Patricio Jeraldo, Tom Butler, Patrick Chan, Badri Athreya, Zhenyu Wang, and Maksim Sipos. I also owe deep thanks to Nigel Goldenfeld for cultivating this sort of environment, not to mention the countless discussions, ideas, suggestions, and other cooperative endeavors.

In particular, I would like to thank Nicholas Chia for teaching me most of the biology that I know and suggesting various lines of inquiry. His understanding of the biologically relevant aspects of my otherwise very abstract work on evolution opened up all manner of lines of inquiry that bore interesting fruit.

It was a class taught by Carl Woese that renewed my interest in artificial life and synthetic evolution. I had been curious about things such as the Tierra and AVIDA projects for some time before, but my investigations had been somewhat groping and unsuccessful. Discussions during Carl Woese's class led me to consider non-Turing-complete simulations where I could more readily understand what was actually being evolved, which led me to write the 'Foodchain' and 'PlantNet' simulations in rapid succession. I would like to thank Carl for the inspiration that led to me getting started in serious evolution research.

The yearly FIBR meetings at the Santa Fe Institute were also very inspiring for the evolutionary aspects of my work. In particular, many of my ideas on the abstraction transition and evolution of biofilms (work which is not strongly represented here as it is still in progress) originated during these meetings after presentations and discussions. In particular, I would like to thank Daniel Segré for discussions about chemical networks, Eric Smith for discussions about the loss of germline in the transition to multicellularity, and John Baross for discussions about biofilms and what I have since termed my set of 'kitchen sink' simulations.

Much of the work that I present on turbulence has at its root our turbulent Friday meetings. Through these meetings I, Nigel, Gustavo Gioia, Pinaki Chakraborty, Tuan Tran, Carlo Zuniga Zamalloa, and Walter Goldburg took turns throwing around hypothetical closures, ways to understand the weirdness that constantly dogged the research, and other such things. Our meetings often became very polarized and loud, and so I would like to thank everyone for their patience and insights. From that chaos, we've managed some sort of emergent theory.

Finally, I would like to thank Stephanie Wendler. She supported me during foul moods of writer's block and tedious editing, and let me use her as a sounding board, despite having no background in physics or biology. Her presence helped me get myself unstuck and into the right mindset to write again.

The work presented in this dissertation was supported by the National Science Foundation under Grants No. NSF-EF-0526747 and NSF DMR 06-04435, and by the University of Illinois Distinguished Fellowship, Edelheit Fellowship, and Drickamer Fellowship.

Table of Contents

List of Figures

Chapter 1

Introduction

In this dissertation, I present work addressing three systems which are traditionally considered to be unrelated: turbulence, evolution, and social organization. The commonality between these systems is that in each case, microscopic interaction rules give rise to an emergent behavior that in some way makes contact with the macroscopic scale of the problem. Thus, it is often not necessary to know exactly what the microscopic interactions are in detail - it is sufficient to understand what macroscopic symmetries they possess. Otherwise, much of the work here would be hopelessly specific to particular microscopic choices in the model building.

1.1 Turbulence

Turbulence is locally just the product of viscosity and Galilean invariance. A problem which would be a simple matter of linear diffusion becomes highly nonlinear when the elements of the system are allowed to be in relative motion with respect to each other. In this case the microscopic rules are fairly clear, but the emergent behavior is very complex. The turbulent flow establishes a range of scales in which advection dominates and transports conserved quantities to larger or smaller scales. The properties of the turbulent flow are then determined by the size of this inertial range. From this one can define a dimensionless number characterizing the flow - the Reynolds number.

Turbulence is hard to pin down - there are very few exact results in turbulence research. Even though the governing equation has been known for some time, its solutions defy exact analysis. Far more success has been obtained with statistical approaches, dimensional analysis, and simplifying pictures that incorporate assumptions about what is and what is not important about a given flow. Intuitively, the rich structure of a turbulent flow should be important in determining how that flow behaves, but that connection has been elusive. In my discussions on turbulence, I will present work that has been done to make those connections, linking the scale-free structure of turbulent flows to macroscopic properties such as their transport behaviors and mean velocity profiles. These connections are similar in construction to the links between the scale-free structure that arises from criticality in statistcal mechanics and the associated scalings of the order parameter and other macroscopic variables with respect to the thermodynamic variables[1].

The approach I will present in Chapter 3 to understanding the connection between the microscopic and macroscopic behavior of turbulence is the most grounded of the three topics. There is a generally accepted equation of motion for the system - the Navier-Stokes equation, and it is supported by an extensive body of experimental work and measurements of turbulent flows. In Chapter 3 I will begin by explaining the momentum transfer theory of Gioia and Chakraborty[2], which connects the friction factor of pipe flows to the energy spectrum of the flow's turbulent structure. I will also discuss how this connects with the picture of turbulence as a nonequilibrium critical phenomenon, as proposed by Goldenfeld[1]. I will present work I did with Nigel Goldenfeld, Gustavo Gioia, and Pinaki Chakraborty in using the basic assumption of momentum transfer theory to predict the velocity profile of the turbulent flow. This relies on the interconnectedness of the set of three quantities: energy spectrum; friction factor; and velocity profile. I extend these results to the case of two-dimensional turbulent flows, in which the energy spectrum may contain both an inverse energy cascade[3; 4] and a forward enstrophy cascade[4]. Because the energy spectrum of two-dimensional turbulence may be different from that of three-dimensional turbulence, the predictions of the momentum transfer theory for the velocity profile and friction factor also change, and so two-dimensional turbulence can be used as a test for the assumptions of the momentum transfer theory. I will also discuss the transition to turbulence and the question of the lifetime of turbulent puffs near the transitional Reynolds number. In work done with Goldenfeld and Gioia, we found a relationship between the lifetime of these turbulent puffs and the energy fluctuation distribution of turbulent flows. The stability of the puff depends on the maximum of a set of energy fluctuations over the puff's volume being above a critical threshold.

In Chapter 4, I will discuss the tool I use to examine two-dimensional turbulent pipe flows in order to test the momentum transfer theory. In order to do this, I constructed a two-dimensional Navier-Stokes solver using conformal mapping to allow for boundaries with nontrivial geometry. I will present the details of this solver, and of the conformal mapping technique. The conformal map causes only minor alterations to the Navier-Stokes equation in the form of two additional local body forces. This can be used to analytically extend the results of smooth pipe flow into the case of rough pipe flow.

In Chapter 5, I will present simulations of two-dimensional pipe flows with smooth and rough walls, in which the energy spectrum was controlled by changing the method of turbulence generation. From these simulations I measured the friction factor and the profiles of various salient quantities of the flow: transverse velocity, energy, dissipation, and shear. These results support the predictions of the momentum transfer theory, and clearly display the spectral dependence of the macroscopic properties of the flow. Furthermore, the friction factor curves collapse onto a universal curve in the same fashion as the data of Nikuradze[5] do in the analysis of Goldenfeld[1], but with the appropriate changes to the data collapse form to take into account the predicted differences in two-dimensional flow. I will compare these results to experiments on turbulent soap film flows performed by the groups of our collaborators Walter Goldburg and Hamid Kellay.

1.2 Evolution

Next, I will discuss the evolutionary systems that I studied as part of this research. The principal question that we will be concerned with is the mechanism by which novel forms and dynamics are generated from a system that is perhaps very simple to begin with. I propose that the generation of new and more complex structures is similar to the cascades of turbulent flow, in which local interactions between the organisms is what gives rise to structure, rather than the response of the evolving organisms to a system which is inherently structured and complex in its own right. If this is the case, it is necessary to determine which interactions between organisms will lead to such complexification, and what will eventually limit it (similar to how the turbulent cascades eventually stop when they reach the scale of the system and the scale of molecular viscosity).

In Chapter 6, I will discuss the nature of evolutionary complexification and the difference between novelty and complexity. It may be possible for an organism to become more complex without actually developing novel modes of interaction. Furthermore, as organisms in the system evolve, there may develop interactions which somehow fundamentally change the way that evolution works. These events I refer to as abstraction transitions. They are abstractions in the sense that the level of representation that exists prior to the transition becomes a background against which some new level of organization operates. The detail of the lower level of dynamics is coarse-grained away by the operation of the new mode of interaction. They are transitions because such changes in the operational modes of evolution open up new degrees of freedom and change the symmetries of the system, and so new regions of phase space open up.

Chapter 7 of this thesis concerns ‘Foodchain’, a model of the evolution of complexity through evolving predator-prey interactions. I will discuss the effects of mutation rate and system size, and describe the necessary conditions for this type of complexity cascade to occur. Similar to the case of turbulent flows, the scaling of quantities in the complexity cascade are all related, and so a data collapse of the saturation complexity of the system is attained as a function of system size and mutation rate. The data collapse scalings are explained in the context of the theory of population genetics.

In Chapter 8, I will discuss PlantNet, which is a model of trees competing for sunlight. The genomes of organisms map onto their shapes in a shared space. As organisms cast shadows upon each other, there is a cost associated with not being as tall as one’s neighbors. The end result is that the system evolves in such a way that the global fitness decreases, as there is a cost associated with being large. The effects of a fixed fitness landscape on the complexity cascade are similar to the effects of mutation rate and finite system size, and fit into the general critical scaling picture.

In Chapter 9, I discuss the evolutionary dynamics of parameters of evolution such as the mutation rate. There is competition between adaptations that have benefits only after time has passed (a change in the evolutionary dynamics) and adaptations that are immediately useful. When a timescale or lengthscale exists in the system, this can tune between the two modes of evolution. I present a way to measure evolutionary pressure, and show the spatial clustering (synchronous selection) and periodic strong selection events have similar effects on meta-evolution.

1.3 Social Organization

The last section of this thesis, Chapter 10, discusses the emergence of social organization from the interactions between individual agents. Here we are concerned with the consequences of shared information, and how the ability to exchange information changes the collective dynamics of a system. This work applies not only to the aforementioned issue of social organization, but also to the dynamics of biological systems that have some form of information sharing: quorum sensing in unicellular organisms[6], horizontal gene transfer, and the foraging behavior of insects.

1.4 My Contributions

In the work on turbulence, I derived the predictions of the momentum transfer theory in two dimensions - for both the friction factor and velocity profile. I contributed to the three-dimensional velocity profile by suggesting that local values of ϵ and η must be used to determine the local transport, and helped in creating a solid framework for deriving the final version of the velocity profile derivation which I present from the Navier-Stokes equation. I also contributed by solving the resulting equations numerically. I derived the method by which the friction factor could be inverted to predict a velocity profile, and performed the comparisons in two dimensions. Furthermore, I wrote the two-dimensional Navier-Stokes solver that was used for all of the turbulence simulations presented in this dissertation. As part of this I developed the conformal mapping technique used to map the rough-walled pipe onto a rectangular boundary.

In the work on evolution, I designed and simulated the in-silico evolutionary models presented - Foodchain, PlantNet, and the simple systems used for the meta-evolution study. I originated the idea of the complexity cascade and its connection to the inertial range in turbulence, and the idea of abstraction transitions in which the mode of evolution changes via the introduction of new degrees of freedom. I designed the method by which selection pressure was measured locally in time in the meta-evolution study.

Chapter 2

Turbulence

Turbulent flows are marked by rich structure over a range of scales—they host fluctuations, vortices, tangles, and other coherent structures that continue to defy a detailed, analytical understanding[7; 8]. When parameterized in terms of the typical flow speed U, characteristic length scale L and kinematic viscosity of the fluid ν, three-dimensional turbulence exhibits universal phenomena as the Reynolds number $\mathrm{Re} \equiv UL/\nu \to \infty$. Most famously, in a theory referred to as K41[9; 10], the dependence of the fluctuation energy spectrum $E(k)$ on wavenumber and mean energy transfer rate $\bar{\epsilon}$ occurs in a way that is independent of ν: $E(k) = \bar{\epsilon}^{2/3} k^{-5/3}$ for values of wavenumber in the so-called inertial range, intermediate between the scales of forcing and the scales where molecular viscosity becomes significant. In this inertial range, turbulent eddies break up into smaller eddies through a mechanism which is to a first approximation Hamiltonian, and results in a cascade of energy to smaller length scales[11].

During the 1930's, Nikuradse undertook a systematic series of measurements of the pressure drop across a turbulent pipe flow as a function of Re[5] and also as a function of r/R, the scale of the roughness of the pipe walls r, normalized by the pipe radius R[12]. The former measurements provided strong support for Prandtl's boundary layer concept[13], and have been replicated and surpassed only recently[14], while the latter measurements, despite recent efforts[15; 16], remain to this day the most complete data set of its kind, spanning three orders of magnitude in Reynolds number and a decade in the dimensionless roughness r. These data reveal that the frictional drag experienced by a turbulent fluid in a pipe with rough walls is a non-monotonic and complicated function of Reynolds number and roughness, which despite intense interest and practical importance (see, e.g. Ref. [13]), has only begun to be understood[1; 2] through two related developments.

First, Gioia and Chakraborty[2] estimated the momentum-transfer between the walls of the pipe and the flow, explicitly taking into account the presence of roughness. Their resultant formula for the dimensionless friction factor (defined precisely below) is expressed in terms of the turbulent kinetic energy spectrum $E(k)$, and thus makes a direct connection between a macroscopic flow property and the velocity field correlations. Second, Goldenfeld[1] pointed out that the power law behavior of Nikuradse's friction factor data in the regimes $\mathrm{Re} \to \infty$ and $r/R \to 0$ was analogous to critical phenomena, where the inverse Reynolds number and roughness play similar roles to, for example, the coupling constant and external magnetic field in an Ising model. Consequently, the dependence of Nikuradse's data on Re and r can be collapsed onto a universal function with sufficient precision for intermittency corrections to be extracted[17]. These results show that the friction factor reflects the nature of the turbulent state through its dependence on the energy spectrum, and that the turbulent state is itself a manifestation of a non-equilibrium critical point at $\mathrm{Re} = \infty$ and $r/R \to 0$.

The purpose of the research that we present in the following chapters is to test the claims of Refs. [2] and [1] in a context where detailed calculations are in principle possible: the case of two-dimensional soap-film turbulence[18; 19]. Here, a soap film is supported between two vertical wires, and the draining flow provides a versatile laboratory for exploring two-dimensional turbulence[19]. It is well-understood that the nature of turbulence in 2D is different from 3D: there is no vortex stretching, for example. Nevertheless, turbulent phenomena exist, and possess the novelty that there are two cascades: an energy inverse cascade that runs from small to large scales[3; 4], and a forward cascade[4] in the enstrophy $\Omega \equiv |\nabla \times \mathbf{v}|^{\mathbf{2}}$, where $\mathbf{v}$ is the fluid velocity field. This enstrophy cascade yields an energy spectrum $E(k) = \beta^{2/3}k^{-3}$, where β is the rate of transfer of enstrophy.

Prior work, dating back to Prandtl and others (for a review see Ref. [13]) is not able to make a prediction about the friction factor in these cases, because it has no specific representation of the nature of the turbulent state, and in particular is disconnected from the energy spectrum. On the other hand, the momentum-transfer theory of Gioia and Chakraborty[2] can reflect the character of 2D turbulent states, as expressed by the energy spectrum, through the dependence of the friction factor on Re and r. It will be shown that the momentum-transfer theory predicts a significant dependence of the friction factor on the nature of the turbulent cascade, one that is observed in direct numerical simulations reported here, and which obeys the scaling predicted by roughness-induced criticality. Thus our direct numerical calculations agree well with the momentum-transfer and roughness-induced criticality picture, and strongly suggest that the standard picture of turbulent boundary layers is incomplete.

Chapter 3

Theory of turbulent pipe flows

3.1 Reynolds Number

The first step towards an understanding of turbulent pipe flows is to characterize the possible states of a pipe flow. At very low velocities, high viscosities, or small pipes, the flow is laminar. There are no spatial or temporal fluctuations in the flow variables. As the flow becomes faster, the pipe larger, or the viscosity smaller, fluctuations with a single period develop, quickly giving way to chaotic motion.

In a pipe with fixed geometry (aspect ratio, shape, etc.), the only relevant variables to describe the fluid motion are the fluid density ρ (units of $[M]/[L]^d$), dynamic viscosity μ (units of $[M][L]^{2-d}/[T]$), average velocity U (units of $[L]/[T]$), the length of the pipe L, and the linear dimension of the pipe's cross section R. These variables can be combined to construct various dimensionless numbers to describe the regimes of flow. As we are interested in the ideal case of infinitely long pipes, L does not contribute to characterizing the turbulent flow. We are then left with only one possible dimensionless number, the Reynolds number[20]:

$$\mathrm{Re} \equiv \frac{UR\rho}{\mu} = \frac{UR}{\nu} \tag{3.1}$$

where $\nu = \mu/\rho$ is the kinematic viscosity. Though there is usually some ambiguity in the lengthscale R, turbulent flows across multiple systems will generally have the same behavior at similar values of Re. The laminar state is stable below Re $\approx 10^3$, beyond which the transition to a turbulent state rapidly occurs.

3.2 Navier-Stokes Equations

Let us begin our treatment of turbulent flows with the Navier-Stokes equations, a continuum model of fluid transport. Consider a domain Ω with boundary S in a volume of fluid. The fluid has associated with it a conserved momentum $\mathbf{p} = \rho\mathbf{V}$ and a pressure P. By taking into account the fluxes $\mathbf{\Phi}$ at the boundaries of the volume and the influence of external forces (in this case, the pressure field), we can write down an equation for the time rate of change of the total momentum in the domain due to advection:

$$\int_\Omega \frac{d\mathbf{p}_j}{dt} d^3r = \int_S \mathbf{\Phi}_{ij}\hat{\mathbf{n}}^i d^2\sigma - \int_\Omega \nabla\mathbf{P}_j d^3r \tag{3.2}$$

The flux $\mathbf{\Phi}_{ij}$ is a rank two tensor $-\mathbf{V}_i\mathbf{p}_j$. We can relate the integral around the boundary with an integral over the volume using Stokes theorem:

$$\int_\Omega \frac{d\mathbf{p}_j}{dt} d^3r = -\int_\Omega \nabla_i\mathbf{\Phi}^i_j d^3r - \int_\Omega \nabla\mathbf{P}_j d^3r \tag{3.3}$$

If we take the limit of an infinitessimal domain, the integrals are replaced by their values at the center of the domain:

$$\frac{d\mathbf{p}_j}{dt} = -\nabla_i\mathbf{V}^i\mathbf{p}_j - \nabla\mathbf{P}_j \tag{3.4}$$

$$\frac{d\mathbf{p}_j}{dt} = -\mathbf{p}_j\nabla_i\mathbf{V}^i - \mathbf{V}^i\nabla_i\mathbf{p}_j - \nabla\mathbf{P}_j \tag{3.5}$$

If the flow is incompressible, then $\nabla_i\mathbf{V}^i = 0$ and ρ is a constant, and so we have:

$$d\mathbf{p}/dt + (\mathbf{V} \cdot \nabla)\mathbf{p} = -\nabla \mathbf{P} \tag{3.6}$$

This is the Euler equation[20], which describes an inviscid, incompressible flow. To add the effects of viscosity, we need only consider the additional diffusive flux Φ of momentum across the boundaries of our infinitessimal cell:

$$\mathbf{\Phi}_{ij} = -\nu \nabla_i \hat{\mathbf{n}}^i \mathbf{p}_j \tag{3.7}$$

Here ν is the kinematic viscosity. This gives rise to an additional term $\nu \nabla^2 \mathbf{p}$, giving us the Navier-Stokes equations:

$$d\mathbf{V}/dt + (\mathbf{V} \cdot \nabla)\mathbf{V} = \nu \nabla^2 \mathbf{V} - \nabla \mathbf{P}/\rho \tag{3.8}$$

$$\nabla \cdot \mathbf{V} = 0 \tag{3.9}$$

We need an additional equation to determine the pressure P and ensure that the incompressibility condition is met. We do this by finding the time rate of change of the divergence $\nabla \cdot \mathbf{V}$:

$$d\nabla \cdot \mathbf{V}/dt + \nabla \cdot (\mathbf{V} \cdot \nabla)\mathbf{V} = \nu \nabla^2 \nabla \cdot \mathbf{V} - \nabla^2 P/\rho \tag{3.10}$$

We now determine P such that this is zero. This gives us the pressure equation:

$$\nabla^2 P = -\rho \nabla \cdot (\mathbf{V} \cdot \nabla)\mathbf{V} \tag{3.11}$$

While the Navier-Stokes equations give a description of the underlying motions of turbulence and there are a number of known exact solutions for specific flows[21; 22], in general these equations are impossible to solve analytically. Some progress can be made with approximate solution via linear-stability analysis, but the results are not consistent with the actual dynamics of the transition to turbulence[23–28].

The Navier-Stokes equations are of primary utility in giving us a framework to build time or space-averaged models, and in numerical simulation of turbulence. Our attempts at understanding the dynamics of turbulence needs more physical insight to proceed.

3.3 Energy Spectrum

Let us turn now to a statistical description of the turbulent state. While we cannot hope to exactly solve the Navier-Stokes equation for any given system, we can perhaps say something about the average behavior over long times and spaces, correlation functions, velocity distributions, and other such statistical quantities.

Let us consider the function of the advection operator $(\mathbf{V}\cdot\nabla)\mathbf{V}$ in the vicinity of a small eddy superimposed on a background velocity field, such that the velocity is $\mathbf{V}=\mathbf{V}_{eddy}+\mathbf{V}_0$. In this case, the advection operator is:

$$(\mathbf{V}\cdot\nabla)\mathbf{V}=(\mathbf{V}_{eddy}\cdot\nabla)\mathbf{V}_{eddy}+(\mathbf{V}_0\cdot\nabla)\mathbf{V}_0+(\mathbf{V}_{eddy}\cdot\nabla)\mathbf{V}_0+(\mathbf{V}_0\cdot\nabla)\mathbf{V}_{eddy} \quad (3.12)$$

These four terms can be considered: the action of the eddy upon itself; the action of the background velocity field upon itself; the action of the eddy upon the background velocity field; and the action of the background velocity field upon the eddy. We will discuss two cases - one in which the background velocity field has turbulent structure only at a significantly different scale than the eddy, and one in which they are both have turbulent structure at similar scales.

In the first case, we will make use of the Taylor frozen turbulence hypothesis[29]: that turbulent fluctuations in a background velocity field are passively advected by it. For one of the fields (the one with smaller structure), the larger field appears to be constant, and so the smaller structure is just advected by a constant velocity field and as such is unchanged. On the other hand, the smaller structure cannot do much to the larger structure, as $\nabla \mathbf{V}$ is very small. So the cross terms between the velocity fields are both small when the scales involved are very different. The self-terms in either case represent the motion of that field independent of the interaction. So it seems that turbulent structures at different scales will not interact strongly, though their own scales may change due to self-interactions.

In the second case, when scales are similar, it is possible for the eddy to be significantly sheared and stretched by the underlying field (and vice-versa). This may either cause the scale to increase (in the direction of stretching) or decrease (perpendicular to the direction of stretching). However, this change in scale is proportional to the current scale of the eddy - there is no fixed scale involved.

It is therefore expected that the energy and other quantities carried by eddies should propagate to larger or smaller scales via a process that is local in scale space. There may also be sources and sinks of these quantities. The advection term conserves energy, however, and so only the viscosity and pressure gradient can change the energy in the system. In the case of the viscosity, energy is dissipated most strongly at the smallest scales, and so in the limit of a large Reynolds number, there should be a scaling range in the energy as a function of scale that is not influenced directly by the viscosity, except as a boundary condition. This is known as the inertial range of turbulence, predicted by Kolmogorov and Obhukov[9; 10]. There is still the possibility of energy injection via a pressure gradient:

$$\frac{dE_P}{dt} = \langle \mathbf{V} \nabla \mathbf{P} \rangle \tag{3.13}$$

This can be rewritten:

$$\frac{dE_P}{dt} = \langle \nabla \cdot P\mathbf{V} - P \nabla \cdot \mathbf{V} \rangle \tag{3.14}$$

For incompressible turbulence, the second term vanishes. Furthermore, for homogeneous turbulence the average value of a spatial gradient will be zero. This means that for the case of homogeneous, incompressible turbulence there is no energy injection due to pressure. As such, the turbulence will eventually decay unless there is some external driving force. The scale at which the driving force operates will interrupt the scaling range of advective energy transfer and set the upper limit of the range. The reason it sets the upper limit rather than the lower limit is that there are no other mechanisms for dissipating energy at larger scales, and so energy must be transported from the source towards the sink - viscosity - at the smallest scales. We will later show that in two-dimensional turbulence this is not necessarily the case due to the addition of a second conserved quantity.

We can use a simple argument to determine what the form of the energy spectrum in the inertial range should be. It has been established that energy - a conserved quantity - is being transported across scales without the presence of any sources. This means that the amount of energy being transported at every scale (in the inertial range) must be a constant ϵ. Once the viscous range is reached, this assumption will fail as that energy will be very quickly dissipated.

Considering the energy contained between two wavenumbers k and $k + dk$, this is: $\int_k^{k+dk} dE = \int_k^{k+dk} E(k)dk$. The argument of the integral, $E(k)$, tells us the part of the total energy of the flow that is contained in the vicinity of a wavenumber k.

Let us determine the relationship between the energy spectrum $E(k)$ and the velocity field in d dimensions. If we take the velocity field and filter it, such that all structure with wavenumber less than k_c is removed, then the energy of the resultant velocity field is $\int_{k_c}^{\infty} E(k)dk$. Writing the Fourier expansion of $\mathbf{v}$, this gives the relation:

$$\int_{k_c}^{\infty} E(k)dk = \frac{\rho}{2} \int d^d\mathbf{r} \int_{k_c}^{\infty} e^{i\mathbf{k}\mathbf{r}} \tilde{\mathbf{v}}(\mathbf{k}) d^3\mathbf{k} \int_{k_c}^{\infty} e^{i\mathbf{k}'\mathbf{r}} \tilde{\mathbf{v}}(\mathbf{k}') d^d\mathbf{k}' \tag{3.15}$$

The integral $\int d^d r e^{i(\mathbf{k}+\mathbf{k}')\mathbf{r}}$ is a delta function $\delta(\mathbf{k} - \mathbf{k}')$, so:

$$\int_{k_c}^{\infty} E(k)dk = \frac{\rho}{2} \int_{k_c}^{\infty} \tilde{\mathbf{v}}(\mathbf{k}) \tilde{\mathbf{v}}(-\mathbf{k}) d^d\mathbf{k} \tag{3.16}$$

This implies that:

$$E(k) \propto k^{d-1}|\tilde{\mathbf{v}}(\mathbf{k})|^2 \tag{3.17}$$

Furthermore, we can relate this to the correlation function of the turbulent velocity field. Let us consider the two-point correlation tensor $S_{ij}(r) = \langle u_i(\mathbf{x})u_j(\mathbf{x}+\mathbf{r})\rangle$. The Fourier expansion of S_{ij} is:

$$\begin{aligned} S_{ij}(r) &= \int d^d\mathbf{r'} \int d^d\mathbf{k} \int d^d\mathbf{k'} e^{i(\mathbf{k}+\mathbf{k'})\mathbf{r'}} e^{i\mathbf{k'}\mathbf{r}} \tilde{u}_i(\mathbf{k})\tilde{u}_j(\mathbf{k'}) \\ &= \int d^d\mathbf{k'} e^{-i\mathbf{k'}\mathbf{r}} \tilde{u}_i(-\mathbf{k'})\tilde{u}_j(\mathbf{k'}) \end{aligned} \tag{3.18}$$

So the energy spectrum is related to the spatial correlation functions by:

$$E(k) \propto \tilde{S}^i_i(k) \tag{3.19}$$

We will use a dimensional analysis argument to determine the form of $E(k)$. Observe that $E(k)$ has units of $[L]^3/[T]^2$. The only other quantities that are relevant to the inertial range are k, with units $[L]$, and ϵ, with units $[L]^2/[T]^3$. If there is complete similarity in terms of the viscosity and the energy injection scale I, then one can use dimensional analysis to say that the energy spectrum must take the following form:

$$E(k) \propto \epsilon^{2/3} k^{-5/3} f(k\eta) g(kI) dk \tag{3.20}$$

Here η is the Kolmogorov lengthscale, at which the inertial range terminates due to viscous forces. The asymptotic behavior of f and g are that they tend to a nonzero constant when their argument limits to zero or infinity, respectively. If one considers the interaction between the transported energy ϵ and the viscosity ν, this suggests that the scale η should be:

$$\eta \propto (\frac{\nu^3}{\epsilon})^{1/4} \tag{3.21}$$

For a finite system - for instance, pressure-driven pipe flow - the energy injection scale I will be related to the system size. This gives the Reynolds number its meaning in that context - the Reynolds number tells you how large the inertial range is. If one considers ϵ to be set by the finite system size L, then by dimension analysis:

$$\epsilon \propto \frac{U^3}{L} \tag{3.22}$$

This implies that $\eta = L\mathrm{Re}^{-3/4}$.

It is possible to make some statements about the unknown functions f and g. There are theoretical predictions for f using models of the energy transfer function between wavenumbers. If one assumes that the effect of turbulent structure is a renormalized eddy viscosity[30], then the effect of turbulent shear can be absorbed into the energy dissipation term. This results in a k^{-7} behavior in the viscous range at large k.

The function f is often taken to be a decaying exponential $\exp(-b\eta k)$, based on empirical observations of spectra[31; 32]. Can this be justified theoretically? If one estimates the rate of transport from dimensional considerations, this gives rise to an energy flux[30; 33]:

$$S(k) = \alpha(k)E(k) = \alpha^{-1}\epsilon^{1/3}k^{5/3}E(k) \tag{3.23}$$

In the steady-state, this gives an energy spectrum of the form:

$$E(k) = \alpha\epsilon^{2/3}k^{-5/3}\exp(-\frac{3}{2}\alpha(\eta k)^{4/3}) \tag{3.24}$$

The explicit dependence of the argument of the exponent on viscosity suggests that α_P should scale linearly with viscosity in order to cancel out that dependence. This result would appear very similar to the $\exp(-b\eta k)$ empirical form.

The limiting behavior of the function g can be determined through an expansion of the integral representation of the energy spectrum following the derivation in Hinze[34]. The asymptotic behavior of the spectrum at small k can be obtained by looking at the relation between the distribution of energy due to the fluctuations along a pair of directions i, j: $S_{ij}(\mathbf{r}, t) = v_i(\mathbf{r}, t)v_j(\mathbf{r}, t)$ and its fourier transform $E_{ij}(\mathbf{k}, t)$:

$$E_{ij}(\mathbf{k}, t) = \int d^{d-1}\Omega \int_0^\infty r^{d-1} S_{ij}(\mathbf{r}, t) e^{i\mathbf{k}\cdot\mathbf{r}} dr \tag{3.25}$$

We have made the assumption that the turbulence is isotropic so $S_{ij}(\mathbf{r}) = S_{ij}(|r|)$. Fixing the direction of k (without loss of generality, due to the assumption of isotropy), this is:

$$E_{ij}(k_x, t) = C \int_0^\infty r^{d-1} S_{ij}(r, t) dr \int d\theta e^{i|k||r|\cos(\theta)} \int d^{d-2}\Omega \tag{3.26}$$

The θ integral is the Bessel function $J_0(k)$ which behaves as $2\pi - \frac{\pi}{2}k^2r^2 + ...$ as $k \to 0$. If one integrates over all possible directions of the $\mathbf{k}$ vector, one finds the connection to the energy spectrum: $E(k) = Ck^{d-1}E_{ii}(k_x, t)$, where C is the solid angle integral in the appropriate number of dimensions (e.g. $\int_0^{2\pi} d\theta = 2\pi$ in two dimensions, and $\int_0^{2\pi}\int_{-\pi}^{\pi} \cos(\phi) d\theta d\phi = 4\pi$ in three dimensions).

So the behavior of the spectrum as $k \to 0$ is:

$$E(k) = C_1 k^{d-1} \int_0^\infty r^{d-1} S_{ii}(r, t) dr - C_2 k^{d+1} \int_0^\infty r^{d+1} S_{ii}(r, t) dr + ... \tag{3.27}$$

The first integral vanishes for incompressible flows[34], so in three dimensions the energy spectrum behaves as $E(k) \propto k^4$ for small k. This is insufficient to determine the crossover function between this regime and the inertial range, but one can write down a function with the correct asymptotic behavior:

$$g(k) = (1 + \gamma k^{-2})^{-17/6} \tag{3.28}$$

This satisfies the criterion that the combined form: $k^{-5/3}(1+\gamma k^{-2})^{-17/6}$ behaves as k^4 for small k and $k^{-5/3}$ for large k.

3.4 Two-dimensional Turbulence

In two dimensions, there is an additional invariant of motion that could contribute to the scaling of the inertial range: the enstrophy $\Omega = (\nabla \times \mathbf{V})^2$. Let us examine the form of the advection operator in three dimensions and in two dimensions in order to see this.

First let us look at the advection of the vorticity $\omega = \nabla \times \mathbf{V}$. Take the curl of the advection operator, one obtains:

$$\nabla \times (\mathbf{V} \cdot \nabla)\mathbf{V} = (\mathbf{V} \cdot \nabla)(\nabla \times \mathbf{V}) - (\nabla \times \mathbf{V} \cdot \nabla)\mathbf{V} \tag{3.29}$$

The first term is just the transport of vorticity, and so conserves vorticity and enstrophy. The second term $-(\omega \cdot \nabla)\mathbf{V}$ corresponds to a rotation of vorticity by the underlying velocity field, and does not necessarily conserve enstrophy (though it does conserve angular momentum). In two dimensions, however, the x and y components of ω are always zero. Because of this, the second term in two dimensions is $-\omega_z \partial_z \mathbf{V}$, which is zero. Therefore, this vortex stretching mechanism does not operate in two dimensions, and enstrophy is conserved.

Once enstrophy is also conserved there are two transport quantities in the inertial range: the transport of energy between scales ϵ, and the transport of enstrophy between scales β. Because of this, one can no longer determine the form of the energy spectrum $E(k)$ from dimensional analysis, because there are multiple ways to construct a dimensionally correct $E(k)$ from these two transports. If, however, there is reason to believe that a given flow's structure is dominated by one transport or the other in a certain range of k, the similarity arguments that give the Kolmogorov-

Obhukov spectrum will still work, but only in that range. For the transport of energy ϵ, the predicted energy spectrum is the same:

$$E_{\epsilon}(k) \propto \epsilon^{2/3} k^{-5/3} \tag{3.30}$$

although the direction of the energy transport is now reversed - from small scales to large[35]. Kraichnan derived the energy spectrum that would result from enstrophy transport[4]:

$$E_{\beta}(k) \propto \beta^{2/3} k^{-3} \tag{3.31}$$

In principle, these spectra could coexist in the same flow, but with each having a range of wavenumber in which it dominates. Spectra of this sort have been observed in atmospheric flows[36], soap film flows[37], and in two-dimensional direct numerical simulation[38; 39].

If we repeat the calculation for the energy range correction in two dimensions, we find that instead of the relation $E(k) \propto k^4$ for small k, the asymptotic behavior of the energy spectrum is $E(k) \propto k^3$. This implies a new energy range correction function for the two-dimensional enstrophy cascade:

$$g(k) = (1 + \gamma k^{-2})^{-3} \tag{3.32}$$

3.5 The Friction Factor

When there is turbulent pipe flow, there is a transport of momentum from the flow to the walls via viscosity and pressure fluctuations. However, if the pipe is of constant cross-sectional area, the mean velocity of an incompressible flow cannot change as one moves down the pipe. Because of this, in order for there to be a steady state flow of non-zero velocity, there must be a pressure drop that adds momentum to the flow to replace what was leaked to the walls.

For a smooth-walled pipe, the longitudinal momentum lost to the walls in a cross section A of the pipe per unit length is entirely due to viscous transport:

$$\frac{dp}{dt} = -l\nu\frac{du}{dy} \tag{3.33}$$

where y is the coordinate perpendicular to the wall and l is the perimeter of the cross section.

The momentum per unit length introduced to the flow due a pressure gradient $\partial_x P$ is:

$$\frac{dp}{dt} = A\partial_x P \tag{3.34}$$

This implies a relationship between the shear at the wall and the pressure drop:

$$\partial_x P = -\frac{l}{A}\nu\frac{du}{dy} \tag{3.35}$$

In order to describe this relationship with dimensionless variables our parameters are the pressure gradient $\partial_x P$, the linear dimension of the pipe cross-section L, the density ρ, and the mean velocity U. We can therefore make a dimensionless number f to describe the friction force due to the flow through the pipe. This definition is only up to a constant of proportionality a that is set by convention:

$$f = -a\frac{L\partial_x P}{\rho U^2} \tag{3.36}$$

The only other dimensionless group we can construct is the Reynolds number, and so we expect that $f = f(\mathrm{Re})$ is a complete description of the scaling of the pressure drop. We can then connect this to the shear $\tau = \nu\frac{du}{dy}$:

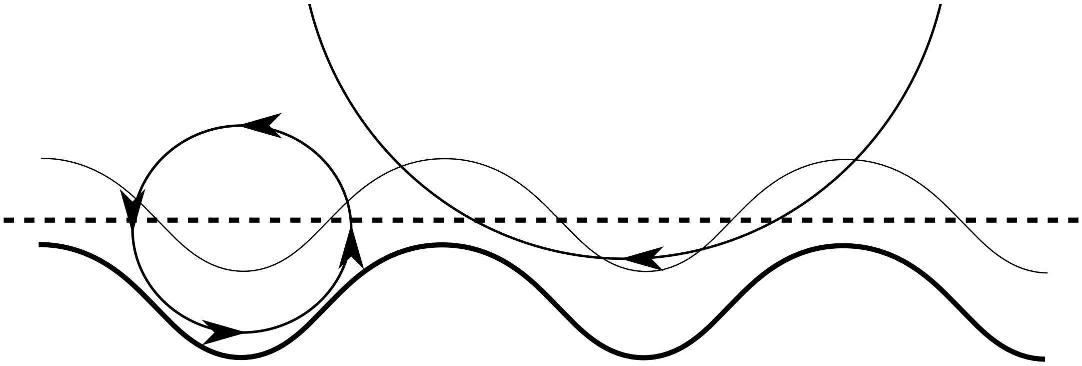

Figure 3.1: Two momentum-carrying vortices near a rough wall. The transport across the dashed line determines the friction factor. In the case of the larger vortex, it intersects the dashed line at a shallow angle, and so the momentum flux is reduced compared to smaller vortices that intersect the dashed line perpendicularly. The scale of the roughness and the boundary layer both determine how large a vortex can be before the intersection can no longer be perpendicular.

$$f = -a\frac{L\partial_x P}{\rho U^2} = a\frac{lL}{A}\tau \tag{3.37}$$

There are two friction factors normally used in three-dimensional flows: the Darcy-Weisbach friction factor, in which $a = 8$[40], and the Fanning friction factor, in which $a = 2$ [20]. Because we will be concerned with the connection between the friction factor and the velocity profile in two-dimensional pipes, it is useful to choose a friction factor definition that trivially connects to the shear τ. For a two-dimensional pipe, the ratio $\frac{l}{A} = 2$ because of the two walls. As such, we will find $a \equiv \frac{1}{2}$ to be a convenient choice of definition for our two-dimensional flows, so that the shear τ may be read off directly. All friction factor data shown in this thesis will use this choice of a.

3.5.1 Momentum Transfer and the Friction Factor

Three Dimensions

The momentum transfer theory of Gioia and Chakraborty[2] can be used to predict the asymptotic scaling of the friction factor. The core assumption is that there is a rate of transfer associated with eddies of a given size, and that the transport of momentum from the main body of the pipe flow into the viscous layer is dominated by a particular scale of eddy. The reason for a single scale to dominate the transport can be seen by looking at the geometric constraints that must be obeyed by an eddy straddling the viscous layer (Fig. 3.1). For large vortices, they intersect the viscous layer at a shallow angle, and so the flux of momentum is reduced by a geometric factor. Below a certain size, any vortex can intersect the viscous layer perpendicular to the layer, maximizing its momentum transfer contribution. However, smaller vortices tend to be less efficient at transporting momentum. The consequence is that there is a vortex size that most strongly transports momentum, set by the scale of the viscous layer η and the roughness r.

We can then look at the characteristic velocity of a vortex of size $s = f(r, \eta)$, where the form of the function f determines cross-over behavior between the fully rough and fully smooth regimes. This function must scale linearly with the larger of its two arguments when that argument is much larger than the other. We will assume for sake of definiteness in the asymptotic analysis that this function is $r + a\eta$, with the understanding that we will incorrectly capture the cross-over regime.

The characteristic velocity can be determined from dimensional analysis. The relevant variables are $E(k)$ and k, from which we obtain $u_s \propto \sqrt{2\pi E(2\pi s^{-1})/s}$[30]. An alternate version of this argument (the one used in the original momentum transfer paper) is to use the total energy contained in all smaller scales as the characteristic energy ($u_s = \sqrt{\int_{2\pi s^{-1}}^{\infty} E(k)dk}$), as opposed to the differential energy at that scale. Both versions give the same asymptotic behaviors, though the intermediate behaviors are somewhat different.

The transport of momentum per unit length of pipe is then $\frac{dp}{dt} = l|u_s|\Delta p$, where Δp is the momentum contrast between the body of the pipe and the wall: $\Delta p = \rho U$. We can use this to immediately evaluate the friction factor:

$$f = f_0 \frac{u_s}{U} \tag{3.38}$$

where f_0 is a dimensionless geometric factor $f_0 = \frac{aL*l}{A}$. For the three-dimensional energy spectrum, this gives us a friction factor:

$$f = f_0 (s/2\pi L)^{1/3} \sqrt{\exp(-\frac{3}{2}\alpha(2\pi\eta/s)^{4/3})(1+\gamma(s/2\pi L)^2)^{-17/6}} \tag{3.39}$$

We can now see the asymptotic scalings of the friction factor in three dimensions. If the pipe is very smooth, then $s \sim \eta = L\mathrm{Re}^{-3/4}$. This gives us the Blasius friction factor scaling[41] at large Re:

$$f(\mathrm{Re}) = f_0 \mathrm{Re}^{-1/4} \tag{3.40}$$

If the pipe is very rough, then $s \sim r$ and the predicted friction factor scaling is the Strickler scaling[42]:

$$f(r) = f_0 (r/L)^{1/3} \tag{3.41}$$

Two Dimensions

In two-dimensional turbulent systems, both the energy cascade or the enstrophy cascade may be observed, or they may occur individually[37] depending on the manner of energy injection and the scale at which it occurs. The two-dimensional inverse cascade friction factor is the same as the case of three-dimensional flows, with a Blasius scaling of $f \propto \mathrm{Re}^{-1/4}$ and a Strickler scaling of $f \propto (r/L)^{1/3}$. The energy spectrum due to the enstrophy cascade leads to a new prediction for the friction factor: a scaling of $f \propto \mathrm{Re}^{-1/2}$ in the Blasius regime and $f \propto (r/R)$ in the Strickler regime. These are the central predictions that we will seek to verify by numerical simulation in the next two chapters.

We can construct a general form of the energy spectrum resulting from an arbitrary conserved quantity, using the same dimensional analysis derivations as we used for the energy and enstrophy cascades. We begin with a conserved quantity with units $[L]^a[T]^b$. The dissipation of this quantity will have units $[L]^a[T]^{b-1}$ and, in the integral range, is related to the pipe geometry by $\kappa = U^{b-1}R^{1+a-b}$. We can therefore construct an energy spectrum:

$$E(k) \propto \kappa^{2\phi/a}k^{2(\phi-1)} \tag{3.42}$$

where $\phi \equiv a/(1-b)$. We also obtain a characteristic lengthscale

$$\eta = (\nu\kappa^{-\phi/a})^{1/(2-\phi)}. \tag{3.43}$$

Combining these results, we obtain that the Blasius friction factor is:

$$f \propto \mathrm{Re}^{-(1-\phi)/(2-\phi)} \tag{3.44}$$

and the Strickler friction factor is:

$$f \propto (r/L)^{1-\phi} \tag{3.45}$$

For the case of the 3D or inverse energy cascade, $\phi = 2/3$. For the enstrophy cascade, $\phi = 0$.

There are some concerns with this derivation which we will address. Can we say that the energy spectrum near the wall is really the same as the isotropic, homogeneous turbulence result that we derived before? This has been addressed in measurements of the near-wall structure functions of turbulence. While the anisotropy due to the presence of the wall has a strong effect on high-order moments, the energy spectrum corresponds to the second order moment, which is not significantly altered from its behavior at the center of the pipe[43].

Another concern is that we have treated the turbulent state itself as its average over the pipe cross-section, but we are interested in transport occuring around a layer near the wall. If, for example, ϵ depended on y, then this could result in a correction factor due to the difference between the local ϵ and the global ϵ. We will address this in the next section, where we will use the momentum transfer argument to derive the flow velocity profile.

A third concern is the determination of ϵ by dimensional analysis: $\epsilon \propto U^3/L$. There is the possibility that ϵ may scale with a dimensionless number such as the Reynolds number. If we consider the energy budget of the flow, the fluid is gaining an amount of energy per unit volume $dE/dt = -U\partial_x P$ due to traversing the pressure gradient. If the flow is in a steady state, then this energy must somehow be removed from the system for the total energy to remain constant. Because the velocity profile approaches a linear profile at the wall, there is no transport of energy to the wall - all energy lost from the flow must be lost due to viscous dissipation. The viscous dissipation can be separated into dissipation from the time-independent structure of the velocity profile (ϵ_0) and dissipation due to the turbulent cascade (ϵ). This implies that $\epsilon_0 + \epsilon = dE/dt = (U^3/L)(f/a)$.

This implies an upper bound on the the turbulent energy dissipation: $\epsilon \leq \frac{U^3}{L}\frac{f(\mathrm{Re})}{a}$. Because the momentum transfer calculation of the Blasius friction factor scaling has assumed a constant turbulent ϵ of the form $c\frac{U^3}{L}$, then this constant turbulent ϵ will comprise the entirety of the energy dissipation when $f = ca$. In the momentum transfer theory, $f \sim \mathrm{Re}^{-1/4}$ in 3D under the assumption of a Reynolds number independent turbulent ϵ. As such, f is predicted to continue to decrease towards zero in the limit of a smooth pipe, and so the predictions of the theory are no longer self consistent beyond some finite Reynolds number at sufficiently small roughness (such that $f(r, \mathrm{Re} \to \infty) \leq ca$, schematically depicted in Fig. 3.2), and we expect that the Blasius scaling is not the true asymptotic behavior when $\mathrm{Re} \to \infty$.

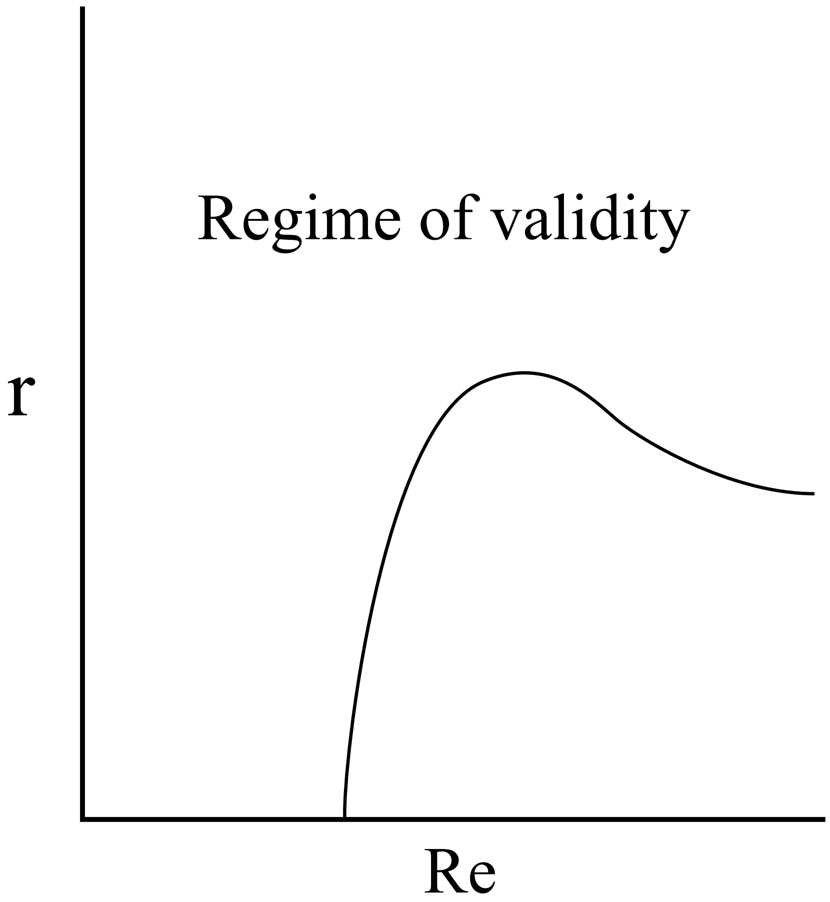

Figure 3.2: This figure schematically depicts the regime of validity of the momentum transfer theory given the shape of the friction factor data of Nikuradse, under the constraint of energy balance. At small Re and r, the energy input to the flow by the pressure drop is too small to account for the energy spectrum that the momentum transfer theory assumes.

3.6 Velocity Profile

3.6.1 Classical Theory

The classical theory of the velocity profile in three-dimensional pipe flows[44] is the basis upon which the modern understanding of the mean behavior of wall-bounded flows has been built. It has been used to derive the friction factor of turbulent pipe flows, to determine the profiles of energy and Reynolds stress, and as a basis for asymptotic matching between the deviation of the velocity profile near the viscous layer and the wake region of the flow[45]. Colebrook and White extended the results of Prandtl to take into account pipes with rough surfaces, and the Colebrook-White equation[46] sees extensive use in engineering in predicting the coefficient of friction in many flow applications.

Despite its successes, the classical theory makes no connection with the unsteady structure of turbulent flows. Turbulent flows consist of a collection of structures at many scales, which all participate in the transport properties of the flow. These structures are known to accelerate the process of mixing[47], which suggests that these structures should be highly important in the transport of momentum and establishment of the structure of the shear near the wall and throughout the pipe.

We will use the connection between the friction factor and velocity profile to shed light on the importance of the self-similar structure of the turbulent flow in the context of the momentum transfer theory[2], which predicts a spectral dependence of the friction factor scaling. This implies that the velocity profile should also depend on the spectrum of the turbulent flow. However, in three-dimensional incompressible flows there is only a forward energy cascade which scales as $k^{-5/3}$[9; 10]. By looking at the case of two-dimensional turbulence, in which there is both a forward enstrophy cascade and an inverse energy cascade[3; 4], we can test the predictions of a spectrally dependent velocity profile. We should expect to find that the velocity profile in the two-dimensional enstrophy cascade should be steeper than three-dimensional velocity profiles, leading to a decrease in the von Kármán constant. Furthermore, the theory predicts a power-law velocity profile with an exponent 1/3, meaning that the von Kármán constant will appear to have a Reynolds number dependence as the size of the turbulent boundary layer increases.

Let us first review the classical theory. We can determine an equation for the balance of momentum transport in a layer of fluid parallel to the wall of an infinite pipe by averaging the Navier-Stokes equation over time and parallel coordinates x (in the streamwise direction) and z (transverse to the flow direction):

$$\langle \frac{d\mathbf{V}}{dt} \rangle + \langle \mathbf{V} \cdot \nabla \mathbf{V} \rangle = \nu \nabla^2 \langle \mathbf{V} \rangle - \langle \nabla \mathbf{P} \rangle \tag{3.46}$$

The average of the derivative of a quantity with respect to a coordinate (time or space) over that coordinate is the difference in the value of that quantity between the end and beginning of the pipe divided by the length of the pipe. For all quantities other than the pressure, this is zero. For the pressure in the streamwise direction, this is the pressure drop per unit length $-fU^2$ related to the friction velocity $u_* = \bar{U}\sqrt{f}$. Furthermore, the time-averaged pressure over a cross-section of the pipe must be constant or the flow would be preferentially driven towards or away from a point on the boundary (which is forbidden by the symmetry of the system). Additionally, the time average of the component of the velocity perpendicular to the fluid layer must be zero for similar reasons of symmetry. Let us now consider only the component of the velocity in the direction of the flow u:

$$\langle \mathbf{V} \cdot \nabla u \rangle = \nu \partial_y^2 \langle u \rangle + u_*^2 \tag{3.47}$$

Because the fluid is divergence free we can write the advection operator as $\partial_x(u^2) + \partial_y(uv) + \partial_z(uw)$. The first and third components of this must average to zero - in the case of the first component, because the velocity profile is the same at the start and end of the pipe, and in the second because the average involves an integral around the perimeter of the pipe, and so the beginning and ending points are the same. Therefore we can express this relation:

$$\partial_y(\nu \partial_y \langle u \rangle - \langle uv \rangle + u_*^2) = 0 \tag{3.48}$$

Integrating this towards the wall from the center of the pipe where we know the boundary condition that the flux must be zero by symmetry, we obtain the exact averaged momentum transport balance:

$$\nu\partial_y\langle u\rangle + \langle uv\rangle = u_*^2(1-y/R) \tag{3.49}$$

In the turbulent boundary layer, the wake term $1-y/R$ is asymptotically constant, and the viscous transport $\nu\partial_y u$ is much smaller than the Reynolds stress. To proceed further, it is necessary to impose a closure condition that relates the Reynolds stress to known quantities. In the Prandtl closure, the Reynolds stress is written in the form $\kappa u_* y\partial_y u$, where κ is the von Kármán constant[48; 49] from dimensional considerations involving only the near-wall scales. This gives rise to a logarithmic velocity profile and the Prandtl friction factor scaling:

$$\partial_y u/u_* = \frac{1}{\kappa y} \tag{3.50}$$

$$u = (u_*/\kappa)\ln(y/y_*) \tag{3.51}$$

In order to exclude the divergence at $y = 0$ we must impose a cutoff. This is due to the previously neglected viscous contribution becoming large at a scale y_* where $\nu \sim u_* y_*$. Because of this, the lower limit of integration is $y_* = \nu/u_*$.

The Prandtl friction factor form results from integrating this velocity profile over the turbulent boundary layer, and relating that to the mean velocity of the flow:

$$\frac{1}{\sqrt{f}} = \frac{1}{\kappa}\ln(\mathrm{Re}\sqrt{f}) + B \tag{3.52}$$

The expression of the Reynolds stress in terms of near-wall variables is an assumption of complete similarity of the turbulent boundary layer with respect to Reynolds number. This is equivalent to assuming that any intrinsic scales of the turbulent flow do not contribute to the problem of the turbulent boundary layer other than the distance from the wall and the shear at the wall. If, however, the structure of turbulence does contribute directly, then there is an undetermined function in the classical theory which modifies the differential equation for the velocity profile.

$$\partial_y u/u_* = \frac{1}{\kappa y} g(y/y_*), y_* = R/\mathrm{Re}\sqrt{f} \tag{3.53}$$

Here we have allowed the viscous lengthscale to contribute to the scaling of the Reynolds shear away from the boundary. In the usual picture, the function $g(x)$ asymptotically approaches a constant as its argument approaches zero. However, if the zero-viscosity limit of turbulence is a singular limit rather than a regular one then $g(x)$ could continue to scale. In this case one would expect the scaling of the viscous lengthscale to determine the characteristics of the flow even as it approaches zero. There is evidence in the form of observed data collapse[1; 17] that there is a critical point at infinite Reynolds number and zero roughness, which would suggest that one cannot safely neglect the scaling of $g(x)$ in determining the form of the velocity profile.

3.6.2 Connection to the Friction Factor

If the scaling of $g(x)$ is not neglected, there is insufficient information to determine the friction factor and velocity profile without further inspection of the physics of the flow. It is however possible to determine one if the other is known. The procedure for finding the friction factor for a given velocity profile is the same as in the Prandtl friction factor case discussed above. We will consider the reverse of that procedure in order to convert observations of the friction factor into predictions of the form of the velocity profile.

We start with the velocity profile:

$$u(y) = u_*/\kappa \int_{y_*}^{y} \frac{1}{x} g(\frac{x}{y_*}) dx \tag{3.54}$$

From here we can substitute $z = x/y_*$ to find that $u(y) = u_*/\kappa \int_1^{y/y_*} \frac{1}{z} g(z) dz$. The important piece of information here is that u can be expressed as a function of y/y_*: $u(y) = u_* G(y/y_*)$. We have absorbed the constant κ into the function $G(y/y_*)$.

Then we can relate this unknown velocity profile to the mean velocity in order to get a relation for the friction factor:

$$\frac{1}{\sqrt{f}} = \frac{1}{R}\int_{y_*}^{R} G(\frac{y}{y_*})dy \tag{3.55}$$

With a change of variables $z = y/y_*$:

$$\frac{1}{\sqrt{f}} = \frac{y_*}{R}\int_{1}^{R/y_*} G(z)dz \tag{3.56}$$

$$\mathrm{Re} = \int_{1}^{\mathrm{Re}\sqrt{f}} G(z)dz \tag{3.57}$$

It is convenient now to think of Re as a function of $\alpha \equiv \mathrm{Re}\sqrt{f}$ so that we may differentiate both sides with respect to $\mathrm{Re}\sqrt{f}$ and extract $G(z)$. We then have:

$$G(\alpha) = \frac{d\mathrm{Re}}{d\alpha} = \frac{1}{\sqrt{f} + \mathrm{Re}/(2\sqrt{f})\frac{df}{d\mathrm{Re}}} \tag{3.58}$$

To make use of this, it is necessary to determine $f(\alpha)$ and $\mathrm{Re}(\alpha)$, and so it is necessary to invert $\alpha = \mathrm{Re}\sqrt{f}$. This imposes certain constraints on the friction factor - namely, $\mathrm{Re}\sqrt{f}$ must be monotonic over the full range of y/y_* needed for a particular pipe's geometry for this procedure to work. Furthermore, $\sqrt{f} + \mathrm{Re}\partial_{\mathrm{Re}} f/(2\sqrt{f})$ must always be greater than zero. These constraints can fail during the drag catastrophe, where the hysteresis of the transition to turbulence allows non-monotonicity in $\mathrm{Re}\sqrt{f}$ (Fig. 3.3). Similar to the case of the liquid-gas transition, a single monotonic curve must be constructed to connect those cases for it to be possible to relate the viscous layer to the rest of the velocity profile. One possibility is to combine the two friction factors with a smooth interpolating function around the point where $f_{laminar} = f_{turbulent}$, resulting in a smooth and monotonic f. Errors due to the interpolation will be localized around the transition between the viscous layer and turbulent boundary layer.

If we have a way to predict the friction factor, we can then proceed to predict the velocity profile. We will use the momentum transfer theory of Gioia and Chakraborty[2], which connects the transport of momentum from the mean flow to the wall to the energy spectrum of turbulence. This is done by modelling the pipe as two layers - the central flow, and the layer in which momentum is coupled to the wall via roughness or viscosity. From the momentum transfer theory we can make predictions for the form of the friction factor in 2D and 3D[50], and as such predict the velocity profile of 2D turbulent flows.

The asymptotic Blasius behavior ($f \sim \mathrm{Re}^{-1/4}$ in the energy cascade and $f \sim \mathrm{Re}^{-1/2}$ in the enstrophy cascade) predicted by the momentum transfer model leads to power-law velocity profiles with exponents 1/7 and 1/3 respectively. The $y^{1/3}$ power law is a novel result, and compares well with numerical simulations of 2D flows[50]. In Fig. 3.4, we show the velocity profile from simulation and calculated from the two-dimensional friction factor.

3.6.3 From Momentum Transfer (3D)

Let us now consider directly how we may derive the velocity profile of three-dimensional turbulence from the momentum transfer theory. We will begin with the Navier-Stokes equation and derive the equations of momentum and energy balance. We will then make a closure using the momentum transfer theory to determine the Reynolds stress $-\langle uv \rangle$ from the energy spectrum. The equation for energy balance determines the relation between the local energy dissipation ϵ and the structure of the velocity profile. This lets us determine how the energy spectrum depends on distance from the wall.

We begin with the Navier-Stokes equation:

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\vec{u} = \nu \nabla^2 \mathbf{u} - \nabla P \qquad (3.59)$$

To derive the energy balance, we take the dot product with $\mathbf{u}$ and convert terms of the form $u\partial_x u$ into $\partial_x u^2/2$.

$$\frac{\partial E}{\partial t} + \mathbf{u}\cdot\mathbf{u}\cdot\nabla\mathbf{u} = \nu\mathbf{u}\cdot\nabla^2\mathbf{u} - \mathbf{u}\cdot\nabla P \tag{3.60}$$

We now write $\mathbf{u}\nabla^2\mathbf{u}$ as $\nabla^2|\mathbf{u}|^2/2 - |\nabla\mathbf{u}|^2$. This second term is the local energy dissipation ϵ. We can rewrite the pressure term as: $\nabla\cdot P\mathbf{u} - P\nabla\cdot\mathbf{u}$. For incompressible flow, the second term is zero.

$$\frac{\partial E}{\partial t} + \mathbf{u}\cdot\mathbf{u}\cdot\nabla\mathbf{u} = \nu\nabla^2 E - \mathbf{u}\cdot\nabla P - \epsilon \tag{3.61}$$

Averaging over time eliminates the time derivative of the energy. We can further split $\mathbf{u}$ into its average value and its fluctuations: $\mathbf{u} = \bar{u} + u'$. From here we will use the component representation of $\mathbf{u}$ as (u, v, w).

All terms that are first order in the fluctuations will average to zero. We will assume that all terms of third order in the fluctuations are very small compared to terms of second order. Also, in a pipe flow situation we know that there is no net flow on average towards or away from the walls, so $\bar{v}$ and $\bar{w}$ are both zero, and that the fully developed flow does not depend on x so that ∂_x of any averaged quantity except pressure will be zero. The surviving terms are:

$$\begin{aligned} u'v'\partial_y\bar{u} + \bar{u}\partial_y u'v' + u'w'\partial_z\bar{u} + \bar{u}\partial_z u'w' &= \\ \frac{1}{2}\nu\partial_y^2\bar{u}^2 + \frac{1}{2}\nu\partial_y^2(u'^2 + v'^2 + w'^2) - \epsilon & \\ -\bar{u}\partial_x\bar{P} - \partial_y(P'v') - \partial_z(P'w') & \end{aligned} \tag{3.62}$$

If we are in two dimensions, or are near the wall, we can average over the coordinate parallel to the wall (z):

$$\begin{aligned}(-u'v')\partial_y\bar{u} + \bar{u}\partial_y(-u'v') + \\ \frac{1}{2}\nu\partial_y^2\bar{u}^2 + \frac{1}{2}\nu\partial_y^2(u'^2 + v'^2 + w'^2) = \\ \epsilon + \bar{u}\partial_x\bar{P} + \partial_y(P'v')\end{aligned} \tag{3.63}$$

We can use the momentum balance equation in the x direction to say that $\partial_y u'v' = \nu\partial_y^2\bar{u} - \partial_x\bar{P}$:

$$\begin{aligned}(-u'v')\partial_y\bar{u} + \bar{u}\partial_x\bar{P} - \nu\bar{u}\partial_y^2\bar{u} + \\ \frac{1}{2}\nu\partial_y^2\bar{u}^2 + \frac{1}{2}\nu\partial_y^2(u'^2 + v'^2 + w'^2) = \\ \epsilon + \bar{u}\partial_x\bar{P} + \partial_y(P'v')\end{aligned} \tag{3.64}$$

So, using $\bar{u}\partial_y^2\bar{u} = \partial_y^2\bar{u}^2/2 - (\partial_y\bar{u})^2$:

$$(-u'v')\partial_y\bar{u} + \nu(\partial_y\bar{u})^2 + \nu\partial_y^2 E' = \epsilon + \partial_y(P'v') \tag{3.65}$$

Let us now use the fact that we know the total shear from global momentum balance: $\nu\partial_y\bar{u}\rangle - u'v' = fU^2(1 - y/R)$:

$$fU^2(1 - y/R)\partial_y\bar{u} = \epsilon - \nu\partial_y^2 E' + \partial_y(P'v') \tag{3.66}$$

It has been empirically observed in three-dimensional flows that in the scaling range of the velocity profile, the final two terms in Eq. 3.66 balance each other in the turbulent boundary layer[20]. If we assume that this balance holds, we arrive at:

$$\epsilon = fU^2(1 - y/R)\partial_y\bar{u} \tag{3.67}$$

We will now use the momentum transfer theory to develop a closure for $-\langle uv\rangle$ in the momentum equation. In the global momentum transfer theory, we posited that transport of momentum arose from a dominant eddy with size s. The transport of momentum was therefore $\rho u_s U$, where ρU is the momentum source (and the momentum sink is at zero). Let us consider the net transport of momentum across some arbitrary boundary, where the momentum sink is not necessarily at zero.

If we have an eddy spinning in a uniform background field, we expect there to be no momentum transport, and so the momentum source term cannot simply be the local momentum. It must instead be a momentum contrast, involving spatial derivatives of the local momentum. The lowest-order term that satisfies this symmetry is $l\partial_y(\rho u)$, where l is some length-scale. If we consider this term to be a momentum contrast, it makes sense that the scale l is the same as the size of the eddy s, as the term is therefore the difference between the momentum at the upper edge of the eddy and the momentum at the lower edge of the eddy.

When we were transporting momentum to the viscous layer, the dominant eddy size was limited by the scale of the viscous layer. However, at some arbitrary distance y from the wall, eddies of size y and smaller can participate in the transport. Out of those eddies, the largest possible eddy will have the highest contribution to the transport, and so we expect that the size of the dominant eddy is proportional to y. This is the Townsend similarity hypothesis[51] and is supported by experiments in three-dimensional pipe flows[14].

We therefore posit the closure for the Reynolds stress: $-\langle uv\rangle = u_y y\partial_y\bar{u}$, where $u_y = \sqrt{\int_{2\pi y^{-1}}^{\infty} E(k)dk}$.

We combine this closure with our local ϵ to derive the velocity profile.

$$\nu\partial_y\bar{u} + (fU^2(1-y/R)\partial_y\bar{u})^{1/3}\sqrt{\int_{2\pi y^{-1}}^{\infty} k^{-5/3}f(k\eta)g(kI)}y\partial_y\bar{u} = fU^2(1-y/R) \quad (3.68)$$

In the limit of infinite Reynolds number, the spectral integral becomes $y^{2/3}$ and terms proportional to ν limit to zero. This reduces the differential equation to:

$$C(\partial_y\bar{u})^4 s^4 = f^2U^4 \quad (3.69)$$

where C is a combination of numerical factors. This gives us exactly the Prandtl differential equation for the velocity profile:

$$(\partial_y \bar{u}) = C u_* / y \tag{3.70}$$

The results of numerically integrating the full differential equation are shown in Fig. 3.5. The viscous layer, log-layer, and wake are all present. Additionally, only when the full energy spectrum including the integral scale corrections is considered does the wake overshoot the log law as seen in the experimentally observed velocity profiles.

3.6.4 From Momentum Transfer (2D)

Let us now perform the previous calculation in the case of the enstrophy cascade. It is necessary to determine the equivalent of the energy equation for the enstrophy, so that we can evaluate the dependence of β on the local profile.

We start from the time-averaged Navier-Stokes in the vorticity formulation (by taking the curl of the usual Navier-Stokes equation):

$$u \cdot \nabla \omega = \nu \nabla^2 \omega \tag{3.71}$$

In order to get an equation for the enstrophy, we multiply through by the vorticity ω:

$$u \partial_x \Omega + v \partial_y \Omega = \nu \nabla^2 \Omega - \nu |\nabla \omega|^2 \tag{3.72}$$

If we discard terms that will average to zero when we take the average over the pipe length x, we obtain

$$\partial_y \langle v \Omega \rangle = \nu \partial_y^2 \langle \Omega \rangle - \beta \tag{3.73}$$

where $\beta = \nu((\partial_x \omega)^2 + (\partial_y \omega)^2)$.

We can rewrite v and Ω in terms of fluctuating and constant parts:

$$\Omega = (\partial_y u + \partial_y u' - \partial_x v')^2 \tag{3.74}$$

$$v\Omega = v'(\partial_y u)^2 + v'(\partial_y u')^2 + v'(\partial_x v')^2 + 2v'\partial_y u'\partial_y u - 2v'\partial_y u \partial_x v' - 2v'\partial_y u'\partial_x v' \tag{3.75}$$

If we assume that all terms of order δ^3 can be neglected, and average over x, this becomes:

$$v\Omega = 2(\partial_y u'v')\partial_y u \tag{3.76}$$

We now have:

$$2\partial_y(\partial_y \langle u'v'\rangle \partial_y \langle u\rangle) = \nu \partial_y^2 \langle \Omega \rangle - \beta \tag{3.77}$$

We know from the momentum balance that $\langle u'v'\rangle = \nu \partial_y \langle u \rangle - fU^2(1 - y/R)$ and so in the infinite Re limit:

$$\beta = -2fU^2/R\partial_y^2 \langle u \rangle \tag{3.78}$$

We can then use the momentum transfer closure for $\langle uv \rangle$:

$$y u_y \partial_y u = fU^2(1 - y/R) \tag{3.79}$$

where $u_y = \sqrt{\int_{y^{-1}}^{\infty} E(k)dk}$.

If we use the enstrophy cascade energy spectrum (Eq. 3.31, 3.32) with an exponential dissipation, we can evaluate u_y:

$$u_y = (-2fU^2/R\partial_y^2 u)^{1/3} \sqrt{\int_{y^{-1}}^{\infty} k^{-3} \exp(-\beta \eta k)(\frac{1}{1 + \gamma k^2})^3} \tag{3.80}$$

If we neglect the dissipative and large-scale parts of the spectrum, and the influence of the wake, this becomes: $u_y = \frac{y}{\sqrt{2}}(-2fU^2/R\partial_y^2 u)^{1/3}$. We then obtain the differential equation for the velocity profile in the infinite Re limit:

$$\frac{1}{\sqrt{2}}y^2(-2fU^2/R\partial_y^2 u)^{1/3}\partial_y u = fU^2 \tag{3.81}$$

Making the substitution $v = \partial_y u$ and cubing, this becomes:

$$\partial_y v^4 = -4\sqrt{2}f^2U^4Ry^{-6} \tag{3.82}$$

Integrating towards the wall from the center of the pipe where $\partial_y u = 0$, we obtain:

$$\partial_y u = \frac{2^{5/8}}{5^{1/4}}\sqrt{f}UR^{1/4}(y^{-5} - R^{-5})^{1/4} \tag{3.83}$$

Away from the center of the pipe, we therefore have:

$$u(y)/U = U_0/U - 2^{21/8}5^{-1/4}\sqrt{f}(y/R)^{-1/4} \tag{3.84}$$

This solution has the property that it goes to zero at some positive value of y determined by U_0. If we attempt to match this solution to the viscous layer, we should have $u(\nu/U\sqrt{f}) = U\sqrt{f}$:

$$U_0/U = \sqrt{f}(1 + 2^{21/8}5^{-1/4}\mathrm{Re}^{1/4}f^{1/8}) \tag{3.85}$$

Let us now use this profile to compute the friction factor, to see if the result is consistent with the original momentum transfer theory. To do this, we integrate the velocity profile to determine the mean velocity. This then gives us an equation for the friction factor:

$$(1 + 2^{21/8}5^{-1/4}\mathrm{Re}^{1/4}f^{1/8}) - 2^{37/8}5^{-1/4}3^{-1}(1 - (\mathrm{Re}\sqrt{f})^{-3/4}) = 1/\sqrt{f} \tag{3.86}$$

Let us write this in terms of $K \equiv \mathrm{Re}\sqrt{f}$:

$$C_2 K^{1/4}(1 - C_2)K + C_1 K^{5/4} = \mathrm{Re}, \tag{3.87}$$

where $C_1 \equiv 2^{21/8}5^{-1/4}$ and $C_2 \equiv 2^{37/8}5^{-1/4}3^{-1}$. If we consider the large-Re limit as we have been doing, the dominant terms are Re and $C_1K^{5/4}$. This tells us that in the large Re limit:

$$f = \mathrm{Re}^{2/5} C_1^{-8/5} \tag{3.88}$$

This is somewhat inconsistent with the previous prediction of $f \propto \mathrm{Re}^{-1/2}$, but not by much. If we examine the friction factor curve obtained by solving the full form of Eq. 3.87, we obtain a significant region of $\mathrm{Re}^{-1/2}$ scaling at moderate Re(Fig. 3.6). Why is this? Note that if we discard the $K^{1/4}$ term, it does not significantly change the shape of the resultant curve at turbulent Reynolds numbers. We can then attempt to solve the equation formed by the remaining terms:

$$C_1^4 K^5 - A_4 K^4 + A_3 \mathrm{Re} K^3 - A_2 \mathrm{Re}^2 K^2 + A_1 \mathrm{Re}^3 K - \mathrm{Re}^4 = 0 \tag{3.89}$$

Where the A's are positive combinations of C_1 and C_2 whose specific form is not germane to this analysis. If we change variables to $u = \sqrt{f}$, this conveniently becomes:

$$\mathrm{Re} u^5 - A_4 u^4 + A_3 u^3 - A_2 u^2 + A_1 u - 1 = 0 \tag{3.90}$$

Note that if $u \sim \mathrm{Re}^{-1/5}$ (as predicted earlier), then the largest terms at large Re are $\mathrm{Re}u^5$ and 1, reproducing our previous limit consistently. If one now considers the case where $u \sim \mathrm{Re}^{-1/5+\epsilon}$, then the next point at which two terms become of the same order in Re is when $\epsilon = 1/20$, corresponding to $f \sim \mathrm{Re}^{-1/2}$. In this case, the terms $A_1 u$ and $\mathrm{Re}u^5$ are of the same order in Re, and balancing them gives us

$f = A_1^{1/2}\mathrm{Re}^{-1/2}$, which is the intermediate behavior that we observe in the Blasius regime. The next crossing of orders is when $f \sim \mathrm{Re}^{-1}$. By this time, the $K^{1/4}$ term that we discarded is likely to become relevant, and so this behavior is not actually observed in the computed solution to the full equation. This does however predict that we should expect a steeper friction factor scaling when leaving the drag catastrophe than the eventual Blasius result.

We have shown that in the case of the enstrophy cascade, the Blasius scaling can be seen as an intermediate Re result of the velocity profile derived from the momentum transfer closure. This prediction also anticipates curvature of the friction factor away from the Blasius scaling at large Re, in analogy to what is seen in the three-dimensional case. It is somewhat mysterious that the intermediate scaling is accessible directly in the enstrophy cascade, and yet is elusive in the inverse energy cascade where the resulting equation for the friction factor is transcendental. However, in both cases the simpler form of the momentum transfer theory seems to be able to predict the intermediate (but not the eventual) scaling ranges.

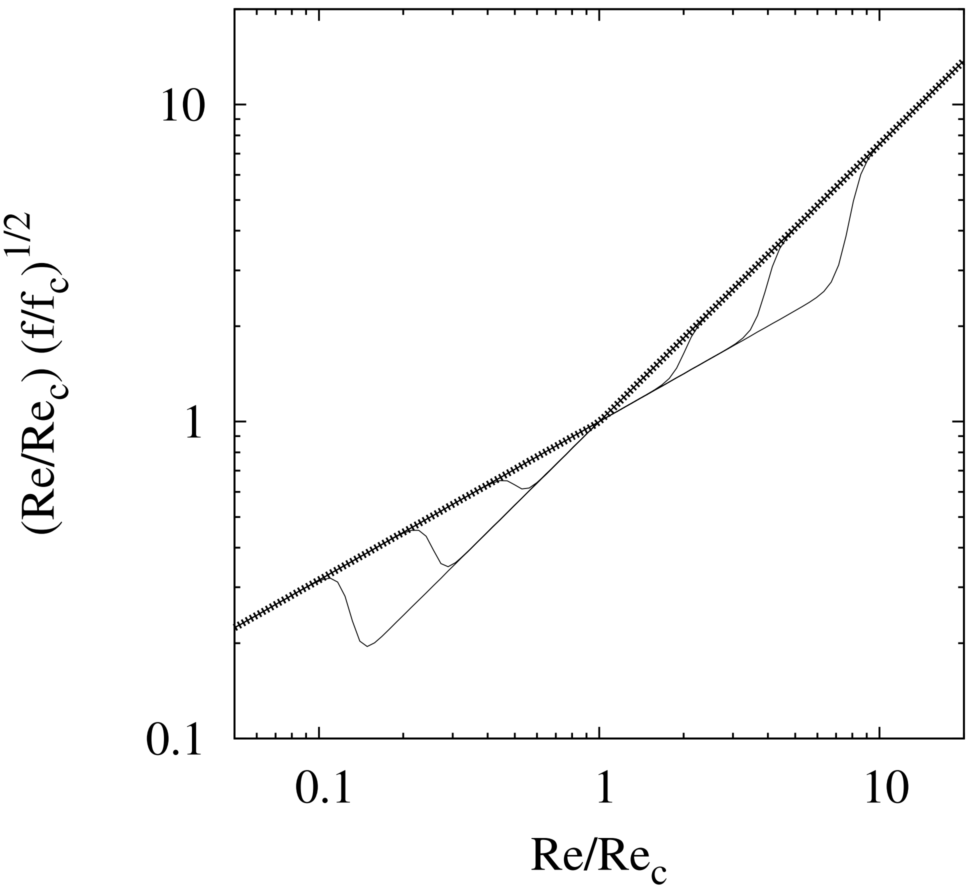

Figure 3.3: Non-monotonicity and extreme derivatives in $\mathrm{Re}\sqrt{f}$ around the drag catastrophe. The light lines are schematic representations of possible trajectories of the friction factor due to hysteresis around the drag catastrophe. Several of these are non-monotonic and thus make inverting $\mathrm{Re}\sqrt{f}$ impossible. The bold line is an interpolating function that bounds the derivative and avoids non-monotonicity.

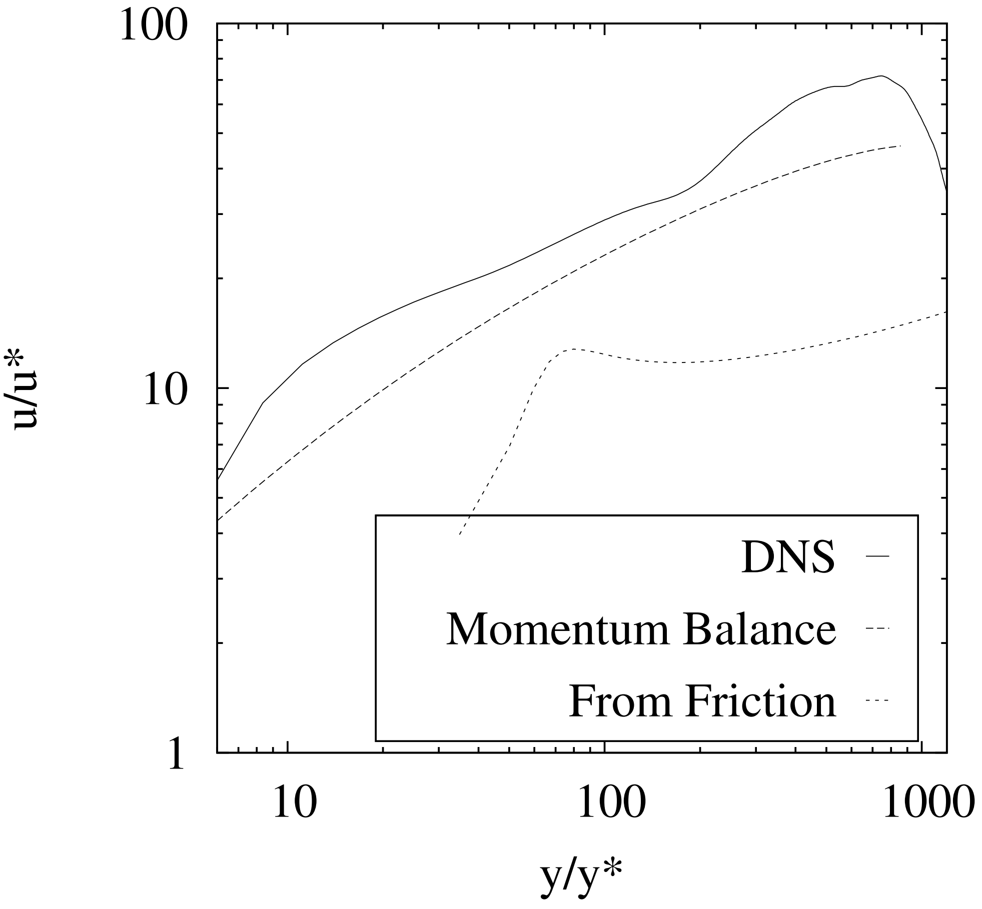

Figure 3.4: Velocity profiles measured from two-dimensional direct numerical simulation, predicted from the momentum balance, and derived from friction factor measurements from the same simulation. The Reynolds number for these velocity profiles is 70000. In the case of the friction factor derived results, an interpolating function $1/(c_1\mathrm{Re}^{c_2}+c_3)+1/(c_4\mathrm{Re}^{c_5}+c_6)$ was fit to the measurements to prevent scatter from dominating the derivative.

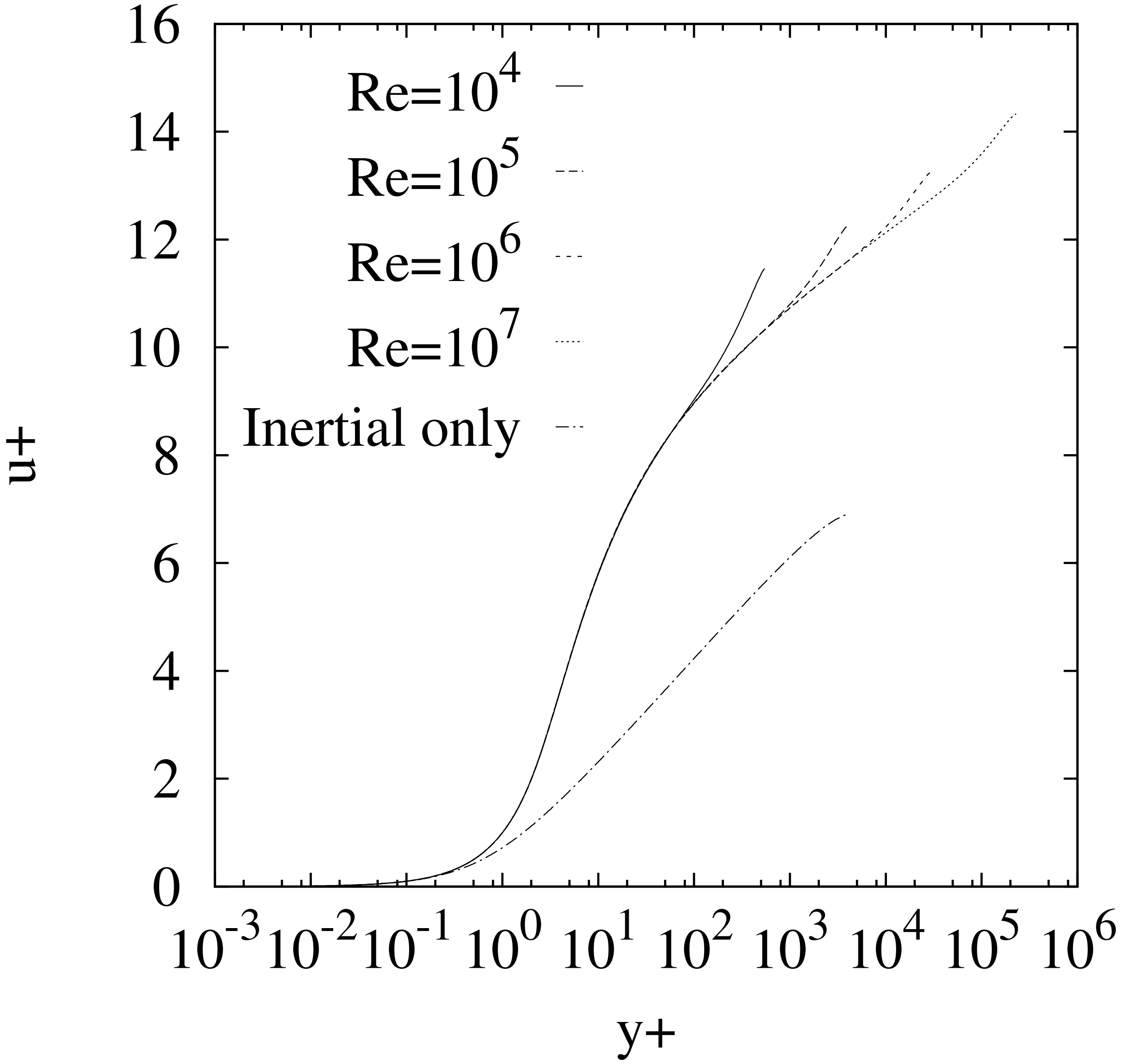

Figure 3.5: Velocity profiles derived from the local momentum transfer theory for three-dimensional pipe flow calculated with the inertial range only, and with the full energy spectrum.

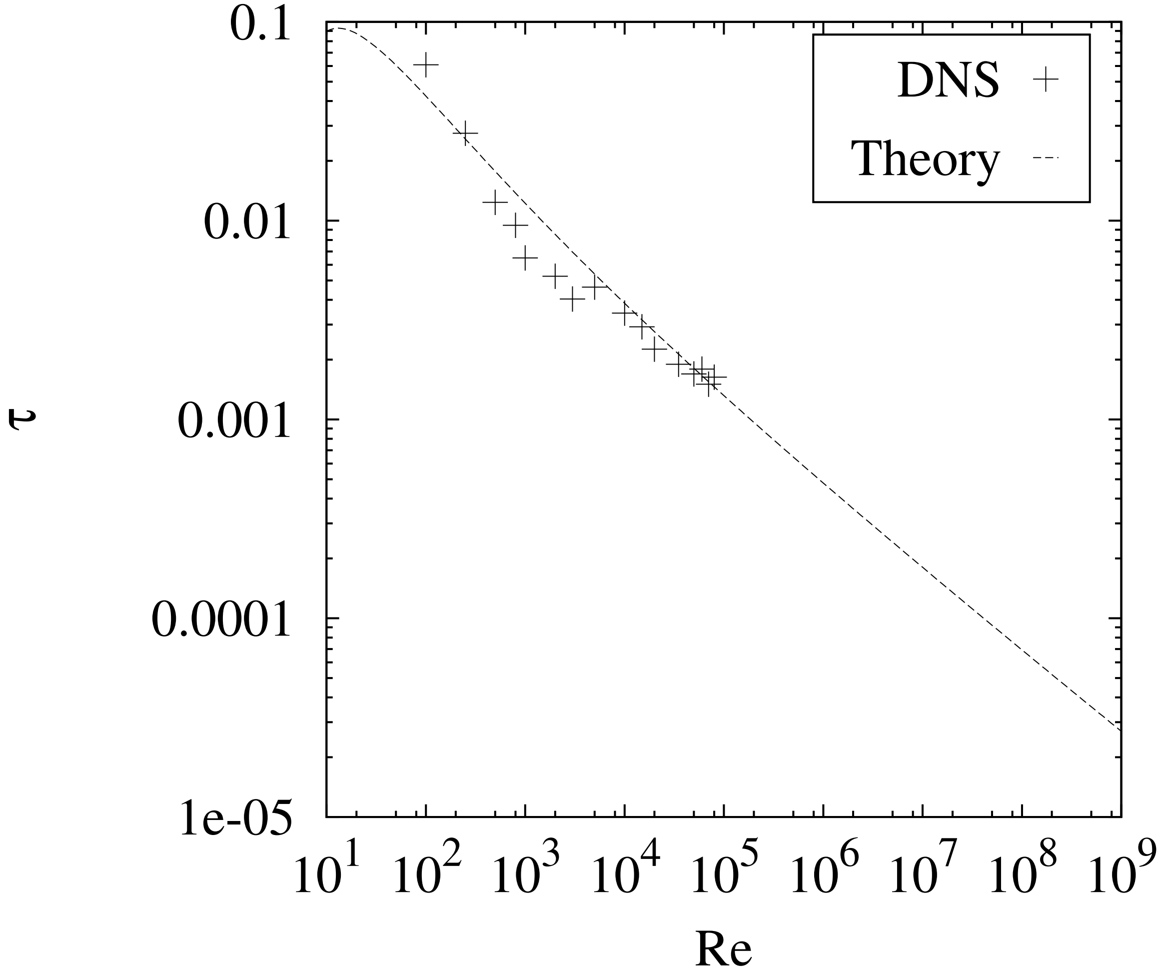

Figure 3.6: This figure shows the two-dimensional enstrophy cascade friction factor predicted from the velocity profile theory. At moderate Re, there is a large region of $\mathrm{Re}^{-1/2}$ behavior. At large Re, this curves upwards and becomes $\mathrm{Re}^{-2/5}$. Plotted for comparison are data from our direct numerical simulations of 2D enstrophy cascade pipe flow turbulence (described in Chapters 4 and 5).

Chapter 4

The Transition to Turbulence

The fundamental nature and stability of the turbulent state of fluids remains an open and challenging question. Fluid flow is characterized by a dimensionless number Re, which depends on the characteristic length L, velocity U and kinematic viscosity of the fluid ν through the relation $\mathrm{Re} \equiv UL/\nu$. As the Reynolds number increases from zero, the flow becomes increasingly structured and eventually statistical in nature, so that at large Re, the flow is said to be turbulent[52]. The conventional assumption—that the turbulent state is absolutely stable—has been challenged recently by a series of theoretical[53; 54] and experimental probes[55–60] of the transition to turbulence. Taken as a whole, these works suggest that turbulence might, in some flow regimes at least, be a long-lived metastable state[25–27; 53; 61]. Such a view would be consistent with the fact that long-lived transient turbulent states can be excited as finite-amplitude instabilities of the laminar state, so that the laminar and turbulent states can coexist (for a review of foundational work in this area, see e.g. Ref. [28]; recent developments are summarized in Refs. [25–27]). However, the question remains as to whether the turbulent state is ever sustainable with an infinite lifetime for finite Reynolds numbers. This is a difficult experimental question to decide, because the lifetime of the turbulent state can become so long that accurate measurements become impossible. With the necessary restriction to a small range of Reynolds numbers, the data have, until recently, been difficult to interpret in a compelling way.

In a set of elegant and remarkably accurate experiments on transitional pipe turbulence[59], Hof et al. have brought into question the idea that pipe flow turbulence is stable at long times beyond a finite critical Reynolds number[54; 62]. The laminar state of a straight smooth pipe flow is linearly stable at all Reynolds numbers (see e.g. Ref. ([63])), but a sufficiently large perturbation triggers localized turbulent puffs that persist for long times. The decay of the transient turbulent state is reported to follow a Poisson distribution, with a lifetime $\tau(\mathrm{Re})$ that increases sharply with increasing Reynolds number. The new measurements of the lifetime of these localized puffs[59] reveal that $\tau(\mathrm{Re})$ apparently only diverges at infinite Reynolds number, scaling in a super-exponential way with Re. Similar observations in another linearly stable flow—Taylor-Couette flow with outer cylinder rotation—have recently been reported by Borrero-Echeverry et al.[60]. The absence of any sort of transition at finite Reynolds number is consistent with roughness-induced criticality—the notion that pipe turbulence is governed by a non-equilibrium critical point at infinite Reynolds number in the singular limit of perfectly smooth flow boundaries[1; 2].

In this article, we show that the form of the experimental data is consistent with a simple and general interpretation using extremal statistics. Our approach is related to the notion that the transient turbulent phenomena reflect escape from a low-dimensional dynamical attractor[64–66], but we conceive turbulence as a spatially-extended phenomenon with a large number of degrees of freedom. The determining factor for the suppression of a puff is the probability that the largest fluctuation in a spatio-temporal interval consisting of multiple fluctuations fails to attain a threshold value. Thus, if we regard such dangerous fluctuations as affecting the amplitude of turbulent velocity fluctuations, we need to calculate the probability that the maximum amplitude of turbulent velocity fluctuations $\delta v(\vec{x}, t)$ falls below some threshold value, which we term A_c. We will assume below that once the turbulence has been sufficiently suppressed, the turbulent state is quenched. Our calculation shows that the super-exponential dependence of the lifetime of the turbulent state is a generic result of extremal statistics, and makes a non-trivial universal prediction about the coefficients extracted from the fits to the experimental data. We confirm this prediction using the newly-available high precision measurements of the lifetime from both pipe and Taylor-Couette flows.

In order to understand the lifetime of turbulent puffs, we assume that turbulent velocity configurations may be regarded as independent over a mean time scale τ_0, and that there is a probability p that the puff will be suppressed within each time interval τ_0. Then, the lifetime statistics will be Poisson. The probability P that turbulence persists to a time t after becoming established at a time t_0 is $P = (1-p)^N$, where the number of intervals is $N\tau_0 = (t - t_0)$. Therefore

$$\ln(P) = N\ln(1-p) = f(t-t_0)\ln(1-p), \tag{4.1}$$

and so it follows that $1/\tau = -f\ln(1-p)$. Since $1 \gg p > 0$, we can estimate $\ln(1-p) = -p$ and therefore express the time constant of the lifetime statistics in the form

$$\tau = \tau_0/p. \tag{4.2}$$

We now determine how p depends on the Reynolds number of the flow and potentially other factors. Within a spatial and temporal interval, multiple fluctuations occur, sampled from the turbulent velocity distribution $P_T(\delta v)$. There is an energy associated with these fluctuations δv^2. When this energy fails to attain a certain threshold at all points in the puff, the turbulent state becomes unstable and decays. If we consider a Gaussian distribution of velocity fluctuations with standard deviation σ, then the probability distribution of the magnitude of the velocity fluctuations

$b = \sqrt{\delta v^2}$ is given by

$$P(b) = S_d(\frac{1}{2\pi\sigma^2})^{3/2} r^2 \exp(-b^2/2\sigma^2), \tag{4.3}$$

where S_d is the surface area of a unit sphere in d dimensions. Since we are interested in whether or not the system has a sufficient level of energy to sustain turbulence compared to a critical amount that scales in some way with Reynolds number, let us consider the probability distribution of $z \equiv \ln(b)$:

$$P(z) = S_d(\frac{1}{2\pi\sigma^2})^{3/2} \exp((d-2)z) \exp(-\exp(2z)/2\sigma^2). \tag{4.4}$$

The behavior at small z, which is the regime of interest, is dominated by an exponential tail. If the distribution of velocity fluctuations were exponential rather than Gaussian (as is the case at high Reynolds numbers), the behavior at large z would change but the exponential prefactor would still dominate at small z. As such, the probability distribution $P_M(x)$ for the maximum x of a set of energy fluctuations $\{\delta v_i^2\}$ is governed by the appropriate Fisher-Tippett distribution[67], where the universality class for P_M must be the Type II Fisher-Tippett distribution, sometimes known as the Gumbel distribution[68]

$$P_M(x) = \frac{1}{\beta} \exp(-(x-\mu)/\beta) \exp(\exp(-(x-\mu)/\beta)), \tag{4.5}$$

where β sets the scale and μ the location of the distribution. The mean and standard deviation are $\mu + \Gamma\beta$ and $\beta\pi/\sqrt{6}$ respectively, where $\Gamma \approx 0.577$ is the Euler-Mascheroni constant. The corresponding cumulative distribution is the probability that $x < X$, and is given by

$$F(X) \equiv \int_{-\infty}^{X} P_M(x)\, dx = \exp(-\exp(-(X-\mu)/\beta)). \tag{4.6}$$

Thus, $p = F(\ln(A_c))$, where A_c is the threshold.

We anticipate that A_c is a decreasing function of Re, reflecting the intuition that at higher Re, turbulence can be more easily sustained by small fluctuations. We will consider the behavior of $\ln(A_c)$, as this sets the threshold in the distribution of energy maxima. The experiments are conducted in nominally smooth pipes within a narrow range of Re, so it is appropriate to expand $B_c = \ln(A_c(\mathrm{Re}))$ about a particular Reynolds number Re_0, leading to $B_c = B_c^0 + B_c^1(\mathrm{Re} - \mathrm{Re}_0) + O(\mathrm{Re}^2)$. We choose Re_0 to be the Reynolds number at which localized turbulent puffs first appear. This onset is not a precisely defined point because, as the puffs are metastable, direct determination of Re_0 depends on the timescale of the measurements. This is familiar from the coexistence point between liquid and gas phases in equilibrium statistical mechanics of fluids.

Collecting results, we find that the average lifetime of a turbulent puff will scale as:

$$\tau = \tau_0 \exp(\exp(-(B_c^0 + B_c^1(\mathrm{Re} - \mathrm{Re}_0) + O((\mathrm{Re} - \mathrm{Re}_0)^2))) \tag{4.7}$$

in agreement with experimental findings. The coefficient τ_0 may in principle depend on pipe length or aspect ratio if these factors change the spatial scale on which regions of the localized turbulent puff are statistically independent, and the timescale on which the state of the puff loses memory of previous states. This analysis suggests that the experiments measuring the lifetime of turbulent puffs have been conducted (understandably) over too small of a range of Reynolds numbers to detect effects of order Re^2 or higher. A similar scaling should also be present, in principle, for the dependence on r; however, it is not clear whether there are any grounds for a more quantitative scaling prediction, along the lines given for pipe friction factor scaling at high Re[1].

We make an additional prediction beyond the universal functional form of the lifetime of the turbulent state. From (4.6) we can write

$$\ln\ln(\tau/\tau_0) = -(B_c - \mu)/\beta = c_1\mathrm{Re} + c_2 \tag{4.8}$$

where we have used the notation of Ref. ([59]) to denote the coefficients c_1 and c_2 of the linear fit to the data. Comparing with Eq. (4.7) we read off that

$$c_1 = -B_c^1/\beta \qquad c_2 = -(B_c^0 - \mu - B_c^1 \mathrm{Re}_0)/\beta \tag{4.9}$$

Now, at Re_0, the lifetime becomes comparable to the correlation time τ_0. We will assume that for $\mathrm{Re} = \mathrm{Re}_0$, $\tau = e\,\tau_0$ for definiteness, although it is straightforward to verify that our results have only a very weak dependence on the precise coefficient used. Then from Eq. (4.7), we see that $B_c^0 = \mu$, and thus the ratio of the coefficients $c_2/c_1 = -\mathrm{Re}_0$. Thus, the Reynolds number at which the localized puffs begin to have a significant lifetime of $O(\tau_0)$ can be determined from measurements of the constants c_1 and c_2.

We now discuss the role of the spatial structure of the localized turbulent puffs within the framework of our approach. The puffs are localized, and thus do not reflect the global geometry of the flow. Thus we expect that Re_0 should be a local property determined only by the flow geometry on the lengthscale of the puff, in contrast to a global transitional Reynolds number associated with the onset of fully developed turbulence. As such, a Taylor-Couette apparatus with a sufficiently large radius of curvature (larger than the scale of the localized puff) should have the same ratio c_2/c_1 as a straight pipe. The Reynolds number must be defined in a consistent fashion for this ratio to be comparable for different systems. For example, in Taylor-Couette flow, it must be chosen such that in the limit of an infinite radius of curvature the Reynolds number reduces to the Reynolds number of the straight pipe. As long as a consistent definition of Re is used for different flows, their transitional behavior is directly comparable, and the predicted value for c_2/c_1 should be the same, with no adjustable parameters. In this sense, the ratio is a universal one, only for a given definition of Re, as is commonplace in fluid dynamics. These considerations also suggest that in short pipes, where the length to diameter ratio is small compared to the puff size, the lifetime statistics should not be described by our approach, because the transient turbulence would then fill the pipe completely, as observed in direct numerical simulations[69; 70]. In such a situation, the lifetime will scale with the length of the pipe, but in the regimes of the experiments on pipe and Taylor-Couette flows, the puffs are smaller than the streamwise direction length, and the lifetime should then be independent of the length of streamwise direction length.

To test this prediction, we compare the two sets of available experimental results, for pipe and Taylor-Couette geometries. The definitions of Re used in both sets of experiments are identical in the limit that the gap between the cylinders remains fixed as the circumference of the Taylor-Couette apparatus becomes large. In the pipe flow experiments of Hof et al. [59], the fitted values are reported to be $c_1 = 0.0057$ and $c_2 = -8.7$, giving $c_2/c_1 = 1526$. In the Taylor-Couette experiments of Borrero-Echeverry et al.[60], $c_1 = 3.57 \times 10^{-4}$ and $c_2 = -0.56$, giving $c_2/c_1 = 1569$. Thus, while the details of the experiments were very different, the ratio of c_2/c_1 is the same (to a good approximation) in both cases, as predicted. Furthermore, the value of Re_0, in the range 1526-1569, is a reasonable approximation to the onset of turbulence in the pipe experiments. The Taylor-Couette experiment and the pipe flow experiments differ in their aspect ratio (length of the flow arena/diameter of flow): in the former it is 46 in the streamwise direction, whereas in the latter it ranges from 143 to 3600. Due to the finite nature of the turbulent puff, we expect that the intrinsic onset value Re_0 should be related only to the local geometry rather than the global geometry, and so we expect it to be different than the reported transitional Re for Taylor-Couette flows.

In conclusion, we have argued that the lifetime statistics of transient turbulent puffs are governed by extreme events, showing that the Reynolds number dependence observed in experiments has the expected form, and verifying that the available experiments are consistent with a parameter-free prediction about the details of the lifetime statistics near the turbulence threshold.

Chapter 5

Simulation of Two-dimensional Pipe Flow

We have established a theory for the difference between the scalings of two-dimensional and three-dimensional flows. In two dimensions, the friction factor and velocity profile should both depend on the particular turbulent cascade that is exhibited by the flow. In the inverse cascade, the results are the same as for three-dimensional flows, but in the enstrophy cascade we expect new scalings. How shall we test these predictions?

One way would be to look at naturally occuring flows. Atmospheric flows are laminated in nature, owing to sharp gradients in temperature and pressure with respect to height, and have been known to exhibit enstrophy cascade scaling[36]. Furthermore, experiments have been done with soap film flows that have also shown the enstrophy cascade scaling[37]. We will also briefly discuss a comparison between my simulations and soap film experiments performed by Walter Goldburg, et al.[71] in the chapter detailing the results of this study.

However, there are many problems associated with either of these systems. Naturally occuring flows are at best quasi-two-dimensional, in that the viscous lengthscale is longer than the third dimension of the flow. Fluctuations out-of-plane may still occur, leading to behavior similar to that of compressible flows — waves and other such phenomena can occur at the surface. Furthermore, even when such fluctuations are well-controlled, we must contend with transport at the interface between the confined flow and its surroundings. In soap film flows, air resistance modifies the shape of the velocity profile[72; 73], and may change the observed friction factor. These problems can be controlled by using a soap film flow in a vacuum, a rather difficult experiment to perform.

Another option is to solve the Navier-Stokes equations computationally. This has the advantage that one has full access to the entire flow field at every point in time, so it is easy to make measurements that would be difficult to make experimentally. It is also possible to simulate a perfectly two-dimensional flow, with no out of plane fluctuations or transport at the interfaces.

The disadvantages of simulation for this problem must also be considered. The friction factor is dominated by the behavior of an extremely thin layer of the flow near the wall. If this layer is under-resolved, it is possible to have a flow field that appears to be correct in terms of measurements made in the bulk, but which gives the incorrect friction factor. Furthermore, the friction factor is very sensitive to errors made in establishing the boundary conditions, and so methods such as the immersed boundary method [74] for handling the roughness may systematically alter the results unless a finite volume method is being used to ensure that they are conservative. Another problem is that of numerical dissipation, and numerical instability. If a low-order advection scheme is used, then it will artificially increase the effective viscosity of the system and prevent resolution of high Reynolds numbers. If, on the other hand, a high-order advection scheme is used, then the simulation tends to be numerically unstable to the formation of waves. Perhaps most significantly, simulating systems large enough to avoid entry effects and exit effects is extremely computationally expensive. In three dimensions, simulating Nikuradze's full apparatus is completely beyond current capabilities. In two dimensions, it would be possible, but prohibitively costly.

Fortunately, we can resolve some of these issues by judicious choices of algorithm. The three big choices to make are: how do we represent the flow, how do we handle advection, and how do we handle incompressibility?

5.1 Representation

When we talk about simulating the Navier-Stokes equation, we've already implicitly made a choice about how to represent the flow. We could in principle use other ways to simulate a fluid flow: molecular dynamics, smooth particle hydrodynamics[75], or an incompressible lattice Boltzmann algorithm[76; 77]. These other representations have advantages and disadvantages that make them ideal for specific problems in fluid dynamics. Molecular dynamics and smooth particle hydrodynamics have the problem that the particle size must be significantly smaller than the viscous scale, so an extremely large number of particles are needed to resolve the flow. This rules them out for high Reynolds number simulations, in which the viscous scale is very small. Lattice Boltzmann instead discretizes space, and allows each grid cell to contain a number of particles with some distribution of velocities. It has the advantage over other algorithms of being unconditionally stable, so there is no issue of numerical instability. Certain artefacts become obvious at high Reynolds numbers, unless the timescale is set to be very small. If particles can move several grid cells without their momentum distribution relaxing to equilibrium, then sheets of particles tend to flow off of any obstacles to the flow. The small requisite timestep for physical behavior makes lattice Boltzmann inconvenient for high Reynolds number flows.

This leaves us with directly solving the Navier-Stokes partial differential equation. The price we pay for having a continuous representation of the velocity field and a fixed particle density is the possibility for numerical instability, and the need for a

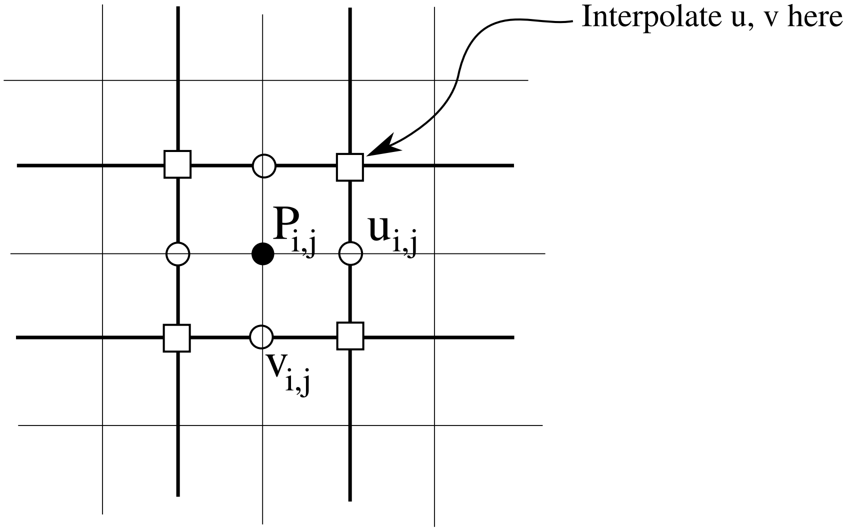

Figure 5.1: In our scheme, the pressure is stored at cell centers. The velocity in the x direction is stored at the center of the left and right sides of the cell. The velocity in the y direction is stored at the center of the top and bottom sides of the cell. For the advection algorithm, the values are interpolated so that the advected quantity is co-located with the advecting velocity.

non-local algorithm to maintain the incompressibility constraint. There is another choice to be made here — how do we discretize the equation, and where do we put the simulation variables on the grid? It is possible to use anything up to a completely unstructured grid (using finite elements methods[78]), but what we gain in generality we lose in increased computational cost and complexity of code. If we can use a rectangular, uniform, orthogonal grid, this will later let use use fast spectral methods to satisfy the incompressibility condition. After having chosen the type of grid, we also need to decide whether the velocity will be placed at the center of the grid cell, or on the faces of the grid cell. There are advantages to either choice, but we will later see that putting velocity on the faces of the grid cell allows us to ensure that the advection code conserves momentum, which is important to our friction factor measurements.

5.1.1 Conformal Mapping

If we wish to later use a spectral method to satisfy the incompressibility constraint, we must have a rectangular, orthogonal, and uniform grid. However, we would also like to simulate pipes with rough walls. We could use an immersed boundary, and treat the roughness geometry as some local body force that tries to keep velocity at zero at their surface. However, when cells within the boundary have their velocity set to zero, more momentum is lost than would normally be lost through viscous transport, and so we strongly overestimate the friction factor. It seems like it is necessary for the roughness elements to be the actual boundary of the flow.

Because we are interested in two-dimensional flows, there is a trick we can use so that we do not have to give up on using a spectral method. In two dimensions, it is possible to use conformal mapping to transform between any two domains in the plane such that all perpendicular coordinate lines in one of the domains remain perpendicular in the new one (that is, the map preserves angles). This is normally used for solving the steady-state diffusion equation $\nabla^2\psi = 0$, because the operator ∇^2 is preserved up to a multiplicative factor by conformal maps, and so the solution in one domain can be transformed to a solution on a new domain.

The Navier-Stokes equation does not transform so simply under a conformal map, though as we will see it does not change by much. However, the equation for the pressure (Eq. 3.11) is almost identical to the steady-state diffusion equation, and so we can use the mapping combined with spectral methods to solve the incompressibility condition on nontrivial domains.

A conformal map is a complex function of a complex argument that is defined solely in terms of analytic functions of that argument. As such, the complex conjugate or any direct reference to the real or imaginary part will not show up in a conformal mapping function. If we start from a domain in which we have coordinates (u, v) on the complex plane, then the complex argument of the map is $w = u + iv$. The mapping is then given as $z = f(w)$ (where $z = x + iy$), which takes us to a new domain with coordinates on the complex plane (x, y). Lines of constant u and v in the original coordinate system will be the boundaries of our computational domain. The physical coordinates of these boundaries are the corresponding $z(u)$ and $z(v)$ curves.

The derivatives of analytic functions of complex numbers with respect to the real and imaginary parts of their argument are related. If we have the complex function $z = f(w) = f(u + iv)$, we can relate dx/du, dx/dv, dy/du, and dy/dv (shortened to x_u, x_v, y_u, y_v for convenience):

$$\frac{df}{du} = -i\frac{df}{dv} \tag{5.1}$$

So:

$$x_u = \mathrm{Re}[\frac{df}{du}] = \mathrm{Im}[\frac{df}{dv}] = y_v \tag{5.2}$$

$$x_v = \mathrm{Im}[\frac{df}{dv}] = -\mathrm{Re}[\frac{df}{du}] = -y_u \tag{5.3}$$

We can also determine a relationship between the second derivatives: $x_{vv} = -y_{uv} = -x_{uu}$.

For convenience we will define the quantity $g^2 = x_u^2 + x_v^2$. This is the local scaling factor of the map.

We can use the definition of the conformal map and these properties to determine how operators transform under the map. As we are concerned with velocities, which are themselves rates of change of a coordinate with time, we will also represent the velocities in the computational coordinates rather than the real space velocities.

The vectors (x_u, y_u) and (x_v, y_v) are the local basis vectors in the physical space that correspond to the basis vectors of the computational map. As such, we can do a change of basis $\mathbf{V_{xy}} = \mathbf{M}\mathbf{V_{uv}}$, where

$$\mathbf{M} = \begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix} \tag{5.4}$$

Inverting M, we find:

$$\mathbf{M}^{-1} = \frac{1}{g^2}\begin{bmatrix} y_v & -x_v \\ -y_u & x_u \end{bmatrix} \tag{5.5}$$

$$\mathbf{V_{uv}} = \mathbf{M}^{-1}\mathbf{V_{xy}} \tag{5.6}$$

Let us define new vectors $\bar{\mathbf{V}}_\mathbf{u}$ and $\bar{\mathbf{V}}_\mathbf{v}$to remove the g^2 scaling:

$$\bar{\mathbf{V}}_\mathbf{u} = g^2\mathbf{V_u} \tag{5.7}$$

$$\bar{\mathbf{V}}_\mathbf{v} = g^2\mathbf{V_v} \tag{5.8}$$

We will now examine how the Navier-Stokes equation transforms under a conformal map by looking at each of its operators. First, we will determine how a partial derivative transforms.

We apply our coordinate transform to the vector of partial derivatives:

$$\begin{bmatrix} x_u & y_u \\ x_v & y_v \end{bmatrix}\begin{bmatrix} \partial_x \\ \partial_y \end{bmatrix} = \begin{bmatrix} \partial_u \\ \partial_v \end{bmatrix} \tag{5.9}$$

So, solving and using the relations between derivatives, this becomes:

$$\partial_x = \frac{x_u\partial_u + x_v\partial_v}{g^2} \tag{5.10}$$

$$\partial_y = \frac{x_u\partial_v - x_v\partial_u}{g^2} \tag{5.11}$$

The gradient transforms as:

$$\nabla_\mathbf{xy} = \frac{1}{g^2}(\hat{\mathbf{x}}(x_u\partial_u + x_v\partial_v) + \hat{\mathbf{y}}(x_u\partial_v - x_v\partial_u)) \tag{5.12}$$

Next, we will look at the transformation of the Laplacian:

$$\begin{aligned}\nabla^2_{xy} =& \partial_x\partial_x + \partial_y\partial_y = \left(\frac{y_u\partial_v - y_v\partial_u}{x_v y_u - x_u y_v})^2 + (\frac{x_u\partial_v - x_v\partial_u}{x_v y_u - x_u y_v}\right)^2 \\ =& \frac{(y_u^2 + x_u^2)\partial_v^2 + (y_v^2 + x_v^2)\partial_u^2 - 2(y_u y_v + x_u x_v)\partial_{uv}}{(x_v y_u - x_u y_v)^2}\end{aligned} \tag{5.13}$$

This is problematic because of the ∂_{uv} term. But because a conformal map is also an orthogonal map, that term vanishes as it is just the dot product between the **u** and **v** basis vectors. For a conformal map, these directions have the same length, and so $x_u^2 + y_u^2 = x_v^2 + y_v^2 = x_v y_u - x_u y_v = g^2$:

$$\nabla^2_{xy} = \frac{1}{g^2}\nabla^2_{uv} \tag{5.14}$$

We will also need the divergence of our remapped vectors in order to compute the right hand side of the pressure equation:

$$\begin{aligned}\partial_x\mathbf{V}_x + \partial_y\mathbf{V}_y =& \frac{1}{g^2}[(y_u\partial_v - y_v\partial_u)\mathbf{V}_x - (x_u\partial_v - x_v\partial_u)\mathbf{V}_y] \\ =& \frac{1}{g^2}\left((x_v\partial_v + x_u\partial_u)\frac{x_u\bar{\mathbf{V}}_u + x_v\bar{\mathbf{V}}_v}{g^2} + (x_u\partial_v - x_v\partial_u)\frac{-x_v\bar{\mathbf{V}}_u + x_u\bar{\mathbf{V}}_v}{g^2}\right)\end{aligned} \tag{5.15}$$

Writing this out in full will give three sets of terms: derivatives of g^2, derivates of the map coordinates, and derivatives of the vector quantities. The result after some cancellation is:

$$\begin{aligned}&(1/g^4)[g^2(\partial_u\bar{\mathbf{V}}_u + \partial_v\bar{\mathbf{V}}_v) - (\bar{\mathbf{V}}_u\partial_u + \bar{\mathbf{V}}_v\partial_v)(g^2/2) + \\ &\qquad \bar{\mathbf{V}}_u(x_v x_{uv} + x_u x_{uu}) + \bar{\mathbf{V}}_v(x_u x_{uv} - x_v x_{uu})]\end{aligned} \tag{5.16}$$

Evaluating the derivatives of g^2, we find that:

$$\partial_u(g^2/2) = x_u x_{uu} + x_v x_{uv} \tag{5.17}$$

$$\partial_v(g^2/2) = x_u x_{uv} - x_v x_{uu} \tag{5.18}$$

These exactly cancel out the other terms in the first part of Eq. 5.16, leaving the final relationship for the mapped divergence:

$$\nabla_{\mathbf{xy}} \cdot \mathbf{V_{xy}} = (1/g^2)\nabla_{\mathbf{uv}} \cdot \bar{\mathbf{V}}_{\mathbf{uv}} \tag{5.19}$$

Combining Eqs. 5.14 and 5.19, we can write down the equation for the incompressibility constraint in the transformed coordinates:

$$\nabla^2_{uv} P = \nabla_{\mathbf{uv}} \cdot \bar{\mathbf{V}}_{\mathbf{uv}} \tag{5.20}$$

This is exactly the same as in the untransformed coordinates, so the pressure equation is conformally invariant.

We now compute the mapping of the advection operator on an arbitrary scalar ϕ: $V \cdot \nabla\phi$.

$$\begin{aligned}(V_x\partial_x + V_y\partial_y)\phi = \frac{1}{g^4}((x_u\bar{\mathbf{V}}_u + x_v\bar{\mathbf{V}}_v)(x_u\partial_u + x_v\partial_v) +\\ (-x_v\bar{\mathbf{V}}_u + x_u\bar{\mathbf{V}}_v)(x_u\partial_v - x_v\partial_u))\phi\\ = \frac{1}{g^2}(\bar{\mathbf{V}}_u\partial_u + \bar{\mathbf{V}}_v\partial_v)\phi\end{aligned} \tag{5.21}$$

So the advection operator on a scalar is scaled by $1/g^2$ but is otherwise unchanged. If we instead operate on a vector, we must also worry about the derivatives of the coordinate transform M of the vector:

$$\mathbf{V_{xy}} \cdot \nabla\mathbf{V_{xy}} = \frac{1}{g^2}(\bar{\mathbf{V}}_{\mathbf{uv}} \cdot \nabla(M/g^2)\bar{\mathbf{V}}_{\mathbf{uv}}) = \frac{1}{g^2}(\mathbf{M}/g^2)\bar{\mathbf{V}}_{uv} \cdot \bar{\mathbf{V}}_{\mathbf{uv}} + \bar{\mathbf{V}}^2_{\mathbf{uv}}\nabla(\mathbf{M}/g^2)) \tag{5.22}$$

Furthermore, this is in terms of the x and y components of the advection operator. Since we want the u and v components, we multiply by $g^2\mathbf{M}^{-1}$:

$$\frac{1}{g^2}(\bar{\mathbf{V}}_{\mathbf{uv}} \cdot \nabla_{\mathbf{uv}}\bar{\mathbf{V}}_{\mathbf{uv}} + \mathbf{M}^{-1}g^2(\bar{\mathbf{V}}_{\mathbf{uv}} \cdot \nabla_{\mathbf{uv}})(\mathbf{M}/g^2)\bar{\mathbf{V}}_{\mathbf{uv}}) \tag{5.23}$$

We have recovered the usual advection operator on the new velocity vector, in addition to a term that takes into account local changes in the orientation of the coordinate system. This extra term is:

$$-\frac{|\bar{\mathbf{V}}_{\mathbf{uv}}|^2}{g^4}\begin{bmatrix} x_u x_{uu} + x_v x_{uv} \\ x_u x_{uv} - x_v x_{uu} \end{bmatrix} \equiv -\frac{|\bar{\mathbf{V}}_{\mathbf{uv}}|^2}{g^4}\mathbf{A} \tag{5.24}$$

where:

$$\mathbf{A} = \begin{bmatrix} x_u x_{uu} + x_v x_{uv} \\ x_u x_{uv} - x_v x_{uu} \end{bmatrix} \tag{5.25}$$

Finally, we look at the viscous diffusion operator.

$$\nabla^2_{xy} = \frac{1}{g^2}\nabla^2_{uv} \tag{5.26}$$

$$\nabla^2_{xy}\mathbf{V}_{xy} = \mathbf{M}\frac{1}{g^2}\nabla^2_{uv}\begin{bmatrix} V_x \\ V_y \end{bmatrix} \tag{5.27}$$

We define: $\phi_x = \nabla^2_{uv}V_x$ and $\phi_y = \nabla^2_{uv}V_y$. Changing $\mathbf{V}_{xy}$ to $\bar{\mathbf{V}}_{uv}$, we have:

$$\phi_i = \partial_j\partial^j g^2\mathbf{M}^{-1}_{ik}\bar{\mathbf{V}}^k \tag{5.28}$$

Let us define the matrix:

$$\mathbf{N} \equiv g^2\mathbf{M}^{\mathbf{-1}} = \begin{bmatrix} x_u & -x_v \\ x_v & x_u \end{bmatrix} \tag{5.29}$$

We have:

$$\phi = \mathbf{N}_{ik}\partial_j\partial^j\bar{\mathbf{V}}^k + \bar{\mathbf{V}}^k\partial_j\partial^j(\mathbf{N}_{ik}) + 2\partial_j(\mathbf{N}_{ik})\partial^j\bar{\mathbf{V}}^k \tag{5.30}$$

We evaluate the various derivatives of the matrix $\mathbf{N}$:

$$\partial_u\mathbf{N} = \begin{bmatrix} x_{uu} & -x_{uv} \\ x_{uv} & x_{uu} \end{bmatrix} \tag{5.31}$$

$$\partial_v\mathbf{N} = \begin{bmatrix} x_{uv} & x_{uu} \\ x_{uu} & x_{uv} \end{bmatrix} \tag{5.32}$$

$$\partial_u^2\mathbf{N} = \begin{bmatrix} x_{uuu} & x_{vvv} \\ -x_{vvv} & x_{uuu} \end{bmatrix} \tag{5.33}$$

$$\partial_v^2\mathbf{N} = \begin{bmatrix} -x_{uuu} & -x_{vvv} \\ x_{vvv} & -x_{uuu} \end{bmatrix} \tag{5.34}$$

$$\tag{5.35}$$

This implies that $\partial_j\partial^j\mathbf{N}_{ik} = 0$.

The viscous operator is:

$$\nabla^2_{xy}\mathbf{V}_{xy} = \mathbf{M}\frac{1}{g^2}\phi \tag{5.36}$$

Combining these results, we have:

$$\begin{gathered} \nabla^2_{xy}\mathbf{V}_{xy} = \nabla^2_{uv}\bar{\mathbf{V}} + \\ \frac{2}{g^2}\left(\begin{bmatrix} x_u x_{uu} + x_v x_{uv} \\ x_v x_{uu} - x_u x_{uv} \end{bmatrix}\nabla_{uv}\cdot\bar{\mathbf{V}} + \right. \\ \left.\begin{bmatrix} x_u x_{uv} - x_v x_{uu} \\ -x_u x_{uu} - x_v x_{uv} \end{bmatrix}\nabla_{uv}\times\bar{\mathbf{V}}\right) \end{gathered} \tag{5.37}$$

Since the flow is divergence free, the viscous operator reduces to:

$$\nabla^2_{xy}\mathbf{V}_{xy} = \nabla^2_{uv}\bar{\mathbf{V}} + \frac{2}{g^2}\mathbf{A}^{\perp}\nabla_{uv} \times \bar{\mathbf{V}} \tag{5.38}$$

where:

$$\mathbf{A}^{\perp} = \begin{bmatrix} x_u x_{uv} - x_v x_{uu} \\ -x_u x_{uu} - x_v x_{uv} \end{bmatrix} \tag{5.39}$$

This is just the perpendicular vector to the vector **A** that we found in the advection term.

The complete conformally mapped Navier-Stokes equation is therefore:

$$g^2\frac{\partial \bar{\mathbf{V}}}{\partial t} + (\bar{\mathbf{V}} \cdot \nabla)\bar{\mathbf{V}} = \nu\nabla^2\bar{\mathbf{V}} + \frac{|\bar{\mathbf{V}}|^2}{g^2}\mathbf{A} + \frac{2\nu}{g^2}\mathbf{A}^{\perp}(\nabla \times \bar{\mathbf{V}}) \tag{5.40}$$

All that has happened is that we have added two body force terms to the Navier-Stokes equation in the transformed (rectangular) domain, in addition to an overall spatial weighting factor g^2 that modifies the time derivative, resulting from the changed volume of each cell.

To simulate a rough-walled pipe, we apply a conformal map of the form $w = z + r\exp(ikz)$, where the aspect ratio is held constant ($rk = 3/4$) and the wavenumber k may be varied to produce roughness of different scales. Note that r plays the role of roughness in Nikuradse's experiments, but our aspect ratio is 3/4 and not unity as in his experiments.

5.2 Advection

We would now like to consider how to handle advection. Advection is the sole nonlinear operator in the Navier-Stokes equation. It is also the source of most of the numerical instabilities, at least at high Reynolds number. The simplest advection

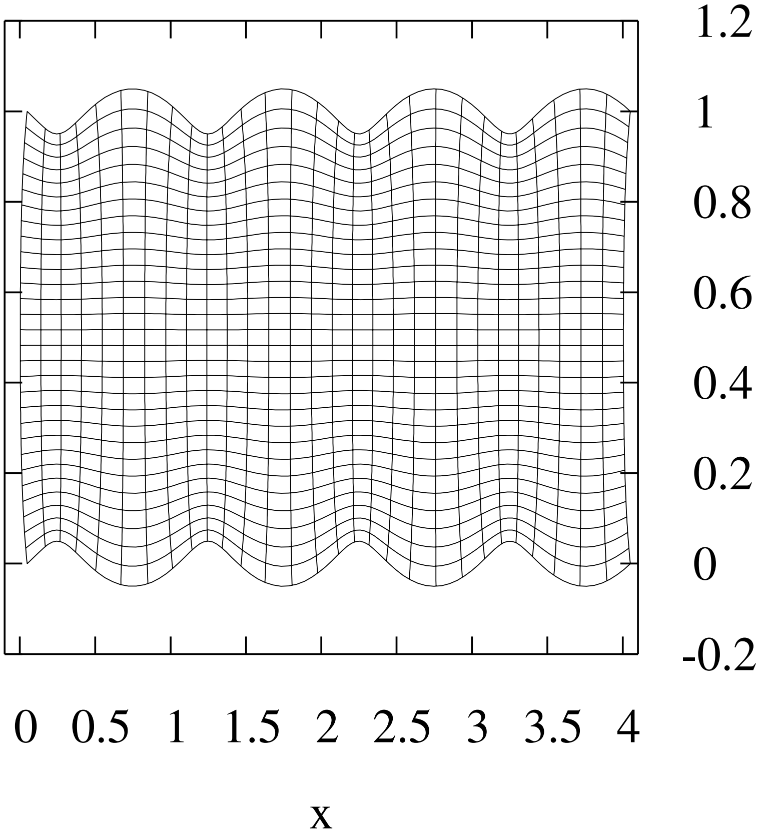

Figure 5.2: Conformal mapping of the rough pipe used for our simulations.

methods are first order upwinding algorithms. These discretize the operator $u\frac{du}{dx}$ around the grid site i in the following way:

$$\begin{array}{ll} u_i\frac{u_i-u_{i-1}}{\Delta x} & u_i \geq 0 \\ u_i\frac{u_{i+1}-u_i}{\Delta x} & u_i < 0 \end{array} \tag{5.41}$$

If we Taylor expand u around u_i, we can see the error we make in approximating the actual operator. We will take the $u_i > 0$ case for simplicity.

$$u_{i-1} = u_i - \Delta x \frac{du}{dx}_i + \frac{\Delta x^2}{2}\frac{d^2u}{dx^2}_i + O(\Delta x^3) \tag{5.42}$$

$$\epsilon = u_i \frac{u_i - u_{i-1}}{\Delta x} - u_i \frac{du}{dx}_i \tag{5.43}$$

$$\epsilon = -u_i \frac{\Delta x}{2} \frac{d^2u}{dx^2}_i + O(\Delta x^2) \tag{5.44}$$

This means that the consequence of using the upwinding operator in place of the actual continuous operator is that we have added a term $u_i \frac{\Delta x}{2} \frac{d^2u}{dx^2}_i$ to the opposite side of the Navier-Stokes equation. This term has the same form as a viscous term with viscosity $u_i \frac{\Delta x}{2}$. As a consequence, the errors we make in having too large of a grid size will take the form of an increase in viscosity. If the errors had the opposite sign, then this would decrease the effective viscosity. If that made the effective viscosity negative, then the simulation would exhibit a checkerboard pattern of instability at the grid scale. This is why the upwinding algorithm chooses the direction in which to evaluate the derivative based on the local velocity — it is ensuring that the error term always increases the effective viscosity rather than decreasing it. This makes the algorithm rather stable, but has the downside of introducing significant numerical dissipation.

To fix this problem, we must go to a higher-order scheme, so that the additional terms will be of order Δx^2 or better. In practice, going beyond third order is somewhat futile as numerical dissipation will give way to lack of resolution as the main barrier to reaching high Reynolds numbers. We could construct a third order scheme by using the Taylor expansion of u around u_i, and combining neighboring values of u until we had eliminated all error terms of order Δx^3 or smaller. This solves the numerical dissipation issue, but adds a new problem: because it is somewhat nonlocal, the third order advection operator may introduce new maxima and minima into the advected field. This gives rise to non-physical wave behavior that eventually blows up, unless the timestep is kept very small.

This problem has been addressed by so-called flux limiter advection schemes, that

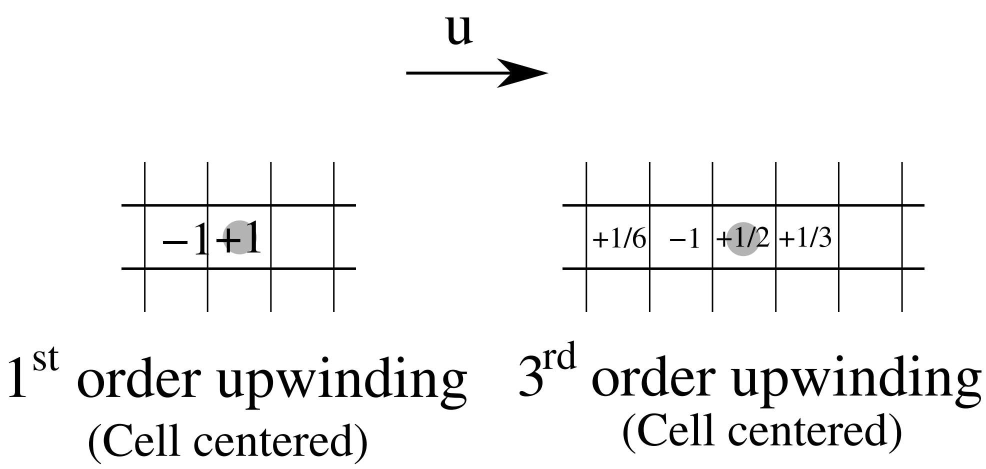

Figure 5.3: Stencils for the first order and third order cell-centered upwinding schemes.

make use of a boundedness criterion in order to ensure that waves are not introduced. For a survey of various flux limiter schemes, their implementations, and the principles on which they work, see [79]. The SMART algorithm[80] is a compromise between the complexity of the operator (which makes it more computationally expensive to evaluate) and stability. The SMART algorithm is a third-order algorithm that interpolates the velocity in a controlled fashion, dependent on the local derivatives, in order to ensure that the resulting interpolated function is bounded above and below by the largest and smallest values in the local region. This prevents that overshoot from manifesting as spurious waves.

Another advantage of the SMART algorithm is that it is written to make use of fluxes at the walls of the computational grid cells, rather than the flow through the cell center. This means that we can guarantee conservation of momentum. If we place the velocities on the cell faces, then we can write an expression for the net flux of momentum through either side of a cell. The difference in the fluxes then gives us our change in momentum. Since the flux into the next cell is the same as the flux out of the cell sharing its face, so long as all the cells are properly connected to each other and the boundaries then the momentum of the system will be conserved. This is in fact the same thing as the advection operator, but with the added benefit that even if the flow field is compressible, it will properly conserve momentum.

In the case of high velocities near a boundary, the SMART algorithm can have symmetry problems due to an arbitrary choice of direction to do upwinding when $v = 0$ exactly. We expect this effect to be minimal in fully developed flows where $v \sim 0$ near the boundary due to viscous effects, but work is being done to test these results with an algorithm that does not suffer this symmetry problem.

5.3 Incompressibility

A large part of the shedding of turbulent eddies from obstacles in a flow comes from the proper consideration of incompressibility. This is what causes the flow to curl up into vortices rather than smear out in long shearing streaks. Satisfying the incompressibility constraint can be the most time consuming part of the turbulence simulation, as it is a non-local problem.

There are a number of ways to approach the constraint. One way would be to use a formulation of the Navier-Stokes equation that eliminates the need for the constraint entirely by using a basis for the flow field that is automatically divergence free. The way to do this is to take the curl of the Navier-Stokes equation to get an equation for the vorticity:

$$\frac{d\omega}{dt} + v \cdot \nabla \omega = \nu \nabla^2 \omega \tag{5.45}$$

This eliminates the pressure, and guarantees that the velocity field is divergence free (because it is specified only by its curl). We still must determine the velocity field at each timestep in order to evaluate the advection operator. This is, unfortunately, a non-local problem:

$$\nabla \times \mathbf{v} = \omega \tag{5.46}$$

So, taking the curl of both sides:

$$\nabla \times \nabla \times \mathbf{v} = -\nabla^2 \mathbf{v} = \nabla \times \omega \tag{5.47}$$

So we must solve a Poisson equation over the entire domain every timestep. The advantage of using a vorticity representation is that anomalies in the solution to the Poisson equation (e.g. by using an iterative solver rather than an exact solver) will not propagate from time step to time step they way they might if the velocity field itself were allowed to have a divergence. However, the boundary conditions for the vorticity representation are very hard to determine correctly and furthermore we would have to work backwards to find the pressure field in order to measure the friction factor. So we will not use this method.

Other approaches usually involve different ways of solving the pressure Poisson equation in the standard representation of Navier-Stokes. If we have just updated the velocity field from a previous step, it may have divergences introduced by the advective operator or by various errors in discretization. We will then perform a projection step which determines the pressure field that renders our velocity field divergence free. We stagger the advection and diffusion steps with these projection steps to ensure that at the end of every timestep we have a divergence-free velocity field.

If our velocity field before projection is $\mathbf{v}$, then our velocity field after projection (implemented via the pressure gradient) will be $\mathbf{v} - \nabla \mathbf{P} \Delta t$. Let us take the divergence of this and set it to zero:

$$\nabla^2 P = \nabla \cdot \mathbf{v} / \Delta t \tag{5.48}$$

This is the pressure Poisson equation that we must solve to ensure that our velocity field is divergence free. One small point is that our results depend on how we choose to discretize the divergence operator and the Laplacian. We should be careful to choose complementary definitions with how we implement advection, or we will introduce errors wherever we assume incompressibility. An example choice for velocities that are stored on the grid cell walls and pressures stored at the cell centers is:

$$\nabla \cdot \mathbf{v} = \frac{(u_{i+1/2,j} - u_{i-1/2,j}) + (v_{i,j+1/2} - v_{i,j-1/2})}{\Delta x} \tag{5.49}$$

$$\nabla^2 P = \frac{P_{i+1,j} + P_{i-1,j} + P_{i,j+1} + P_{i,j-1} - 4P_{i,j}}{4\Delta x^2} \tag{5.50}$$

These two representations are consistent because we can arrive at the representation for ∇^2 by applying the representation for ∇ to itself.

Once we have this equation, we can use a number of different solution techniques. The problem we are solving is linear, so it would be possible to set the problem up as a matrix and then exactly invert the matrix. While this has the advantage that the resulting pressure field is an exact solution to guarantee a divergence-free velocity field, it has the problem that exact matrix inversion scales as $O(N^3)$ in the dimension of the matrix, which for a two-dimensional domain means that it scales as L^6.

If we are willing to sacrifice accuracy, we can use any number of relaxation methods to solve the problem. The immediate connection to our previous attempt is to use a conjugate gradient solver to iteratively approximate the inverse of the matrix. We may also use a multigrid method to quickly converge to a solution. Multigrid methods work by representing the problem at multiple scales, and first solving it at the coarsest scales. The reason that Poisson's equation takes a long time to converge is that there is a limit to how fast information can be allowed to propagate or the process becomes numerically unstable. This limit is the size of the stencil used to represent the Laplacian operator — basically, if the stencil connects to nearest neighbors, then the continuous solution cannot call for information to move more than one grid cell per iteration.

If, however, we coarsen the problem to a grid which only has four cells, then moving one grid cell per iteration is not a very stringent constraint. We can then take the coarse solution and use it for an initial condition for the next level of refinement, and iterate this process recursively until we are back at the true grid resolution. This process takes $O(N \log N)$ steps to converge for a one dimensional system of length N. This is very competitive. The main downsides to the multigrid method are that it can be hard to handle certain cases around the boundaries when the coarsening destroys boundary detail, and that it can lead to more complex code with more room for errors. It is also not an exact solver.

It is possible to solve the pressure equation exactly in $O(N \log N)$ steps in certain cases. Note that the operators in the pressure equation are translationally invariant and are linear. This suggests that taking a Fourier transform of the equation will convert the partial differential equation into an algebraic equation. If we do this to the continuous pressure equation, we see that:

$$-k^2 \tilde{P} = i\mathbf{k} \cdot \tilde{\mathbf{v}} / \Delta t \tag{5.51}$$

We can immediately solve this for $\tilde{P}$ and then invert the Fourier transform to find the pressure field. In practice we will want to be careful to take the Fourier transform of the actual operators we are using for discretization, rather than the transform of the continuous operators:

$$\frac{\cos(2k\Delta x) - 1}{\Delta x^2} \tilde{P} = F[\nabla \cdot \mathbf{v} / \Delta t] \tag{5.52}$$

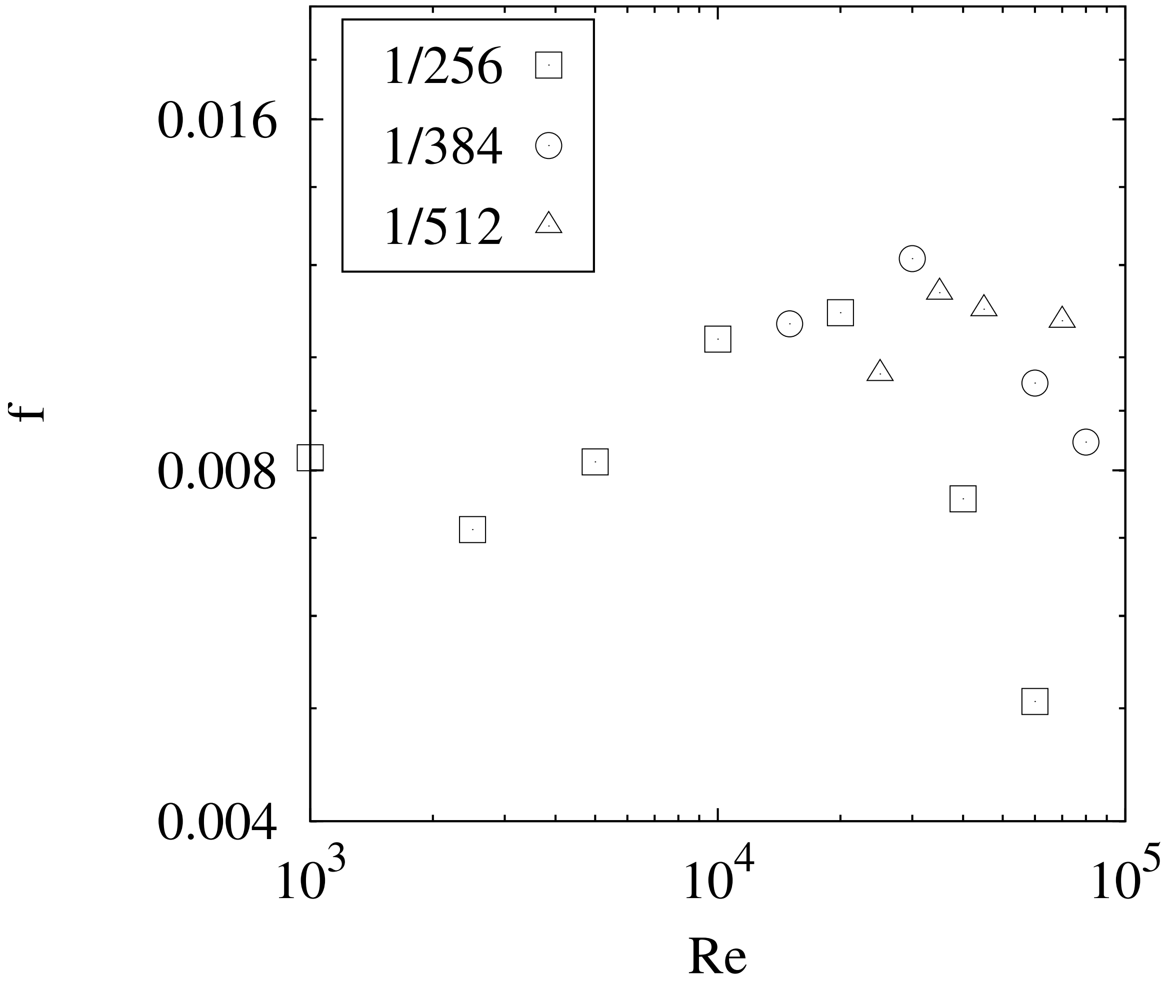

Figure 5.4: Plots of the friction factor at high roughness computed in simulations at resolutions $\Delta x = 1/256$, $\Delta x = 1/384$, and $\Delta x = 1/512$. The behavior at high Reis expected to be in the saturated part of the Strickler regime. Beyond the limit of grid convergence, the friction factor begins to drop as 1/Re as the viscous layer becomes under-resolved.

5.4 Simulation Details

Now that we have chosen the algorithm to use, we can discuss the detailed setup of the simulation. It is necessary to be able to represent the viscous lengthscale, or the effective value of the viscosity will be higher than the intended value, resulting in a decreased Reynolds number. More importantly, we will under-resolve the scaling of the viscous layer, and so will not correctly determine the transport of momentum. It is useful to understand what our results will look like in this case.

If we have under-resolved the viscous layer, then all viscous transport will take place across one computational grid cell. In the viscous layer, the velocity profile is linear, with a slope $\bar{U}^2 f/\nu$. There is a certain velocity proportional to the mean velocity at which the viscous layer starts to become turbulent — let us call this $U_c = a\bar{U}$. If l is the width of the viscous layer (which normally would depend on Reynolds number[20]), then this implies that $f = \nu U_c \bar{U}^{-2} l^{-1}$, or $f = a\mathrm{Re}^{-1}(R/l)$. The scaling of l then determines the friction factor scaling. If now we fix the width of the viscous layer to be Δx, this means we will see a friction factor that scales as $f = a\mathrm{Re}^{-1}(R/\Delta x)$ — in other words, we will see the flow exhibit a laminar friction factor scaling. This means that if we are under-resolved at a certain Reynolds number, the friction factor will appear to sharply drop below its normal behavior.

We have performed a series of simulations at different grid sizes to test our grid convergence. The friction factor curves from these parallel runs are plotted in Fig.(5.4). It appears that in order to approach Reynolds numbers of 60000 or so and still be converged, we need a grid spacing $\Delta x = 1/512$ or smaller.

As we are making our domain periodic, there is also the question of how long of a section of pipe we need. If we only used a segment one pipe diameter long, we are implicitly saying that the largest turbulent structure should be of about that length-scale. Furthermore, if there are interactions between adjacent roughness elements, we are saying that those interactions should always see a periodic flow structure. How long then do we need to make the pipe? Simulations at a variety of pipe lengths show that when the pipe length is only two pipe diameters, the periodic symmetry causes pairs of stationary vortices to form, one in a cove at the top part of the pipe, and one in a cove at the bottom part, offset by one roughness element. This structure weakens at a length of three diameters, and appears to go away at a length of four diameters.

We use a simulation domain of 2048×512 to simulate a section of pipe of diameter 1 and length 4. After initializing the velocity field we allow the system to evolve for a sufficient number of pipe transits so that the system is fully turbulent (one pipe transit corresponds to four units of time as the mean flow velocity is set to 1 in the simulation units). The smaller the roughness, the more transits are needed. This results in roughness-generated turbulence, in which case the observed energy spectrum is dominated by the inverse cascade, as shown in Fig. (5.5).

In order to attain an enstrophy-dominated flow, we used a technique suggested by the observations reported by Rutgers[37]. We simulated grid-generated turbulence, by placing a series of cylinders at the mouth of the pipe; in each cylinder we set the velocity field to zero every timestep. After one pipe transit the velocity field is fully developed. We then remove the grid and allow the turbulence to decay for a transit before we begin to measure the friction factor and other flow properties. We have observed energy spectra dominated by the enstrophy cascade in this system, as shown in Fig. (5.5).

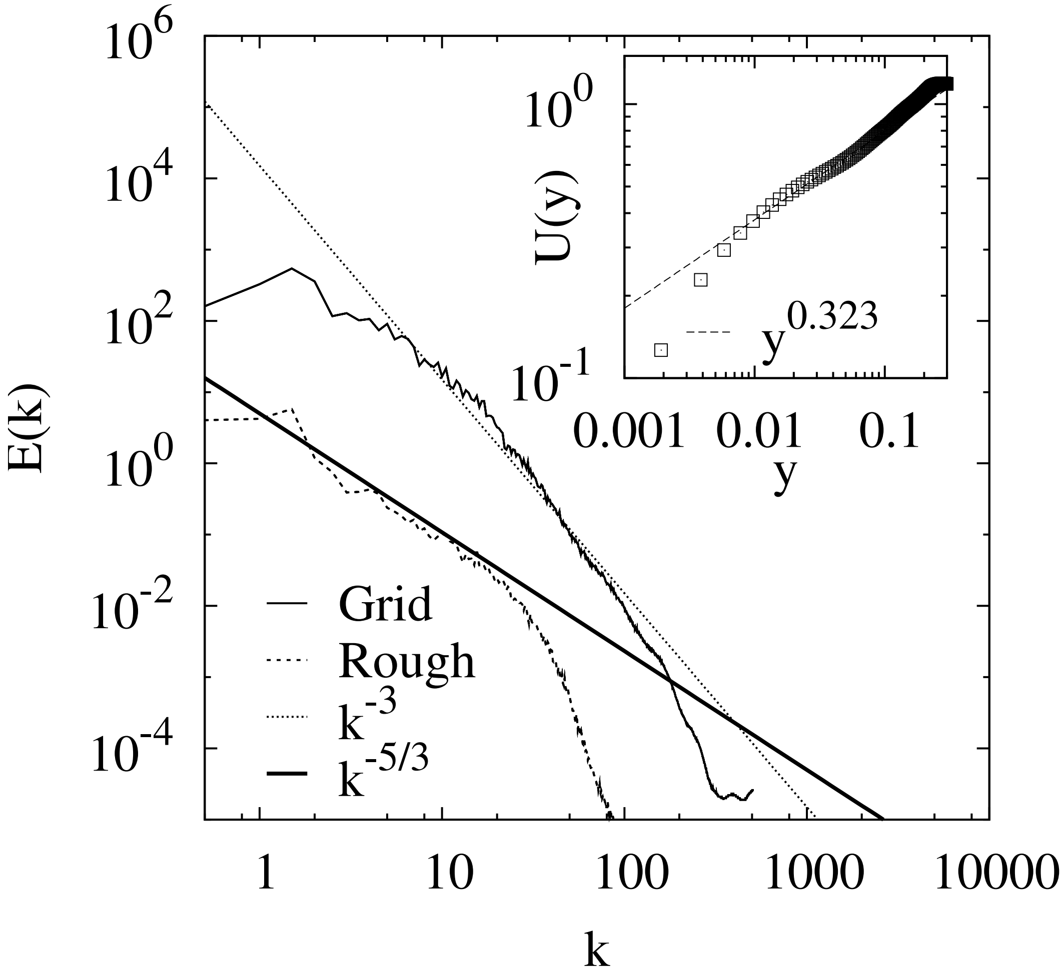

Figure 5.5: Energy spectra for grid- and roughness- generated turbulence. Grid-generated turbulence exhibits the k^{-3} enstrophy cascade, compared to the $k^{-5/3}$ inverse cascade scaling of roughness-generated turbulence. Inset: simulated wall velocity profile of grid-generated turbulence in a smooth pipe at $\mathrm{Re} = 60000$. The profile is consistent with a power law with exponent 0.323 ± 0.005. We predict an exponent of $1/3$ for enstrophy cascade turbulence.

Chapter 6

Two-dimensional Pipe Flow Results

In this chapter we will present the results from our simulations of two-dimensional pipe-flow turbulence. We examine the friction factor scalings and velocity profiles for both the inverse cascade and enstrophy cascade energy spectra. We find that the predictions of momentum transfer are supported for the friction factor scalings and for the enstrophy cascade velocity profile, but that there are inconsistencies from the predicted scaling of the friction factor fluctuations and the inverse cascade velocity profile.

We have examined a range of Reynolds numbers from 1000 to 80000, with roughnesses (defined by the parameter r in our conformal map) between 0.05 and 0.2. In these ranges, we see primarily the laminar and Blasius regimes, and the very onset of the Strickler regime. This is sufficient to examine the possibility of data collapse with respect to the predicted Strickler scaling.

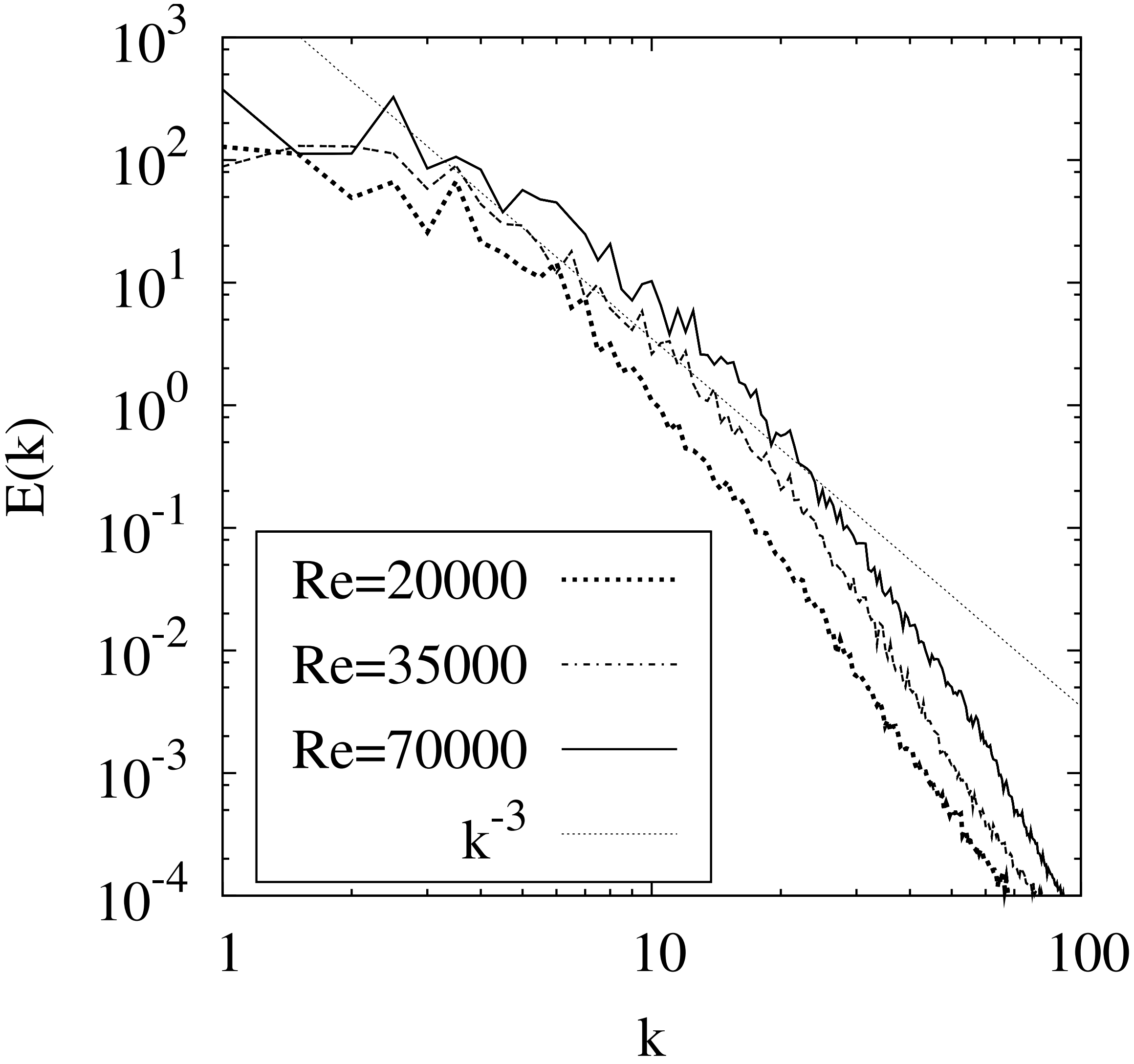

Figure 6.1: Energy spectrum of grid-generated turbulence at three Reynolds numbers. The solid line indicates k^{-3}. At higher Reynolds numbers, the range of the k^{-3} scaling is obeyed increases.

6.1 Energy Spectra

Here we discuss in detail the measurements of the energy spectra (the energy contained in an infinitessimal range of wavenumbers between k and $k + dk$) in our two-dimensional simulations. In order to measure the energy spectrum, we relate it to the velocity field using Eq.3.17. We can obtain the longitudinal component of $E(k)$ by taking the power spectrum of the Fourier transform of the longitudinal velocity along the pipe at specific distances from the wall. The transverse component is not well defined in this flow as the flow is not periodic or infinite in extent in directions parallel to the wall. We can therefore see that the energy spectrum retains its form even close to the wall, as required by the momentum transfer theory.

For roughness-generated turbulence, we see a spectrum with a $k^{-5/3}$ range, a sharp fall-off at large k, and the far edge of the integral range (Fig. 5.5). It is questionable whether the behavior at large k is the enstropy cascade or the dissipative range, or some combination. It is significantly steeper than the k^{-3} scaling we would expect of the enstrophy cascade, however. It is closer to the k^{-7} or exponential scalings that have been predicted for the dissipative range. Furthermore, we see that as we change the Reynolds number, this range moves outwards as we would expect for the dissipative range.

For grid-generated turbulence, we see both an enstrophy cascade and a small section of inverse cascade (Fig.6.1). As the friction factor scaling in the smooth pipe is dominated by the smallest scales, this small inverse cascade contribution should not influence the Blasius scaling. When we combine grid-generated turbulence with roughness, however, the inverse cascade range expands and dominates the spectrum. Consequently, we cannot immediately measure the roughness scaling in the enstrophy cascade unless we can prevent that from happening.

We find however that if we add a noise term to the Navier-Stokes equation at the smallest scales, this enhances the enstrophy cascade and sustains it in the presence of roughness. We use this trick to measure the dependence of the friction factor on roughness in the enstrophy cascade.

6.2 Friction Factor

Our simulation results at small values of the dimensionless roughness ($r/R = 0.067$) are plotted in Fig. (6.2). These results were obtained by averaging over 5 full pipe

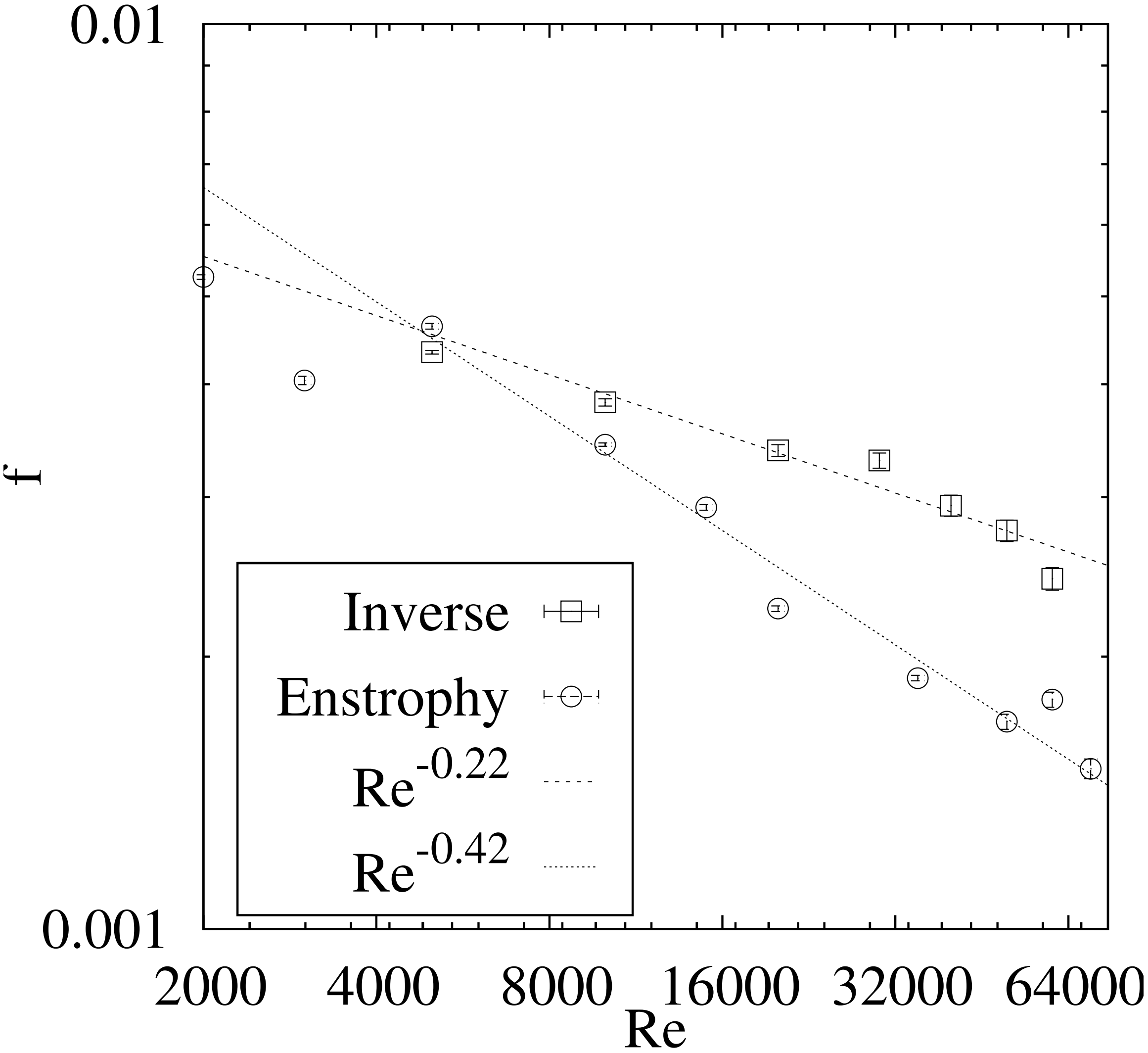

Figure 6.2: Scaling of the friction factor with respect to Re for inverse cascade and enstrophy cascade dominated flows in 2D. The roughness is $r/R = 0.067$, and the data have been averaged over a time of 5 pipe transits.

transits, yielding reproducible values for the friction factor, with error bars determined from bootstrap resampling of the data using 1000 replicates each consisting of 50% of the timeseries data. For this flow we observe an approximate power-law scaling of the friction factor with Reynolds number, with an exponent -0.22 ± 0.03 (error bar determined by the uncertainty in the nonlinear fit) together with an energy spectrum dominated by the inverse-cascade. In the case of grid-generated decaying turbulence, corresponding to an enstrophy-cascade dominated spectrum, we observe an exponent of -0.42 ± 0.05. These results are within satisfactory agreement with the scalings of

$-1/4$ and $-1/2$ respectively, predicted for the 2D Blasius regime on the basis of a momentum transfer argument.

We cannot reach sufficiently high Reynolds numbers to observe a pure Strickler regime, but we can verify the Strickler scaling exponent with data collapse. In three dimensions, or in a system dominated by the inverse cascade, we expect data collapse when plotting $f\mathrm{Re}^{1/4}$ against $(r/R)\mathrm{Re}^{3/4}$[1]. For the enstrophy cascade, these variables should be $f\mathrm{Re}^{1/2}$ and $(r/R)\mathrm{Re}^{1/2}$ respectively. We have observed previously that in the presence of roughness, the spectrum is dominated by the inverse cascade. However, we have found that by adding a small amount of random forcing to the velocity field, the enstrophy cascade may be observed even in a rough pipe, though it may be coexistant with an inverse cascade. Using this method we can obtain the roughness dependence of the friction factor in an enstrophy cascade dominated flow. The collapse of the friction factor curves using the enstrophy cascade variables is shown in Fig. (6.3). The collapse is quite good, despite an apparent shallowness to the Blasius regime in the raw data. This shallowness is likely caused by the presence of a small amount of roughness, modifying the expected $\mathrm{Re}^{-1/2}$ scaling at larger Reynolds numbers. We have neglected intermittency, which may be negligible or at least very small in 2D compared to 3D[81].

6.3 Velocity Profile

Following Prandtl[44], we have calculated the mean velocity profile $u(y)$ as a function of distance from a wall y, and for the enstrophy cascade this yields $u(y) \sim y^{\alpha}$ with $\alpha = 1/3$, corresponding to the Blasius regime. For a general conserved quantity, $\alpha = (1-\phi)/(3-\phi)$. This relation depends on the zero roughness limit. In [82], it has been shown that roughness modifies the velocity profile so as to increase the apparent scaling exponent. Other work[83; 84] also considers the influence of rough walls on the velocity profile and near-wall scaling.

In our simulations of smooth-pipe enstrophy cascade turbulence, we have measured the velocity profile and found the power-law scaling exponent $\alpha = 0.323 \pm 0.005$ between $0.01R$ and $0.1R$, as shown in the inset of Fig. (5.5), close to the predicted $\alpha = 1/3$. In the case of our rough-pipe simulations, the velocity profile yielded an exponent of 0.333 ± 0.002, significantly steeper than the predicted $\alpha = 1/7$ that applies in the smooth, inverse cascade case. Our interpretation is that this is due to spectral contamination from an enstrophy cascade, as in the case of the simulations with random forcing that we presented. The momentum transfer theory integral has an upper limit that is comparable with the Kolmogorov lengthscale at low roughness, and so in that case the small-k part of the energy spectrum controls the friction factor scaling. Because of this, we would expect to see a velocity profile consistent with the enstrophy cascade until the roughness or Reynolds number were high enough to place the crossover between the inverse cascade and contaminant enstrophy cascade below the scale of the roughness.

Our results for the power-law Blasius regime in 2D enstrophy-dominated turbulence show convincingly that this regime is more than an empirical fit, and has a dynamical significance. The fact that the friction factor scaling changes when we change the spectrum of our flows further supports the connection between the scale-free structure of turbulence and its macroscopic transport properties. We additionally see that the velocity profile in two-dimensional enstrophy cascade turbulence behaves differently from the three-dimensional case, as we predicted in Chapter 3.

Two-dimensional turbulent pipe flow is experimentally accessible in the form of gravity-driven soap film flows. Experiments by Tran, et al.[71] have investigated the behavior of the velocity profile and the friction factor in such soap films. They observe an enstrophy cascade spectrum, and friction factor scalings consistent with the predictions of the momentum transfer theory. Additionally, they observe a von Kármán constant that is significantly lower than the three-dimensional value, consistent with our predictions.

These results support the fundamental connection between spectral structure and friction factor scaling, which is manifested in the observed roughness-induced criticality.

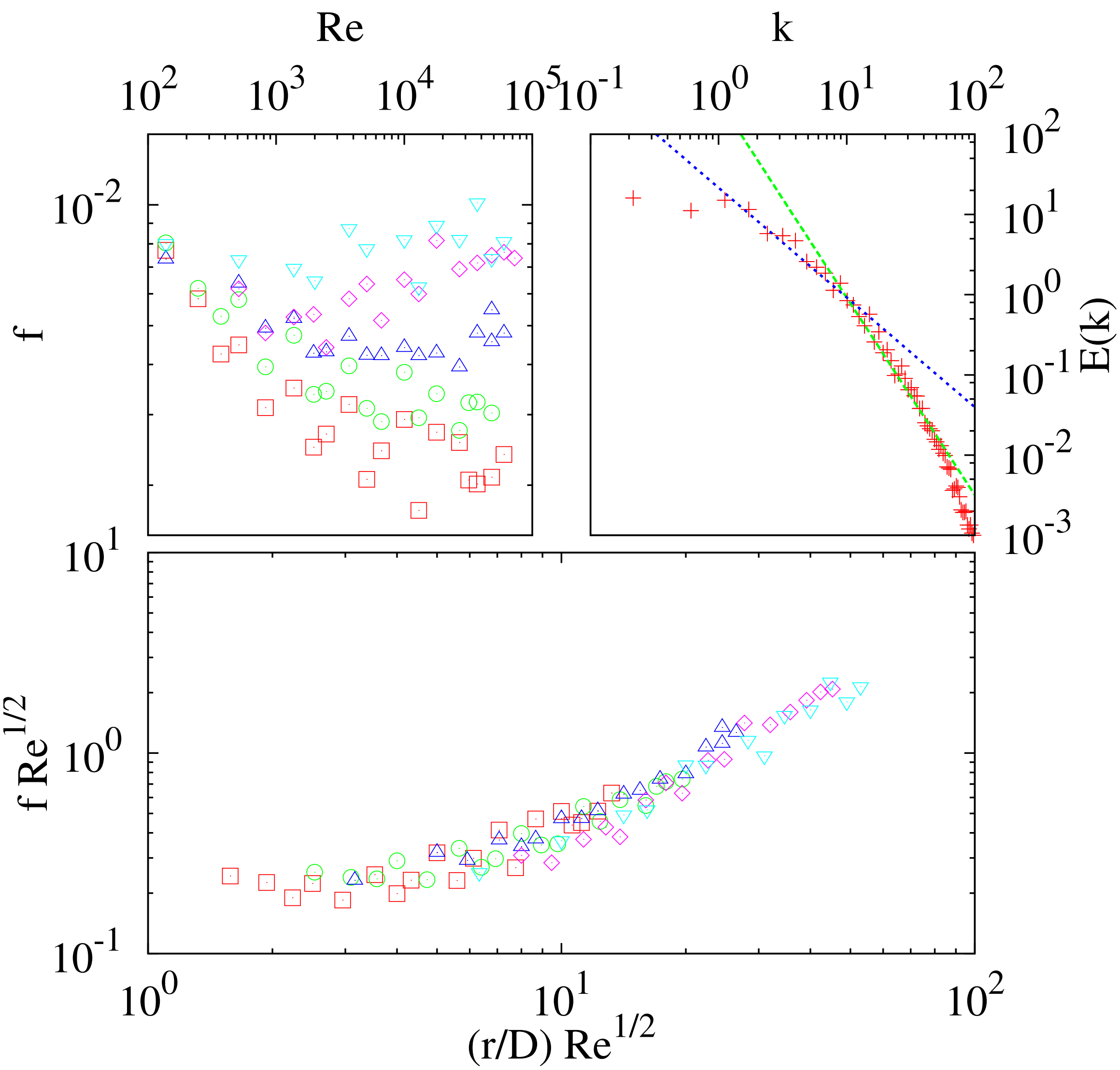

Figure 6.3: The bottom inset shows the enstrophy cascade data collapse of the friction factor curves for nondimensional roughness 0.05 (○), 0.08 (□), 0.1 (△), and 0.2 (▽) over a range of Reynolds numbers from 1000 to 80000. The top left insent shows the unscaled friction factor data. The top right inset shows the energy spectrum at $r/R = 0.08$ and Re = 80000. The straight lines correspond to $k^{-5/3}$ and k^{-3}.

Chapter 7

Biological Complexity

Biological organisms exhibit many different levels of organization and function. Even at the level of single-celled organisms, there are layers of structure — proteins with direct functional consequence such as structural proteins and enzymes, regulatory proteins that control the expression rates of other proteins, ribosomal RNA to translate from the encoding of information in DNA to the actual expressed proteins, and (particularly in biofilms) emergent behaviors such as colony-wide structures and cell differentiation[85]. In multicellular organisms, there are further levels of organization, some of which are clearly advanced forms of what is seen in biofilms. Differentiation between somatic cells and functional cells occurs. Organisms are macroscopic, and interact with the world mechanically rather than primarily chemically. Furthermore, secondary levels of information storage appear in the form of neurons, which encode information taken from the environment during the lifetime of the organism and allow for a faster adaptation to changing conditions.

The lengthscales across which living (in the sense of being self-replicating and of undergoing evolution) structures exist span nine orders of magnitude, from viruses at a scale of 10 nanometers to the largest ocean-dwelling mammals at 10 meters. A detailed explanation for this wide range of coexisting scales has not been given. In principle, evolutionary processes will allow organisms to find optima for survival and fast replication in their given environments, but why would macroscopic organisms with orders-of-magnitude slower replication times and orders-of-magnitude higher nutritional requirements for survival be able to coexist with the smallest bacteria, which easily out-replicate the larger organism?

7.1 Separation of Scales

I propose that the answer to this is a separation of scales arising from predominantly local interactions in the space of organismal complexity and size. This picture of local interactions can lead to a cascade in which the complexity of organisms in the system increases even at the cost of decreasing the total fitness. One might suggest that the rules should be different for different kinds of organisms — why would insects, mammals, and avians share the same scaling relationships? However, other scaling relationships in multicellular organisms have been shown to hold across multiple phyla[86]. In multicellular predator-prey systems, one can use a scaling argument to support the idea that this sort of locality exists. Organisms that are much smaller than a given predator don't provide enough nutrients to be worth pursuing, and organisms much larger are impossible to consume. As a consequence, there is a narrow range of scales in which predation is efficient.

We will begin by considering a predator of scale l and prey of scale s. This predator cannot effectively eat things much larger than itself, though it may eat larger things within the same order of magnitude of scale. The amount of food that the predator requires to survive is proportional to its body mass, and so scales as l^3. Furthermore, the food provided by a given prey species scales as its body mass s^3. If we consider that the distance that a given predator may range to find food scales as its length, then a predator that lives in d dimensions covers a volume l^d when searching for prey, where the density of prey in that volume will scale as $1/s^d$. Consequently, the efficiency ϵ of a given predator in acquiring food by pursuing prey of scale s is:

$$\epsilon = (l/s)^{d-3} \tag{7.1}$$

For predators existing in a three-dimensional environment, it appears that the size of prey is irrelevant under these assumptions. This is consistent with the observation of very large oceanic creatures that eat many many very small creatures - whales and krill for instance. However, within two dimensions, such as predators ranging on land, the efficiency scales linearly with the size of the prey species s. This means that the most efficient choice is to specialize in eating something on the same scale as itself, and so one can predict that on land there should be a cascade of organisms at many scales, each eating things of similar scale to itself.

While this argument may work for very late-appearing forms of life, it is necessary to explain the separation of scales without resorting to multicellular, land-bound creatures. The first large change of scale corresponds to the formation of biofilms and multicellular plants and fungi, which do not have predator-prey interactions. However, these systems enjoy direct advantages from cooperative effects compared to isolated unicellular organisms. Biofilms can establish local chemical environments that are more efficient for metabolism and other cell functions[87], and protect the cells from external fluctuations via an exterior layer of structural insulating cells. The improvement in replication due to higher cell density is observed across many unicellular systems, not only biofilms, and is known as the Allee effect[88; 89].

On the other hand, biofilms are less efficient in times when environmental food supplies are short. Individual cells have higher motility and can cover a larger volume in search of food compared to a biofilm which is fixed in place. As such, it is understandable that simple quasi-multicellular forms would coexist with unicellular organisms (and in fact, that biofilms would use regulation to switch between cohesive behavior and unicellular behavior depending on the availability of food or presence of environmental stresses).

Though replicating structures smaller than a single-celled organism exist (viruses), they are all parasitic in nature, and so require the existance of single-celled organisms in order to replicate. There seems to be a minimal scale necessary for efficient self-replication without recourse to external replication machinery. This explains the lower cutoff of the range of scales observed.

There are putative precursors to cellular organisms - self-replicating metabolic cycles[90], and the RNA world hypothesis[91]. It has been observed that DNA on its own can replicate without the intervention of enzymes in suitable environments[92]. These forms of replication seem to all be much slower than cellular replication, however. Encapsulation allows for chemical concentrations much higher than would exist in an open environment. So while early life may have bootstrapped its way up to cells, it is unsurprising that all currently known life is either cellular or parasitizes cellular life.

7.2 Complexity

Beyond the issue of multiple coexistant scales in existing life, there is the question of the reason for the observed complexity of life and how it may be explained. This issue is at first apparently less troublesome, as there are many possible sources for the complexity of actual life. The laws of physics and chemistry, the environment, and the many scales of fluctuations all suggest that the complexity of life may simply be the complexity needed to be an optimal replicator — that is, an organism that fills up its local environment with copies of itself as quickly as possible. This argument breaks down when one notices that there are also many different scales of complexity in living organisms. Simply looking at the active component of genome content, it varies across six orders of magnitude[93].

We will attempt to address this by looking at simulations of artificial organisms. In simulations, the inherent complexities of chemistry, physics, and the fluctuating environment are not by default present. It is also of interest to determine what is necessary to have open-ended evolution in such simulations, in order to drive evolutionary algorithms and explore possibilities in artificial intelligence. An intelligent system and a system with open-ended evolution share the property that they are both self-directed, and so perhaps insight in one version of the problem could lead to insight in the other.

In general, no computer model can hope to capture the actual complexity of the chemistry and environment that real organisms experience. Certain things will be very difficult to predict through model building due to this inherent background complexity. On the other hand, many behaviors of systems in the real world stem from extremely simple underlying mechanisms, such that even when details are omitted, these mechanisms remain more or less unchanged. We propose that open-ended evolution is qualitatively different from evolution towards a fixed point and neutral evolution, and that this qualitative difference stems only from a small set of necessary elements. By strictly controlling what goes into the model, the minimal set of necessary elements to have self-replicating organisms that evolve in an open-ended fashion may be discovered. In this sense, the idea is to build a series of minimal models to isolate the root causes for different evolutionary behaviors.

The earliest attempts to model artificial self-replicating life may be the cellular automata of John von Neumann[94]. He described a way in which a self-replicating system could be built via the implementation of a universal constructor. He initially designed a set of rules for the transitions of sites on a two-dimensional grid that would give rise to such an object, though he could not get it to copy itself entirely. Much later, these rules were improved such that a self-replicating universal constructure could be built[95]. Evolution in this system was possible with the addition of randomness to the rules, but the eventual behaviors were strongly bounded by the choice of rules.

The question of whether or not artificial replicating systems would complexify has been broadly considered [96]. While the potential to complexify is present in almost all such artificial life simulations, baring those whose rules are so simple that only replication and some fixed functionality is possible, whether or not it actually occurs is another issue entirely. The rules of the von Neumann cellular automaton were comparatively simple, but the structure itself was very rigid as far as what changes could be made that would not destroy its function entirely.

Following from the basic idea of computation as life, other attempts at artificial life were made. One approach to achieving a rich set of behaviors is to create a very flexible and general set of rules for the mapping between synthetic organisms' genetic code and their dynamics, which abstracts away any sort of biochemistry. The organisms then interact and evolve within the artificial ecosystem represented in the computer according to these simplified rules. The Tierra[97] and AVIDA[98] projects use computer languages as a way to produce a general set of rules. Given a sufficient minimal set of operations, a computer language is Turing complete and can emulate any other such computer language. As a result, any behavior can in principle be developed by the artificial organisms. One problem is that many of these behaviors may be more or less difficult to evolve depending on the specifics of the language chosen to represent the dynamics.

In the Tierra model it became evident that the dynamics were not neutral with respect to the size of replicating programs. Evolutionary pressure favored programs becoming smaller as they would then replicate more quickly and out-produce the larger programs in the system. This led to the development of interesting parasitic behavior in which a program would use a neighbor's replication code to decrease its length[97; 99]. Though corrected in later work, this was insufficient to obtain open-ended complexity growth. Work by Standish[100] analyzed the complexity of organisms in Tierra and attempted to pinpoint the cause for the eventual limit in organismal complexity; but the results were inconclusive.

It is not necessarily the case that the artificial organisms need to be Turing complete in order for the simulation to produce interesting evolutionary behaviors. A minimal model can be more informative as it is easier to control representational effects, and the causes of specific features of evolution can be isolated. An example is the work of Vetsigian and Goldenfeld[101] in studying the effects of homologous recombination. The result was that certain structures were predicted to occur in the genomes of bacterial populations (and were reportedly observed by comparative genomes), simply as a result of the genetic mechanisms, independent of any sort of selective pressure.

In the following two chapters I will present my work in making abstract minimal models that can be simulated on a computer over evolutionary timescales. The models implement the genetic operations extant in real biological systems, but the interactions between organisms and the environment in the system are highly simplified. The goal is to find the necessary structure of a simulation in which the complexity of organisms has open-ended growth.

There are two features necessary to accomplish this. The first is that the evolutionary dynamics must have an invariance with respect to changes in the complexity of the evolving organisms. That is, if there are inhomogeneities which encourage organisms to have a specific complexity, then that will act to prevent the complexity of the system from continually increasing. This feature is related to the origin of the Kolmogorov cascade in turbulence — in the turbulent cascade, a hierarchy of length-scales exists due to a transport of energy by scale-invariant processes between a large length scale and a small length scale. In this case, processes invariant to changes in complexity will produce a hierarchy of complexities in the system. An overview of the effects of different genetic operations will be presented with respect to this invariance criterion. This criterion can also apply to the way that the fitness of an organism is determined in the dynamics, either explicitly or implicitly.

The second feature is that there must be some advantage which can only be gained by an organism in the system being more complex than the average population. This feature occurs naturally in competitive interactions. This has the same function as viscosity in turbulent flows. That is, it sets the directionality of the relevant transport.

7.3 Abstraction Transitions

Once the open-ended evolution of complexity is established, there is still the question of the evolution of qualitatively different features. It is in principle possible for a given species to evolve a high complexity in a single facet of its interaction with the environment or with other organisms. That is, a pair of species could for instance develop more and more sophisticated poisons and immunities in an attempt to compete for space. However, this is fundamentally different than, for instance, one of the species developing some new mode of interaction that allows it to access or even create new ecological niches.

One point of view is that these qualitatively different modes of behavior are simply due to the existance of unexplored niches, and that they are eventually limited to the total number of niches that the underlying physics creates. However, this leaves out the issue of niches that exist solely due to the presence of organisms in the system, such as things that consume the waste products of other organisms. The structure of niches themselves can fluctuate due to what the organisms in the system are doing.

In a broader sense, however, there are certain qualitative changes that fundamentally change the way evolution itself works. The transition from pre-genetic life to organisms that stored their template in DNA and then made use of the translation apparatus to convert that stored information into functional proteins is an example of one such change. Before the existance of a polymer with an invariance that allowed substitution of different monomers at different sites without changing the polymer's behavior, any hereditary information in the system must have been distributed somehow across the varying chemical concentrations of the environment. It was not necessarily the case that changes could be made in a way that they were inherited, nor that many pathways were open to evolution. Even if one considers the RNA world picture[91], changes to the ribozymes that would catalyze their own self-replication would have immediate impact on their ability to replicate, and so it would be hard to evolve secondary ribozymes that would aid replication but would be insufficient on their own.

Another example is the evolution of gene regulation. Without a regulatory apparatus, a given organism would always have one set of behaviors, and so would be unable to use environmental information to change its mode of operation without building it into the protein chemistry of the cell directly. By attaching complexes to DNA such that certain regions could be activated or repressed, it became possible to change the cellular state in response to external stimuli. This also provided a new layer of information in the genome, and a new mode of evolution. Now, rather than constructing a certain function by producing the necessary proteins to enact it, functions could be produced by having timed combinations of proteins. This again changed the way that evolution operated.

In studying the evolution of the translation apparatus, Carl Woese pointed out that "evolution becomes an expansion of an informational universe"[102] as a way to understand the new degrees of freedom accessible due to the development of translation-based protein synthesis. I propose that these transitions share a universal structure, though the details of the transitions will be different each time. Each of these transitions consist of the development of a new level of information upon which evolution can operate in the sense of Woese's expanding "informational universe", such that the new level is partially decoupled from the lower level (in the example of regulation, evolution can operate on the proteins, or the regulatory network that relates their expression levels). As such, there is a degree of abstraction of the lower level that is similar to the emergence of quasiparticles in many physical systems. I will refer to these transitions as abstraction transitions.

The actual operation of evolution at a given level requires more machinery than simple point mutation to be efficient. Mechanisms such as error correction[103], homologous recombination[104], and illicit horizontal gene transfer[105] vastly speed up and stabilize the operation of evolution[106; 107]. When a new level of evolution becomes available, these mechanisms, such as they exist at the current level, will not necessarily function in a way that is efficient for the new one. Thus, in these transitions, the evolutionary efficiency of the system as far as phenotypic changes is predicted to be low at first. In order for the transition to be viable, pressures that allow the evolution of new evolutionary mechanisms must be present.

When there is competition between something immediately beneficial and something that only has benefits several generations down the line, evolutionary pressure operates more strongly on the thing with immediate benefit unless some other time or space scale exists in the system. However, if there is a timescale in the selection pressure much longer than the lifespan of an individual organism, or if there is larger spatial organization or collective effects that allow organisms to depend on each other for their survival, this can change the balance of evolutionary pressure and put emphasis on longer-term benefits such as the development of new evolutionary mechanisms. As such, we would expect to see the seeds of abstraction transitions in environments with long timescale fluctuations or places where collective effects have produced groups of mutually supporting organisms.

Looking at the various forms of life on Earth, there are several times where an abstraction transition may have occured in addition to the origin of DNA as an information-bearing molecule and the emergence of regulation. In the origin of multicellular organisms, the formation of spatially-varying structures became possible - the new layer of evolutionary information in this case is cellular differentiation. In multicellular organisms, the development of neurons allowed information to be stored and processed on a faster timescale than the replicative timescale. The mechanisms of learning then comprise the new evolutionary level. Once nervous systems are in place, organisms can develop communities, pack structure, and the like. With communication and later language providing the mechanism for information exchange and heritability, this leads to the formation of societies and culture as the next evolutionary level.

Chapter 8

Foodchain

Experiments on digital organisms represent one of the most accurate and informative methodologies for understanding the process of evolution[108]. Systematic studies on digital organisms are especially informative, because the entire phylogenetic history of a population can be tracked, something that is much more difficult—but not impossible[109]—to do with natural organisms. Experiments on digital organisms can be performed over time scales relevant for evolution, and can capture universal aspects of evolutionary processes, including those relevant to long-term adaptation [110; 111], ecological specialization[112; 113] and the evolution of complex traits[114].

Despite this progress, the way in which evolution leads to ever increasing complexity of organisms remains poorly understood and difficult to capture in simulations and models to date. Is this because these calculations are not sufficiently realistic, extensive, or detailed, or has something fundamental been left out? In this Chapter, we argue that two fundamental aspects of evolutionary dynamics, with the character of symmetries, have been omitted, thus causing complexity growth to saturate.

The first feature is that the evolutionary dynamics must be invariant with respect to changes in the complexity of the evolving organisms. That is, if there are inhomogeneities which encourage organisms to have a specific complexity, then these will act to prevent the complexity of the system from continually increasing. This invariance is similar in spirit to that which lies at the heart of the Richardson cascade in turbulence[9; 11]. Here, a hierarchy of length-scales exists due to a transport of energy by scale-invariant processes between a large length scale and a small length scale. The largest and smallest features of the flow are determined by where the invariance is broken. In the biological case, processes invariant to changes in complexity will allow the dynamics to produce structures of arbitrarily high complexity. We will see below, in an explicit model, the effects of different genetic operations with regard to this invariance criterion. This criterion can also apply to the way that the fitness of an organism is determined in the dynamics, either explicitly or implicitly.

The second feature is that there must be some advantage which can only be gained by an organism in the system being more complex than the organisms it competes with. Competitive interactions can drive such a dynamic; for example, if competition can be thought of as one organism setting the environmental problem that the other organism must solve. The resulting co-evolution favors an increase in complexity over a decrease, because for the problem-setter, simplifying the problem does not exclude an organism already able to solve the problem. This factor has the same function as viscosity in turbulent flows: it sets the directionality of the relevant transport.

These two features have precisely the same mathematical role in evolutionary models as the mechanisms of energy transfer and viscous dissipation do in fluid turbulence. Thus, the open-ended growth of complexity in our model, and the existence of a hierarchy of structures at all scales in turbulent flows are mathematical consequences of the same underlying dynamics. It is not important for this argument what is the direction of energy flow in the turbulence case: in fact, the direction depends on dimensionality, with the possibility of the accumulation of large-scale structures in two-dimensional turbulence through the so-called inverse cascade.

The implications of this dynamical systems argument are far-reaching, and impose constraints on how digital evolution models should be built. For example, despite its popularity, the "fitness landscape" [115–117] picture of evolution does not satisfy these constraints, and is conceptually insufficient to account for the open-ended growth of complexity. To illustrate our points, we now show how open-ended growth of complexity emerges from underlying dynamical rules in a simple caricature of an evolving ecosystem.

8.1 Complexity Saturation in Digital Ecosystems

Tierra[97] and Avida[98] are systems of digital organisms, which are represented as self-replicating programs in a Turing complete language. In principle any program or behavior can then be encoded with a sufficiently large genome. In Tierra, organisms exist in a linear space for which each point in space is associated with an instruction and replication occurs via a loop which copies the contents of the space at an offset. In early work on the Tierra model it became evident that the dynamics were not neutral with respect to the size of replicating programs. Evolutionary pressure favored smaller programs as they replicate with fewer instructions and out-produce the larger programs in the system. This led to the development of interesting parasitic behavior in which a program would use a neighbor's replication code to decrease its length, i.e. the complexity of organisms did not increase. When this was corrected by a change in the way in which resources were allotted, the length of organisms was observed to increase in bursts, but eventually saturated for longer and longer intervals[97], a finding attributed to insufficient richness of the environment[100].

In Avida, there is a two-dimensional grid, each cell of which contains a program, and replication occurs between cells. Selection is based on an organism's ability to solve a particular mathematical problem. Avida uses an information-theoretic definition of complexity based on the information learned by the organism from its environment[114]. For evolution occurring in a single niche, it is found that this complexity increases for some time, then saturates to a value determined by the maximum information associated with the niche (the potential complexity) [118].

A similar pattern of complexity saturation is found in 'WebWorld'[119–122]. Here, species are described by a set of features that may be either present or not, and the total rate of predation between species is determined by summing over a random interaction matrix for each feature possessed by the predator and each possessed by the prey. The total number of features possessed is found to increase in the presence of interactions above the neutral case. However, the increase in complexity is eventually limited by the predefined set of features, there being no possibility of creating new features in the model.

In summary, these and other digital ecosystems appear to lack the drive to increasing complexity that arguably is present in real biological systems.

8.2 The Model

We now present an abstract minimal model of an evolving predator-prey system, which we call "Foodchain". This model exhibits the potential for an open-ended growth of complexity. Organisms in this model exist in a two-dimensional space and interact with each other. The detailed mechanics of replication are abstracted away (unlike Tierra and Avida)—during replication, genetic operators (point mutation and gene duplication) are applied to the genomes, which are of fixed length 2048, to produce the genome of the offspring. In 'Foodchain', fitness is determined solely by interactions between organisms, as they attempt to eat a random neighbor each timestep. A certain amount of energy is introduced to each living organism every time step, and replication occurs when an organism has an adjacent empty grid cell and a sufficient amount of energy.

Each organism has a fixed-length string of letters as its genome. These letters can be upper or lower case, so that each letter is one of 52 possible letters. All but eight letters are inactive and do not influence the interactions between organisms. Of the eight active letters, four are offensive (A,B,C,D) and four are defensive (a,b,c,d).

The predator-prey interactions are determined by organisms' genomes. A particular organism is not predisposed to be predator or prey, and may even be able to eat its own offspring. The comparison between genomes consists of matching contiguous substrings of offensive letters in the organism attempting predation with defensive letters in the prey. If the predator has a sequence of offensive letters that is not matched in the prey by a corresponding defensive string, the prey organism dies and the predator gains a percentage of its energy. A neutral letter or letter of a different type ends a sequence.

This interaction rule satisfies the condition that fitness in the system should depend only on relative quantities as well as the condition that in interactions between different complexities, higher complexities produce a benefit for the organism. If a particular organism only has a defensive string of length L, then a predator with an offensive string of length at least $L+1$ will always be able to eat it; thus there is always a structure at a higher complexity which can bypass a particular defense.

When an organism replicates, its genome is subject to change from mutation and other genetic operations. Point mutations occur at a rate r_m per letter and set the mutated letter to a random letter, which may be the same as the original. Gene duplication occurs at a rate r_d. In gene duplication, three random values between zero and the length of the genome are generated: a start position i_{start}, ending position i_{end}, and an offset i_{ofs}. The sequence between i_{start} and i_{end} is stored in memory and written back into the genome starting at $i_{start} + i_{ofs}$. The genome is treated as being periodic as in microbial DNA, so if $i_{end} < i_{start}$ the reading process proceeds through the end of the genome and wraps around to the beginning.

In this system the complexity is taken to be the longest functional string (separated into attack and defense complexities). The motivation for this choice is that it is directly related to the capabilities of the organism. It also represents the interaction between pieces of information in the organism's genome: together, a sequence of multiple letters have a certain functionality that, apart, they would not.

Point mutations do not satisfy the condition that the dynamics should be invariant to changes in complexity. If an organism has a particular active string of length L, there are L chances for a point mutation to decrease the complexity, and 2 chances for a point mutation to increase the complexity. More specifically, if a mutation occurs at the first letter before or after the string, there is a 1/13 chance that the length of the active string increases by 1. If a mutation occurs anywhere within the string, there is a 12/13 chance that the active length will decrease. The average resultant length L' of an active string initially of length L after a single point mutation is given by:

$$\langle L' \rangle = \frac{3}{4}L - \frac{1}{2} - \frac{1}{4L} \tag{8.1}$$

The dynamics of point mutations tends to decrease the active length because there are many more ways to decrease it than to increase it. This entropy pressure competes against the selection pressure due to the advantage that results from having a sequence of higher active length. The magnitude of the advantage, and thus the selection pressure, is independent of the absolute sequence length, whereas the entropy pressure scales with the sequence length. Therefore, there is an equilibrium active string length (complexity) at which the entropy pressure is balanced against the selection pressure.

Gene duplication on the other hand operates equally on sequences of different active lengths so long as the active length is much smaller than the total genome length. The probability that the gene duplication region cuts a sequence of length L is L/L_{genome}. If a particular sequence is captured, its length will at least be preserved and may increase by an amount proportional to the average sequence length in the organism if the write region is adjacent to another sequence of the same type.

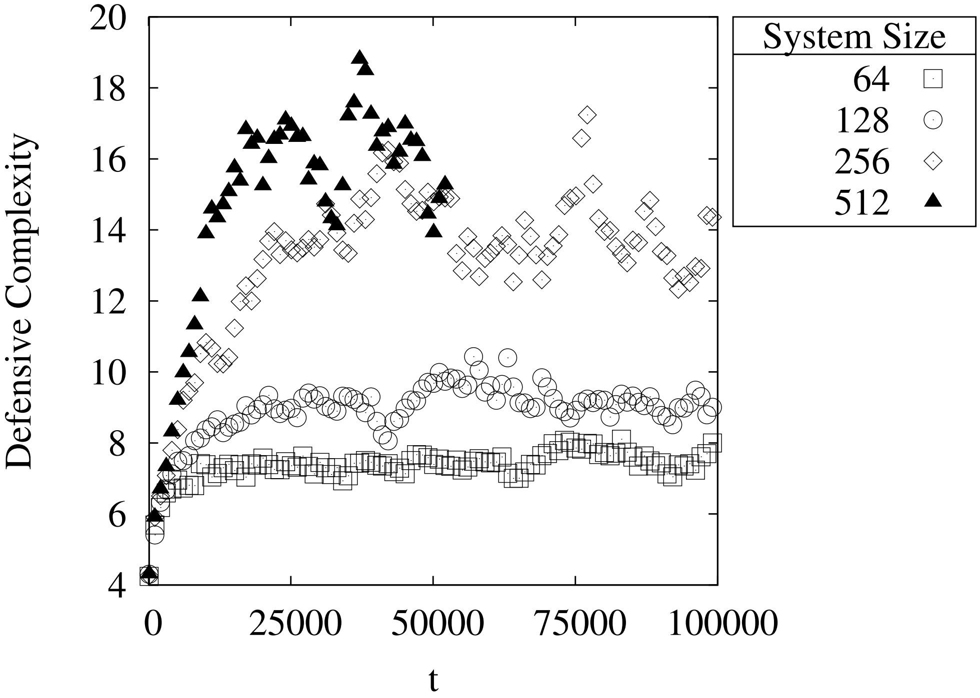

Figure 8.1: Defense complexity versus time in Foodchain for system sizes 64, 128, 256 and 512 square grids. Duplication rate is set to 0.1 and mutation rate is set to 0.01.

Point mutations are necessary to fully explore the genetic space, but if the point mutation rate is too high, the complexity cascade is inhibited. The next section examines the results of simulations for a variety of point mutation rates and system sizes in order to probe this effect.

Every hundred timesteps the system-wide population, average energy, average attack complexity, and average defense complexity are stored for analysis. The attack and defense complexities are taken to be the longest contiguous string of attack and defense functionality. The simulation is run for different initial random seeds in order to extract the mean behavior of these quantities with simulation time.

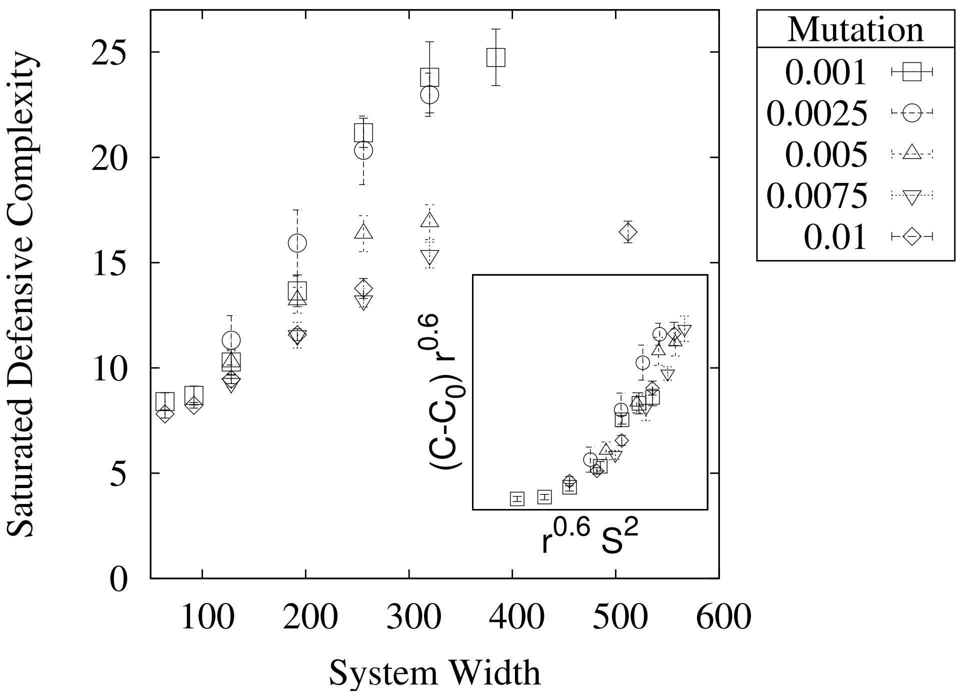

Figure 8.2: Dependence of maximum defensive complexity on system size and mutation rate. The inset shows that the data collapse onto a single curve when plotted with a dependent variable $(C - 6.65)r^{0.6}$ and independent variable $r^{0.6}S^2$.

8.3 Results

The average defensive complexity of organisms in the system as a function of time is plotted in Fig. 8.1 for different system sizes. These simulations use a gene duplication rate (per replication) of 0.1 and a mutation rate of 0.01 per letter. The complexity increases with time for short times, but then saturates at a value which depends on the system size. We observed that in a system with no gene duplication, the increase in complexity was logarithmic with time, whereas the system with gene duplication exhibited super-logarithmic complexity growth. Increasing the system size beyond 256 has diminishing returns, as the change from 256 to 512 is less than the change from 128 to 256.

When the mutation rate is decreased to 0.001, the saturation at low system sizes is unchanged, but at high system sizes the saturation point increases. These results

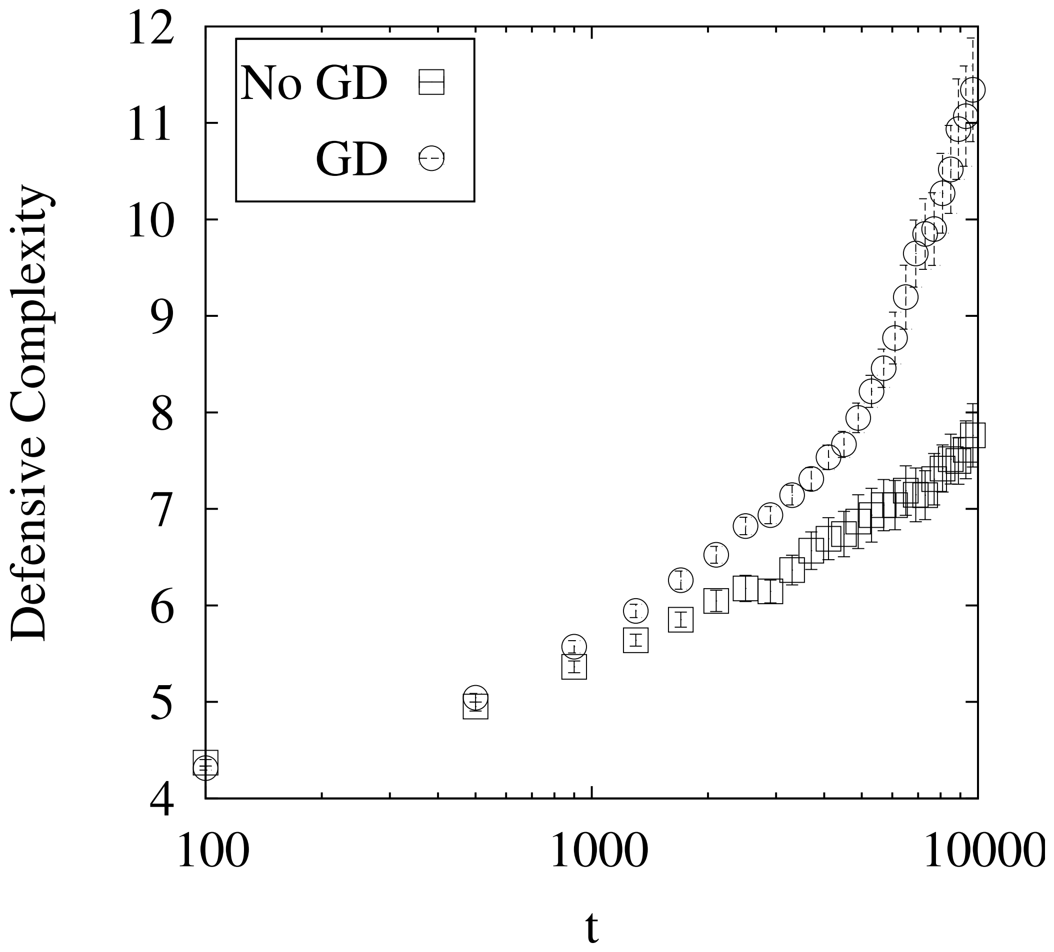

Figure 8.3: This figure shows the growth of the average maximum defense complexity as a function of time with and without the gene duplication mechanism, averaged over many runs. In the case without gene duplication, the complexity grows logarithmically with time. With gene duplication (at a rate of 0.1 per generation), the complexity grows super-logarithmically.

are shown in Fig. 8.2. This suggests that a large mutation rate creates a specific maximum complexity value due to entropy pressure, and that a small system size creates a different specific maximum complexity value. Thus the system will increase in complexity until it reaches the first of those maxima. When the data are plotted in terms of variables which reflect the asymptotic complexity scaling, they collapse onto a single curve. This is analogous to finite size scaling around a critical point in which the system size creates a departure from criticality and causes the scaling to saturate.

The data collapse takes the form of $r^a(C - C_0) = f(r^a S^b)$ where $f(x)$ scales as

x when $x \to 0$ and $f(x)$ approaches a constant when $x \to \infty$. The data are found to collapse for $a = 0.6 \pm 0.2$, $b = 2 \pm 0.1$, and $C_0 = 6.65 \pm 0.1$. The error in these quantities was determined by varying them around the point of best collapse and monitoring the quality of the collapse. The S^2 dependence is indicative that the total population is the relevant quantity when determining finite-size effects. The value of C_0 is consistent with the complexity one would generate by randomly generating strings of length 2048 with a proportion of defense characters to alphabet size equal to that observed in the smallest systems. That is to say, at the asymptote corresponding to high mutation rate and low system size, the complexity of strings is due entirely to evolutionary pressures on the relative proportions of the different characters, rather than spatial organization within the genome.

The saturation due to large point mutation rate can be understood as being due to its complexity dependence as discussed earlier and in terms of the Eigen error threshold [123; 124], but the observed scaling exponent is not at this time understood. The system size scaling is surprising as it is not obvious a priori that the complexity of an organism's genome should be related to the size of the space the organism lives in (in contrast with turbulence, in which the complexity of the flow is expressed in the distribution of velocity throughout the system).

It is possible that the connection between system size and complexity in 'Foodchain' is a result of the fixation of complexity-decreasing mutations. For a finite population of organisms with a set of traits that may be present or absent in each organism, the fluctuations in the population and the dynamics of reproduction will eventually cause the trait to be either present or absent in every member of the population. The probability of a particular mutation going to fixation is $P(s) = (1 - \exp(-2s))/(1 - \exp(-4Ns))$, where s is the selective advantage and N is the population size[125–127].

In the context of the Foodchain model, each organism may have many strings of varying complexities only a few of which are responsible for the organism's reproductive success. The pivotal strings are not necessarily those of the highest complexity (short defense strings can still be important in defending against short attack strings held by other organisms, for instance). However, a mutation to the most complex string may turn it into a pivotal string even if it is not currently experiencing selective pressure. In the low mutation rate limit fixation of complexity-decreasing mutations imposes a limit on the maximum sustainable length L of a particular string. We balance the rate of fixation of complexity-increasing mutations (which occur at a constant rate) with the rate of fixation of complexity-decreasing mutations (which occur at a rate proportional to L): $P(s) - LP(0) = 0$, where $P(0) \propto 1/N$[126]. This results in the scaling $L \propto N$, consistent with the system size scaling exponent observed in the data collapse.

Another point of interest in the dynamics of the Foodchain model is the presence of large fluctuations in the time series of the complexity. Compared to the case of neutral mutation, the probability distribution of complexity fluctuations has long tails that follow a decaying exponential for small fluctuations, but a power-law for larger fluctuations. The defensive complexity fluctuation distribution function for several runs at different system sizes and mutation rates is plotted in Fig. 8.4. We observe that for large fluctuations, the probability of a given fluctuation x scales as $x^{-3.5\pm0.7}$.

This can be understood as the consequence of the blocking dynamics of the interactions. An organism survives so long as it has a defense string for all of its neighbors. As a consequence, the large-scale dynamics of the system can become locked up, modulo small fluctuations due to mutations during replication. However, once something evolves a new attack string that does not have a corresponding defense string in its neighborhood, it will expand until either a mutation provides a defense string, or everything around it is extinct. This causes large changes in the composition of the system. Furthermore, if something with a very complex defense string is blocked, but

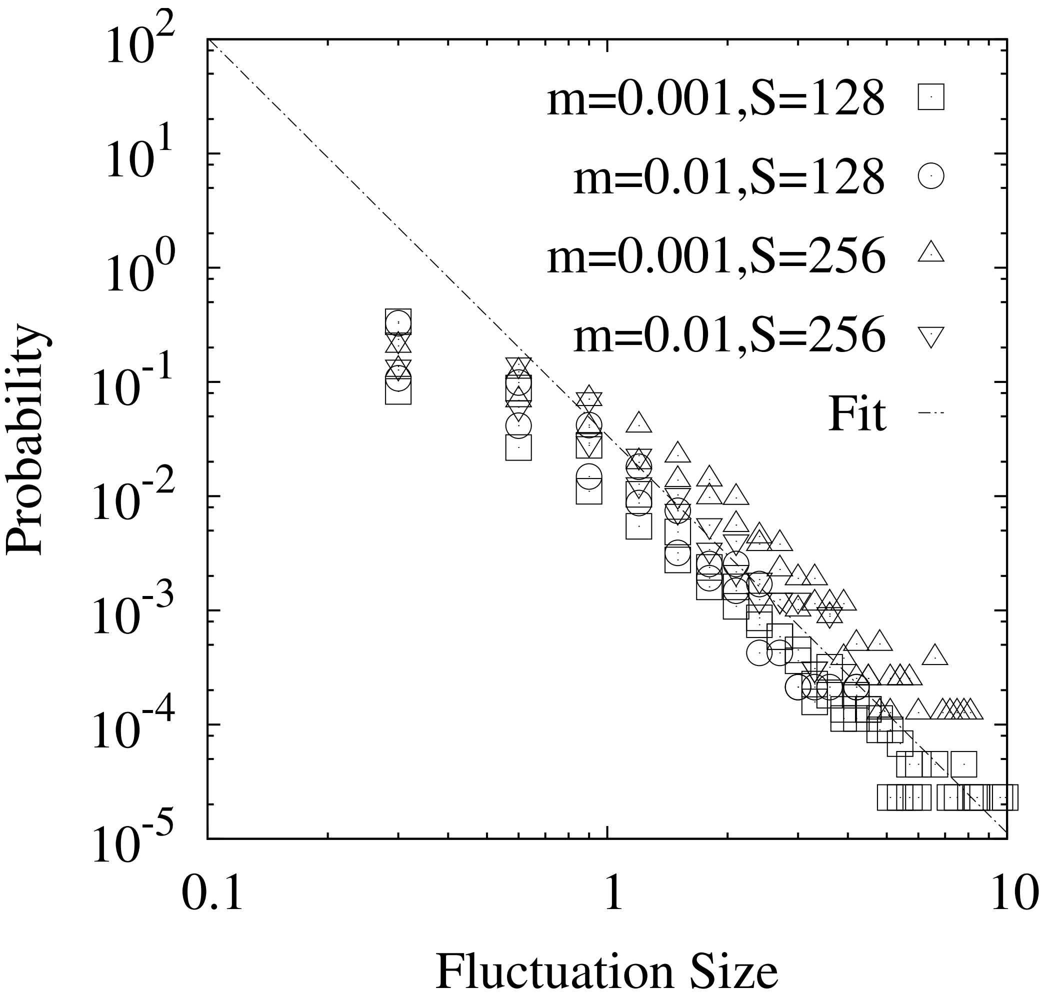

Figure 8.4: Here we plot the distribution of average maximum defense complexity changes from one generation to the next for different system sizes and mutation rates. The fluctuations are power-law distributed and do not strongly depend on either system size or mutation rate.

then mutates so that it can eat its neighbors, its representation in the system will grow quadratically (due to the dynamics of an expanding front in two dimensions). Consequently, the average complexity of the system will change dramatically in a short time.

If the large fluctuations arise from changing representation in the system, one would expect that the defense complexity fluctuations would be symmetric about the mean (as it is the attack complexity that determines whether something will grow, and so one would not expect the defense complexity to preferentially increase or decrease). On the other hand, if the large fluctuations arise due to the dynamics of gene duplication and mutation, the fluctuations should be skewed as the two mechanisms act with different strengths to increase and decrease complexity. We observe that the fluctuations are symmetric about the mean, suggesting that they are a result of the dynamics of changing representation of organisms in the system.

These long-tailed fluctuations in system representation may cause the diversity of the system to obey a different scaling law than in the case of either neutral evolution or evolution towards a fixed point. The representation of the system can change rapidly. This is similar to results seen in bioreactors[128; 129], in which the dominant OTU changes on a very fast timescale even if the throughput of the system remains constant.

8.4 Conclusions

We have shown that in the absence of a complexity scale, an evolving system can cascade towards ever-increasing complexity. A finite rate of point mutation, or finite system size can both interrupt this cascade, creating a point of saturation in analogy to the saturation of complexity in previous evolutionary simulations.

Unlike the case of turbulence, however, we do not see a scale-invariant distribution of organisms in complexity space. In the simple 'Foodchain' model presented here, there is no separation between primitive organisms that compete with each other using structures of low complexity and organisms with very complex offensive and defensive strings. In order to generate a rich hierarchy of structures, some form of trophic structure would need to be represented in the system[130].

The mechanism for the cascade presented here in competition. We can also see a complexity cascade in symbiotic interactions, by making a small change to the 'Foodchain' model. If we imagine that there is some chemical energy input to the system that is broken down into waste products, which can then be processed as a

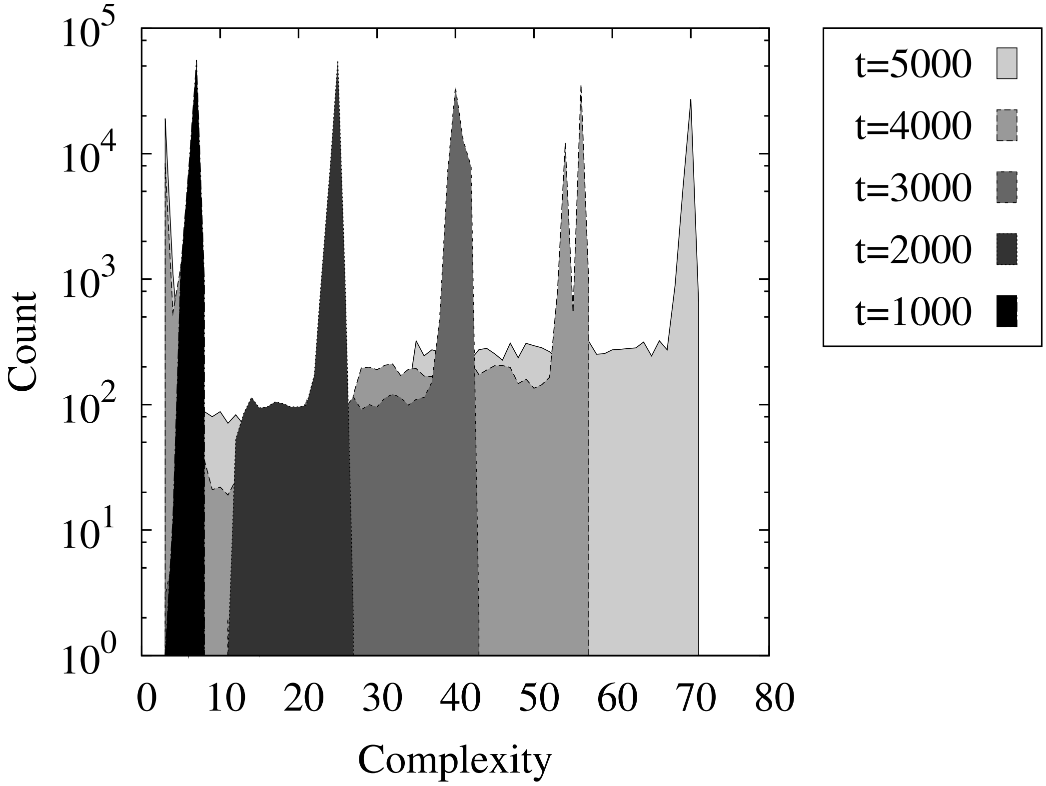

Figure 8.5: Population histogram with respect to string complexity in a symbiotic complexity cascade. The system develops a travelling peak at the maximal complexity. Organisms of lower complexity can still exist stably in the system. At later times, secondary peaks at lower complexity can develop.

secondary energy source that produces its own waste products, and so on. The model equivalent for this picture is that organisms with a given active string can metabolize the waste products of any matching shorter string (in a non-destructive way). The strings will lengthen by the exact same mechanism as in competitive 'Foodchain', but in a non-competitive fashion. This gives rise to a diverse population of organisms with many different active string lengths (Fig. 8.5).

We therefore see that the complexity cascade can be a mechanism for niche formation, and give rise to ecological structure in an otherwise homogeneous system (that is, one with no inherent niche structure).

Despite these results, we have not seen the emergence of qualitatively different modes of interaction in this model, as the organisms do not have any way to develop a new mode of interaction other than comparing more complicated strings.

Chapter 9

PlantNet

In this chapter I will discuss a second model system that exhibits a complexity cascade. In the previous chapter we saw the role of mutation rate and finite system size in preventing an infinite cascade from occuring. We have postulated that a fixed fitness landscape will also inhibit the cascade, as the complexity scales inherent in the lengthscale may cause the system to become stuck. If this is a strong constraint, though, then no natural systems can be reasonably thought of as having a complexity cascade. There are minor differences in efficiency associated with almost any detail of an organism's structure, so it is inconceivable that certain organisms truly have a complexity-invariant fitness landscape.

To resolve this dilemma, we will consider the effects of perturbations to the fitness landscape of different size and complexity scale. It is not necessarily the case that an infinitessimally small term added to the fitness landscape with a fixed scale will completely disrupt the cascade — one expects that small terms should have smaller influence in determining where the cascade saturates, just as one obtained part of a complexity cascade even at finite mutation rate. The prediction is then that organisms whose dynamics are predominantly evolution towards a fixed point will behave differently than organisms whose dynamics are predominantly interactions with the population, but that there is the possibility to go from one regime to the other as evolution continues.

The model we will use to test this idea is one of competition between organisms for a shared resource - food or light that is being transported down from the top of the system. The organisms in the system will evolve to become larger in order to not be blocked by their neighbors. However, the more mass an organism has, the less efficient they are — the consequence is that there is an absolute fitness landscape, where the shortest organism (that manages to get access to the food) is the optimal one. We will see that even when this fitness landscape exists, the slope of the landscape (how much penalty associated with a certain amount of added mass) determines the saturation point of the complexity cascade. Here, unlike in Foodchain, the organismal complexity reduces simply to the organism's height relative to its neighbors (though there may be many different ways of attaining that height). Because of this, there is an analogy here between the complexity cascade picture and Red Queen evolutionary dynamics[131], of which this is a classic case.

We can thus understand the complexity cascade as a broader version of Red Queen dynamics. It is not simply that two particular channels of evolution are co-evolving in an arms race, but that as these avenues saturate out (it is impossible to become, e.g., a faster runner) then the organisms must explore new possibilities and modes of competition.

We consider a model in which the three-dimensional structure of organisms is uniquely derived from its genome. Our model is a variation of a model used to construct realistic virtual environments[132; 133], in which coevolutionary pressure on the height of the virtual trees was observed. The primary difference is that, unlike in the case of the Lindenmeyer systems used in their model, the size of an organism in our model is correlated with the length of the ordered regions in the organism's genome. The demands we impose on our organisms in order for them to change shape are much more stringent than those of real multicellular organisms. Rather than determining the size of an organism by associating the various dimensions and lengths with scalars that are controlled by gene expression or regulation (which would imply a fixed type of geometry to some degree) or by connecting it to the degree of recursion, we allow the structure to be completely general. In this way it is hoped that we can see more than the simple scalar Red Queen situation, because it is necessary to add new information to the genome in order for the organism to continue progressing in a given direction (for it to become taller, have more branches, etc.)

9.1 Theory

We can understand the effect of a fixed fitness landscape by asking a similar question as we did in Foodchain to understand mutation and system size effects: what do the evolutionary dynamics do to the average string length in the genome? The complexity cascade then saturates at the fixed point of the evolutionary dynamics.

Let us consider a tree of height h with a total number of cells C, a subset Φ of which receive light. As a first approximation, only the canopy of the tree can receive light, and so we have the bound $C \geq h + \Phi_{\max}$. The fitness function is:

$$F = \Phi/C - \alpha \tag{9.1}$$

The first term is the amount of the radiant energy absorbed from the environment that each cell in the organism receives. The second term is the amount of energy that each cell needs to survive. Consequently, the optimal state at nonzero α in the absence of other organisms is a mat with minimal height, such that $\Phi = C$ and the fitness is $1 - \alpha$. This implies that for any organism in the system to survive, $\alpha < 1$. If $\alpha = 0$, then any tree that receives at least one cell worth of light will survive, and as such the optimal state is to be infinitely tall. The maximum survivable height of an organism with canopy area A is strictly limited by:

$$h_{\max}/A = \frac{1}{\alpha} - 1 \tag{9.2}$$

Due to the presence of other trees, Φ may be reduced. If there is no neighboring tree which is taller, then the tree will be completely illuminated. Otherwise, some portion of its light will be blocked. If its height is less than the maximum survivable, then so long as it is taller than every other nearby tree, it is guaranteed to survive. Only a few trees will have this distinction, however. On average, a given tree will experience some amount of shadow based on its height with respect to the distribution of tree heights in the system.

If we assume a Gaussian distribution of heights in the system, characterized by an average height $\langle h \rangle$ and a standard deviation $\langle h^2 \rangle - \langle h \rangle^2$, then the amount of shadow a given tree receives is distributed according to the probability that some number of its neighbors are taller than it. If the tree has N neighbors, and each neighbor can shadow $1/N$ of the tree's canopy, then we will generically have a Poisson distribution of the amount of shadow S:

$$P(S) = \lambda \exp(-\lambda S) \tag{9.3}$$

The sharpness of the distribution depends on where the tree's height is with respect to the average. If there is a probability P_h that a given neighbor is taller, then $\lambda = N \ln(P_h/(1-P_h)) \approx N \ln(P_h)$. The probability of survival is then the probability that the tree receives at least αC light, or that the amount of shadow is less than $A - \alpha C$

$$P_s = 1 - \exp(-\lambda(A(1-\alpha) - \alpha h)) \tag{9.4}$$

The optimum height is therefore:

$$h = A(1/\alpha - 1) + \frac{\sqrt{\pi}\sigma_h N \ln(2)}{2N} \tag{9.5}$$

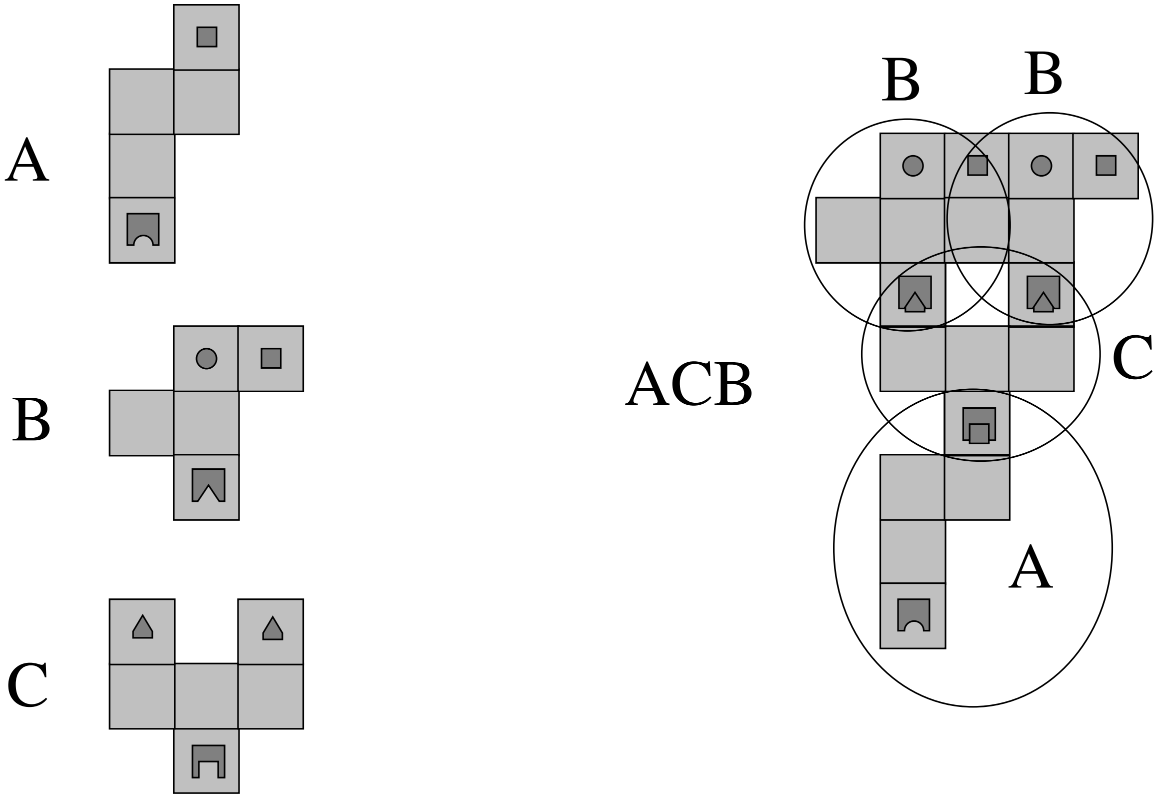

Figure 9.1: This is an example system in which there are three modules: A,B, and C, shown on the left. These modules have different attachment sites (△,∘,□) which are specific to each other. On the right, the resultant tree from expressing the genome ACB is shown. Cells with matching attachment sites are placed so that they overlap.

So we again see a scaling of the equilibrium height with $1/\alpha$.

9.2 The Model

Organisms in the PlantNet model have three-dimensional structures that are uniquely determined by their genome. The structures are composed of a series of modules with attachment points. These modules are small, connected sets of cells that are generated at the start of the simulation by randomly filling in a fixed-size region of space and ensuring that the resultant structure is fully connected - that is, there are no floating nodes. Each module corresponds to a particular base in the genome. The genome is an ordered variable-length string of these bases, which represents the sequence of expression as the structure grows. Cells in a given module may be attachment sites or otherwise inert. All cells receive light and cast shadows.

The structure begins by placing whatever module corresponds to the first base in the sequence. Then, the genome is expressed base by base in order. When a given module is expressed, it attaches at all open sites. If there are multiple possible attachment pairs on the two modules, the pair that produces a minimum overlap with the existing structure is chosen. Modules that cannot be placed without generating more than a threshold level of overlap are not placed. In cases where two open sites would create an overlap with each other, there is an fixed, arbitrarily chosen ordering to which has precedence. This process of construction continues until the genome has been fully expressed. An example set of modules and resultant organism is shown in Fig. 9.1. One consequence of this structuring is that a subset of the genome moved elsewhere will produce the same sub-structure, but will produce it at a different stage of the organism's growth. As such, this representational scheme is innately modular. Gene duplication is a very effective mechanism in this system because of that modularity.

Each organism has a root cell where the first module is placed. When an organism replicates, it picks a random new grid site. This is then used as the root for a new organism with a slightly mutated genome. Organisms are allowed to overlap, but overlapping cells do not contribute energy to the organism. Another version of this model involves special replicative cells such that the offspring is generated directly beneath a random replicative cell. This introduces the effect of spatial clustering of the organisms, but also introduces artefacts as organisms attempt to structure themselves so as to not be overshadowed by their own offspring. The central results are similar in both cases. In the interest of keeping the number of mechanisms at work small, we will only discuss the simpler version with random offspring positioning.

The genome of the offspring is changed via three mechanisms: mutation, deletion, and gene duplication. Mutation has a chance of operating on every base in the genome. When it operates, it randomizes the base that it is operating on. Deletion has a chance of operating once every replication. If it does so, it deletes a randomly chosen subset of the genome. Gene duplication copies a random subset of the genome to a new position, re-inserting the string at that location. It has a fixed probability of operating once during per replication. Deletion and gene duplication are made to have the same distribution of substring lengths in order to ensure that there is not a preferred genome length.

Every iteration, each organism has a chance to expire based on its fitness as well as a random culling process. If the fitness of any organism is less than or equal to zero, that organism will die. Additionally, there is a fixed chance per organism that it dies regardless of its fitness. This prevents the system from freezing if it happens to reach a state in which no organism shadows any other organism. For each organism that dies, an offspring is generated from a random parent to replace it and keep the population fixed.

For the fitness function, we use Eq. 9.1.

$$F = \Phi/C - \alpha \tag{9.6}$$

9.3 Results

The system we considered had 80 building blocks each existing within a cube of side 5 grid cells. The templates each contained from 4 to 8 attachment sites and any number of inert cells, such that all cells were placed adjacent to each other. When constructing the trees from the genomes, the overlap threshold beyond which we did not allow a module to be expressed was 5. Organisms initially had random genomes of lengths from 16 to 32 bases long. Each organism was bounded to a column with base 20 grid cells and height 120 grid cells. The system was a space of size $30 \times 30 \times 120$

grid cells.

We simulated this system at mutation rates of 0.001 and 0.01 per base, with α varying from 10^{-4} to 10^{-1}. The deletion rate and gene duplication rates per replication were held constant at 0.01. For these parameter ranges, there was no noticeable effect of the mutation rate on the limiting heights of the organisms. For the smallest values of α, organisms in some runs reached the system size limit, and so we expect a deviation from scaling in that range. The heights of the organisms appeared to converge after a few hundred generations (Fig. 9.2), with some sudden changes due to gene duplication events. In determining the limiting heights of organisms and their error bars, we used height values from multiple initial conditions averaged from $t = 1000$ to $t = 1500$.

We observe that the limiting heights increase with decreasing α as expected (Fig. 9.3. For $\alpha \sim 10^{-1}$, the scaling asymptotes to our prediction of $1/\alpha$ (Eq. 9.2). However, at smaller values of α it quickly deviates, growing only logarithmically with decreasing α. This could be due to a combination of factors. The limiting height value is strongly initial-condition dependent, however, suggesting that the system has many local minima. As such, it may be that much longer simulation times would give rise to increased limiting height values over the ones we report due to rare events knocking the system out of its local minimum. Furthermore, the values of h at which the scaling deviates from the prediction are comparable with the system size, and so system size effects are likely interrupting the cascade.

Also of note is the connection between genome length and height. Because of the way that the structure of the organism is tied to its genome, it is necessary for the genome to elongate in an ordered manner (such that each module has a connection site that matches to one on the previous module) for the tree to become taller. The correlation between genome length and organism height is plotted in Fig. 9.4. For short times, the curves of height versus genome length for different parameter values

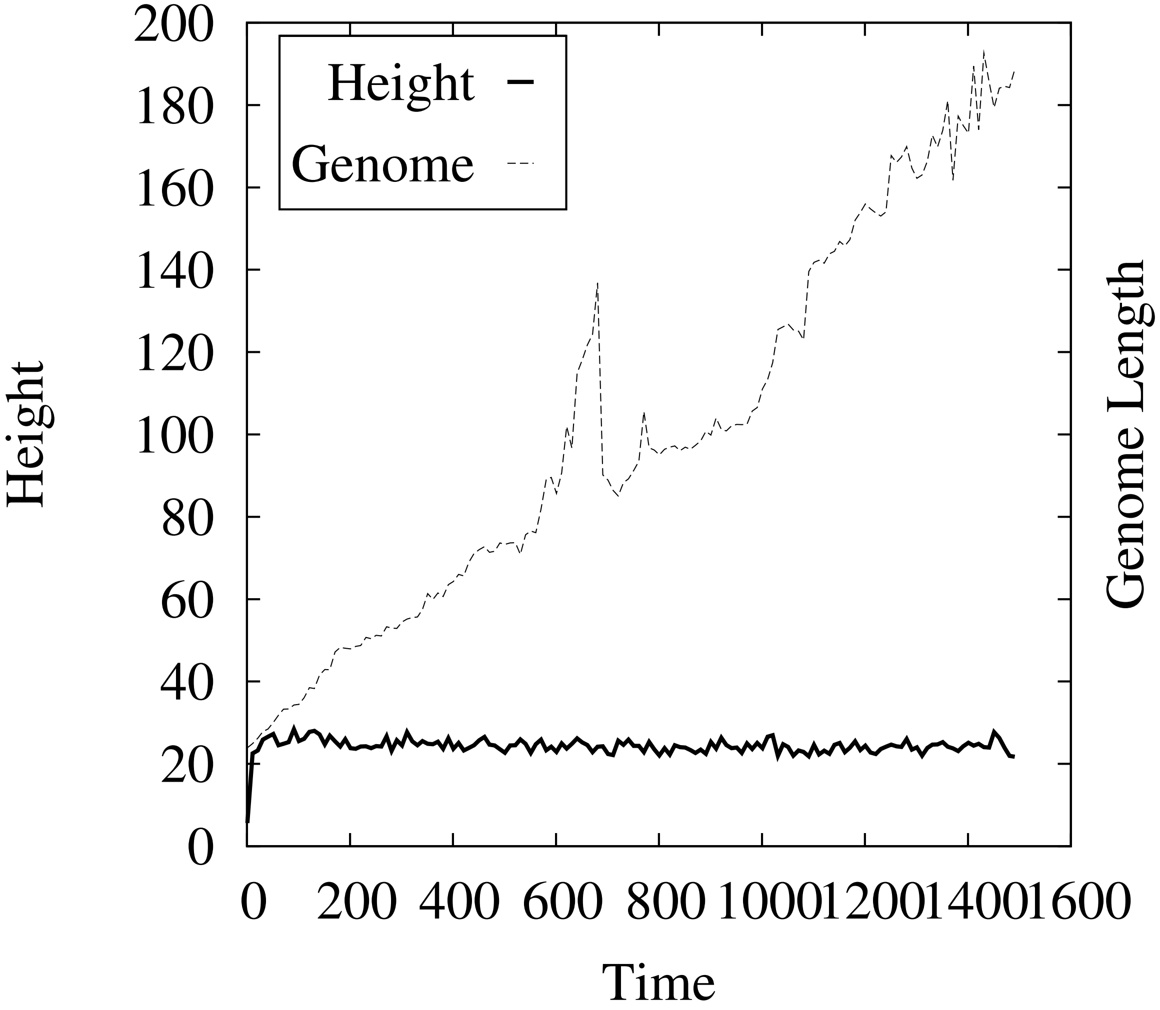

Figure 9.2: Time dependence of the average organism height and average genome length for a system with $m = 10^{-3}$ and $\alpha = 6 \times 10^{-4}$. The average height saturates to its limiting value quickly, though the genome length continues to fluctuate.

collapse onto a single curve - here, growth of the genome and growth of the organism's structure are strongly connected. At longer times, the genome continues to grow through the evolutionary operators, but the height becomes fixed by balance from the penalty term.

9.4 Conclusions

The way in which the organism's genome connects to its structure in this model is more similar to the construction of large proteins than the actual phenotypic evo-

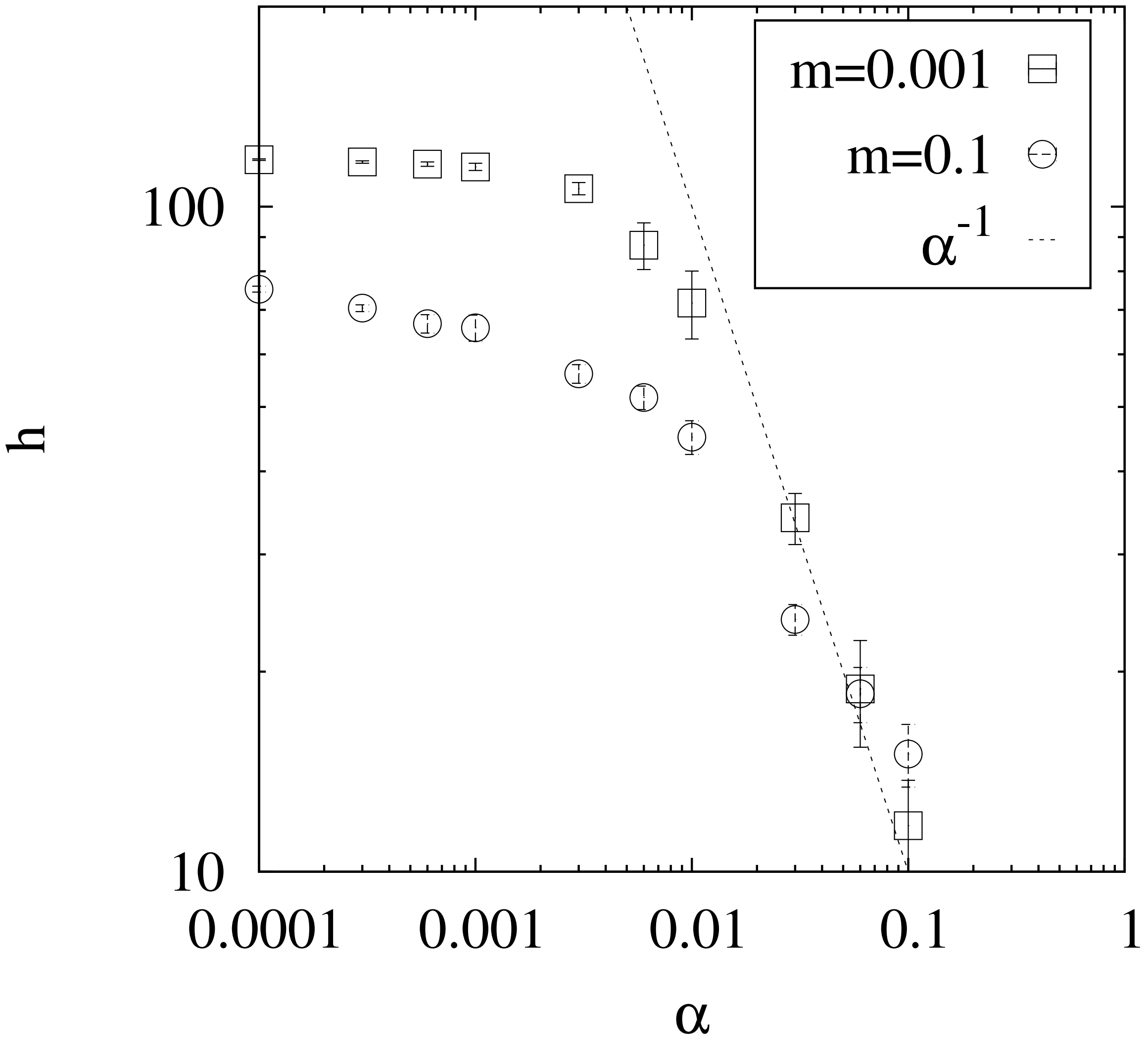

Figure 9.3: Scaling of the limiting height of organisms with respect to mutation rate and α. At large α, the predicted $1/\alpha$ scaling is observed.

lution of trees, algaes, or other macroscopic organisms that this system appears to model. This would be a deficit were one to try to use this model to study, for example, the height distribution of trees in a jungle. On the other hand, this particular organization allows this system to be used to study the complexity cascade described in the previous two chapters. Here, changing the height of the organism is not just a matter of changing expression levels of a particular fixed set of genes, but rather requires an increasingly large section of the organism's genome to be ordered in a specific fashion. A mutation somewhere in the middle of that section could remove a module whose connection point was needed for the rest of the genome to make sense.

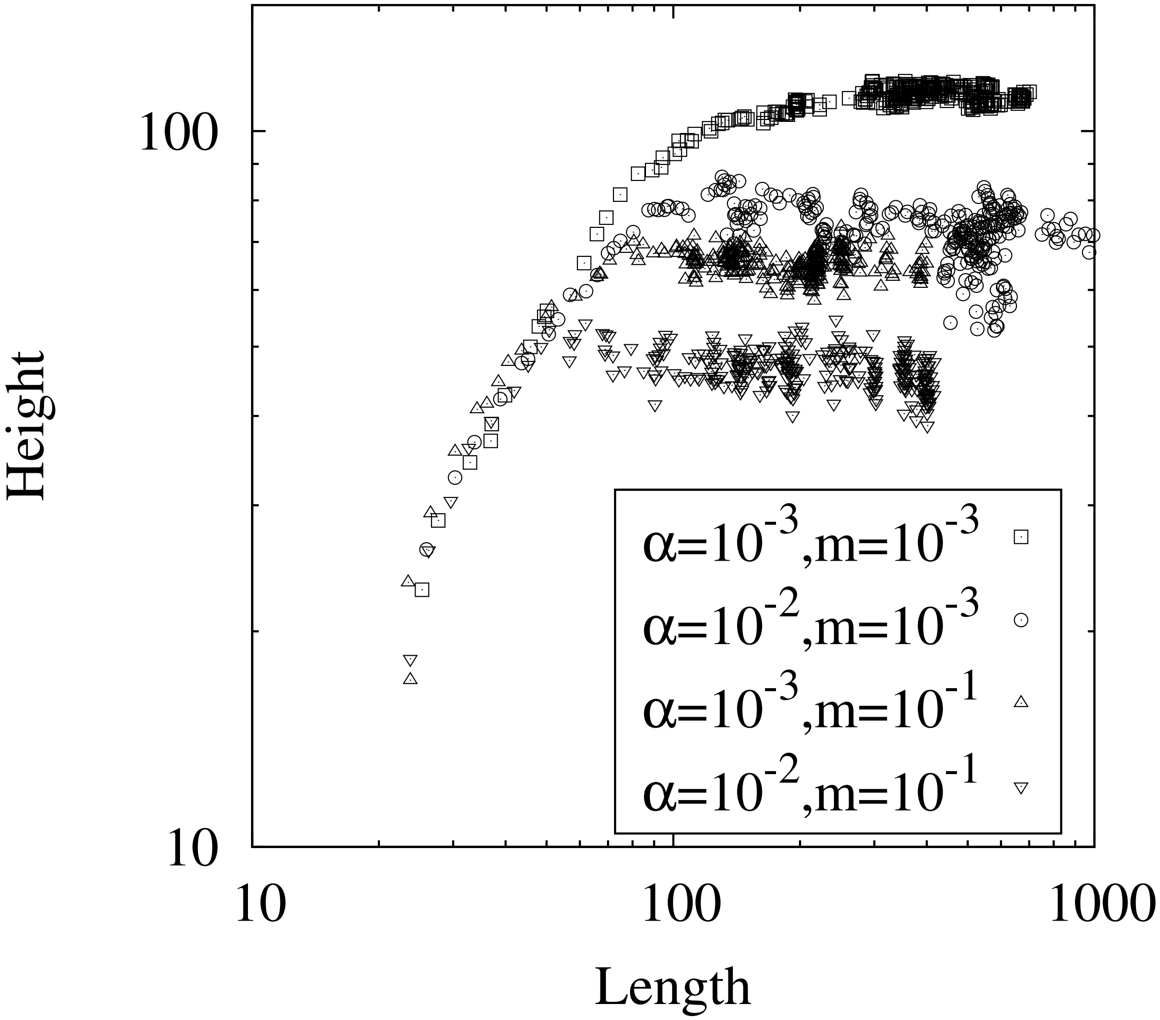

Figure 9.4: Average organism height versus average genome length for four systems. Though there is a wide variance, there is a clear correlation between genome length and organism height that shows the connection between internal organization and external structure in this model.

As such, the increasing height of trees in the Plantnet model is tied to an increase in genome complexity, and is perhaps better thought of as similar to the introduction of increasingly developed structures in the evolution of embryogenesis.

Chapter 10

Meta-evolution

We turn now to a question of evolvability. It is well known that the evolutionary dynamics of the parameters of evolution itself is markedly different from the evolution of phenotype[134; 135]. Selection tends to work immediately upon acquisition of a harmful phenotype, and somewhat immediately upon acquisition of a beneficial phenotype. On the other hand, changes to the dynamics of evolution do not manifest selective advantage until many generations later. How then did the variety of evolutionary mechanisms beyond simple point mutation manage to evolve?

From another point of view — that of designing genetic algorithms — it is interesting to consider the evolution of evolutionary mechanisms because one would like to find the optimal set of parameters for a given problem. If the problem is not well-known before-hand, it is necessary to find that optimum via some method. If the optimum could be found as part of the iteration of the genetic algorithm itself, that would be advantageous. On a further level, if new evolutionary operators specifically tuned to the relevant problem space could be developed by the dynamics of evolution, then the solution times for those problems would be much improved. Much of the convergence time of a genetic algorithm depends on setting up the problem in the right representation[136], so a way to automatically find such a representation would be very useful.

We would like to concern ourselves with the conditions in which evolutionary pressures on the mechanisms of evolution is comparable to the pressures on phenotypes. It is in those conditions that the evolutionary operators themselves have the greatest chance of changing significantly. These conditions are also a necessary part of the abstraction transition picture, as new levels of evolution demand new evolutionary mechanisms as well.

There are two candidate conditions that are likely to produce selection pressures on meta-evolutionary parameters. Firstly, a system which has some sort of long-time periodic selection might be able to emphasize long-time evolutionary benefits over short-time evolutionary benefits. Long-time periodic selection could arise from the day-night cycle, tides, seasonal variations, or rare disasters.

Secondly, a system in which multiple organisms are supporting one another may be able to emphasize long-term benefits, as some subset of the organisms in the system are free to evolve without having to support as much evolutionary load (that is, only some organisms need to focus on phenotype, and they stabilize the environment and decrease the effective selection pressure on phenotype for the others). In essence, the dynamic is that a subset of the organisms in the system live or die as a whole — the selection of that subset is synchronized. As such, different subsets are effectively competing with each other. Synchronous selection is most likely to occur in biofilms, but may also result from spatial isolation into subsystems (pools, for instance).

In order to test these conditions, we will perform simple simulations in which the conditions are enforced by hand, and both the mutation rate and a scalar phenotypic parameter are allowed the vary. We will then measure the selection pressure using a technique detailed in the Simulations section. A consequence of the nature of meta-evolutionary parameters is that even though they do not appear in the fitness function, they experience selection pressure. While the change in these parameters may be computed by careful application of the statistics of the alterations they introduce, the ability to include them as an effective term in the fitness function is useful in order to connect them to traditional models from population genetics.

The simulations suggest that both time delays and synchronous selection can give rise to pressures on evolutionary parameters, and that there are optimum time delays and optimum group sizes in order to maximize such pressures. Furthermore, we see that it is necessary for the system to be in a constant state of evolution towards an unattained goal in order for meta-evolutionary aspects to develop. As such, the open-ended evolution of the complexity cascade may be an ideal system to drive the development of new evolutionary dynamics.

10.1 Theory

10.1.1 Price's Theorem

To predict the scaling of selection pressure on mutation rate and other meta—evolutionary quantities, we will make reference to results from the theory of population genetics. For sake of completeness, and because the derivations are simple and elegant, I will rederive here the central result that we will make use of: Price's equation[137]. I will reproduce here a version of the derivation from [138], but using language familiar to statistical physicists.

We begin with an infinite population of organisms that each have some set of traits, encoded by the vector $\vec{\psi}$. We furthermore propose that we can determine the probability that a given organism survives a selective event, proportional to its fitness $P \propto F$. We then wish to determine the new average value of some combination of the traits $\langle \phi \rangle$. Between selective events, there may be drift or other mechanisms operating on the trait. These effects are summed up in a term δ_ϕ, which corresponds to the change in ϕ between the previous selection event and the present one. We must be careful to include these before the effects of selection.

In order to properly normalize the new averages, we must consider the change in the total number of organisms (even if this number is infinite). For a population of size N, the average number of organisms after a generation has passed is $\sum_{i=1}^{N} F_i = N\langle F\rangle$. We can then use this to determine the new value of any system average:

$$\langle\phi\rangle' = \frac{\sum_{i=1}^{N}(\phi_i + \delta_i)F_i}{\sum_{j=1}^{N} F_j} = \frac{\langle\phi F\rangle}{\langle F\rangle} \tag{10.1}$$

So the change in the average value of ϕ is give by Price's Theorem:

$$\Delta\langle\phi\rangle = \frac{\langle\phi F\rangle - \langle\phi\rangle\langle F\rangle + \langle\delta_\phi F\rangle}{\langle F\rangle} \tag{10.2}$$

It is important to note that this must be the fitness evaluated in terms of the variables at the moment of selection. As such, δ_ψ will in practice show up in the correlations as well.

10.1.2 Time delays

Let us consider a system with a single scalar phenotype s and a scalar mutation rate r. The system consists of N organisms which each undergo a series of T neutral replications before selection occurs. Each replication, the scalar phenotype is offset by a gaussian-distributed random amount with standard deviation r, and the mutation rate is offset by a gaussian-distributed random amount with fixed standard deviation Δ.

After T replications, the new value of s and r for a particular organism will be:

$$s' = s + \sum_{i=1}^{T}(r + \Delta\sum_{j=1}^{i-1}\eta_{rj})\eta_{si} \tag{10.3}$$

$$r' = r + \Delta\sum_{i=1}^{T}\eta_{ri} \tag{10.4}$$

If the fitness function is $F_i = f_0 + f_1 s_i + f_2 s_i^2 + ...$, then we can use Price's theorem (Eq. 10.2) to find the average value of the mutation rate after a selection event $\langle r' \rangle_S$:

$$\Delta\langle r\rangle = \frac{\langle rF\rangle - \langle r\rangle\langle F\rangle + \langle \delta_r F\rangle}{\langle F\rangle} \tag{10.5}$$

We will expand the fitness in terms of its value at the time of the previous selection event, compared to its value immediately before selection occurs:

$$\Delta\langle r\rangle = \frac{\langle r(F+\delta_F)\rangle - \langle r\rangle\langle (F+\delta_F)\rangle + \langle \delta_r (F+\delta_F)\rangle}{\langle F+\delta_F\rangle} \tag{10.6}$$

We assume no a-priori correlation between mutation rate and fitness to start with, and so the covariance term is zero. Furthermore, by the construction of our evolutionary rules, there is no correlation between the change in mutation rate and the fitness outside of selective events. This leaves us with:

$$\Delta\langle r\rangle = \frac{\langle r\delta_F\rangle - \langle r\rangle\langle \delta_F\rangle + \langle \delta_r \delta_F\rangle}{\langle F+\delta_F\rangle} \tag{10.7}$$

Where δ_r and δ_F are:

$$\delta_r = \Delta \sum_{i=1}^{T} \eta_{ri} \tag{10.8}$$

$$\delta_F = f_1 \delta_s + f_2 \delta_{s^2} + ... \tag{10.9}$$

$$\delta_{s^n} = \sum_{k=0}^{n-1} \frac{n!}{k!(n-k)!} s^k \left(\sum_{i=1}^{T} (r + \Delta \sum_{j=1}^{i-1} \eta_{rj}) \eta_{si}\right)^{n-k} \tag{10.10}$$

To second order in F, we have:

$$\Delta\langle r\rangle = \frac{f_2(\langle r\delta_{s^2}\rangle - \langle r\rangle\langle \delta_{s^2}\rangle) + f_1\langle \delta_r \delta_s\rangle + f_2\langle \delta_r \delta_{s^2}\rangle}{f_0 + f_1\langle s\rangle + f_2\langle s^2\rangle + f_2\langle \delta_{s^2}\rangle} \tag{10.11}$$

We can determine these correlations:

$$\begin{aligned}\langle r\delta_{s^2}\rangle =&\langle r(\sum_{i=1}^{T}(r+\Delta\sum_{j=1}^{i-1}\eta_{rj})\eta_{si})^2\rangle + 2\langle rs(r+\Delta\sum_{j=1}^{i-1}\eta_{rj})\eta_{si})\rangle \\ \langle r\delta_{s^2}\rangle =&\langle r^3\rangle T + \langle r\rangle\Delta^2\frac{T(T-1)}{2}\end{aligned} \tag{10.12}$$

$$\begin{aligned}\langle \delta_{s^2}\rangle =&\langle (\sum_{i=1}^{T}(r+\Delta\sum_{j=1}^{i-1}\eta_{rj})\eta_{si})^2\rangle + 2\langle s(r+\Delta\sum_{j=1}^{i-1}\eta_{rj})\eta_{si})\rangle \\ \langle \delta_{s^2}\rangle =&\langle r^2\rangle T + \Delta^2\frac{T(T-1)}{2}\end{aligned} \tag{10.13}$$

$$\langle \delta_r\delta_s\rangle = 0 \tag{10.14}$$

$$\begin{aligned}\langle \delta_r\delta_{s^2}\rangle =&\Delta\langle (\sum_{k}^{T}\eta_{rk})(\sum_{i=1}^{T}(r+\Delta\sum_{j=1}^{i-1}\eta_{rj})\eta_{si})^2\rangle \\ \langle \delta_r\delta_{s^2}\rangle =&2\langle r\rangle\Delta^2\frac{T(T-1)}{2}\end{aligned} \tag{10.15}$$

Combining these results, we have:

$$\Delta\langle r\rangle = f_2T\frac{\langle r^3\rangle - \langle r\rangle\langle r^2\rangle + \langle r\rangle\Delta^2(T-1)}{f_0 + f_1\langle s\rangle + f_2(\langle s^2\rangle + \langle r^2\rangle T + \Delta^2\frac{T(T-1)}{2})} \tag{10.16}$$

We can compare this to the lowest-order change in the phenotypic parameter s:

$$\Delta\langle s\rangle = \frac{\langle s(F+\delta_F)\rangle - \langle s\rangle\langle F+\delta_F\rangle + \langle \delta_s(F+\delta_F)\rangle}{\langle F+\delta_F\rangle} \tag{10.17}$$

$$\Delta\langle s\rangle = f_1\frac{\langle s^2\rangle - \langle s\rangle^2 + \langle r^2\rangle T + \Delta^2\frac{T(T-1)}{2}}{f_0 + f_1\langle s\rangle} \tag{10.18}$$

Note that the lowest order term in the change in mutation rate is proportional to the time between selection events, whereas the corresponding term in the change in the phenotypic parameter is not proportional to T at all. This means that when T

is very small, there will be no selection pressure on the mutation rate. If the fitness function is concave down, the first order in T term is actually negative, selecting against mutation rate. However, if the fitness function monotically increases in terms of the phenotypic variable s, then there must be a higher order derivative of $F(s)$ which is positive, with the consequence that a sufficiently long time between selection events will eventually drive an increase of the mutation rate.

Let us consider the effect of balancing the mutation rate against a direct penalty. If there is a cost associated with implementing an evolutionary operator, we would like to know how high that cost can be before the evolutionary operator is selected against. This is a more direct measurement of the selection pressure, as opposed to the rate of evolution.

We use the fitness function: $F^* = F - \alpha r$ and see what additional terms are generated:

$$(\Delta\langle r\rangle)^* = \Delta\langle r\rangle + \alpha\frac{\langle r\delta_r\rangle - \langle r\rangle\langle\delta_r\rangle + \langle\delta_r^2\rangle}{\langle F + \delta_F\rangle} \tag{10.19}$$

Because r and δ_r are uncorrelated, the only surviving terms are:

$$(\Delta\langle r\rangle)^* = \Delta\langle r\rangle + \frac{\alpha\Delta^2 T}{\langle F + \delta_F\rangle} \tag{10.20}$$

We can then find the value of α such that the mutation rate does not change:

$$\alpha = \frac{f_2(\langle r^3\rangle - \langle r\rangle\langle r^2\rangle + \langle r\rangle\Delta^2(T-1))}{\Delta^2} \tag{10.21}$$

The selection pressure has a T dependent term and a T independent term. The T independent term depends entirely on the third-order variation in the distribution of mutation rates. These results suggest that for an arbitrary evolutionary operator, it might be characterized by the scaling of its selection pressure with time delay in the presence of a concave-up fitness function of phenotype in absence of intrinsic variation. Terms that have direct impact on the fitness will scale with an exponent of 0, whereas local operators such as point mutations will scale with exponent 1. Operators such as horizontal gene transfer that involve the structure of the entire population might scale with even larger exponents. We must be careful however — this simple picture involves continuous scalar quantities, whereas the dynamics of biological evolution operate on strings of symbols instead. Unlike in this case, there is likely to be a finite optimum mutation rate for a given evolutionary problem.

The dependence on concavity of the fitness function furthermore suggests that the evolution of meta-evolutionary operators will be strongest when new phenotypic degrees of freedom are being opened up. When a new degree of freedom is opened up, the gains from exploiting it are likely to increase the further one moves in that direction. On the other hand, once the system is near its optimum, the possibility of overshooting the optimum and the diminishing returns on optimization will cause the fitness function to become concave down.

10.2 Simulation

We will discuss two simulations of meta-evolution in order to test the predictions of the theory. One will vary the timescale between selection events, while the other will vary the group size and implement synchronous selection. In both simulations, the fitness is determined by the degree of matching to a target string, which is either held constant or allowed to change with time. In this fashion, we can see what influence the quantized nature of the genetic code has on the dynamics of meta-evolution, compared to the simple scalar model we discussed in the Theory section.

10.2.1 Measuring Selection Pressure

We would in general like to know the instantaneous strength of selection pressure on an evolutionary parameter ϕ, given the ability to evolve the system forward in time many times over from any given state. If we were to simply measure the rate of change of ϕ with time, we would not know whether what we observed was a result of selection pressure, or rather some combination of pressure with the intrinsic rate drift, a change in the rate of replication of organisms in the system, or any other such dynamics. Instead, we would like to measure the selection pressure in terms of the fitness function, using the known parts of the fitness function to set a scale for the relative strength of selection. For example, if we know the local slope of the fitness function with respect to a pair of explicit parameters ψ_1 and ψ_2, we can compare the selection strength on those parameters by comparing their slopes. It may be that due to a lower intrinsic mutation rate, ψ_1 changes more slowly, but if we can measure the fitness function we can actually determine their precise relative selection strengths.

When we add meta-evolutionary parameters, they do not show up explicitly in the fitness function. However, they do have an influence on the eventual changes in fitness that the system undergoes. If we explicitly link these parameters to the fitness function, we can find the point at which our explicit term balances the term that we cannot directly measure:

$$F^* = F - \alpha\phi \tag{10.22}$$

Since $d\phi/dt \propto \partial F^*_{ef}/\partial\phi$, where F^*_{ef} is the effective fitness function that we would have were we to know the exact influence of the meta-evolutionary effects, if we can change α such that $d\phi/dt = 0$, we have measured the strength of the effective selection on ϕ. That selection strength is given by the value of α such that $d\phi/dt = 0$.

10.2.2 Time Delay

In this simulation we have a population of N organisms, each of which has a genome L bases long. Each base is allowed to either be 0 or 1. Each organism also has a scalar parameter r that controls its mutation rate. The fitness of a given organism is a function of the number of 1's (n_1) in their genome and the mutation rate parameter:

$$F = g(n_1) - \alpha r \tag{10.23}$$

Every generation, the organism's mutation rate and genome evolve. Each base has a probability of changing given by:

$$P_{mut} = \frac{1}{1 + \exp(-r)} \tag{10.24}$$

The mutation rate r changes each generation via a random walk with coefficient M. The reason for constructing this map between the actual probability of mutation and the rate parameter is that otherwise, we would have to artificially bound r to lie between 0 and 1. If we do this, there is an entropy pressure from the edges of the interval. Furthermore, it becomes hard to achieve precise yet small rates in such a model — the rate of change of the mutation parameter would roughly set the smallest possible mutation rate one could achieve. We would like the system to as easily maintain a mutation rate of 10^{-12} as one of 10^{-2}, and so it makes sense to use a function such that a value of $-\inf$ is mapped to zero and a value of $+\inf$ is mapped to one.

Each organism replicates in place (replacing its parent) for T generations, after which the fraction f (50% in these simulations) of the population with the lowest fitness die, and are replaced randomly by replication of the remaining organisms. In terms of our fitness function from before, this selection mechanism produces a step-function around a particular fitness value. As such, it is hard to evaluate exactly what f_2 would be in the context of the above model; for the lower end of the system, it will be very positive, whereas for the upper end of the system it will be very negative.

This process is iterated for 10 cycles of selection around a given initial condition, and the change in the mutation rate is measured. This is repeated for many values of the counter-pressure α to generate the function $\Delta r(\alpha)$ (see Fig. (10.1). The zero of

$\Delta r(\alpha)$ is then found by a local linear fit, and the resultant value of α is the selection pressure for the corresponding value of T.

It is necessary to consider the initial conditions in examining the effects of meta-evolution. For any given set of initial conditions of interest, the simulation will immediately relax unless all of the moments of the distributions begin at their equilibrium values. This can dominate over any long-term pressures on the mutation rate by weakening the importance of mutations in changing the fitness of the system. Since we are interested in how the selection pressure scales with T, it makes sense to choose initial conditions that maximize the selection pressure so as to produce the best signal. In order to do this, each organism is initialized with an identical genome. This way the value of r determines the rate at which the standard deviation of the population grows, which in turn controls the speed of evolution. In terms of Eq. (10.16), we are extracting the term $2f_2T\langle r\rangle\Delta^2\frac{T-1}{2}$, so we expect a quadratic dependence of the selection pressure on T. We have determined the selection pressure curves for initial fitnesses of 0%, 25%, and 50%, for initial values of $r = -6$ (corresponding to a mutation rate of 0.0025) and $\Delta = 0.5$. These results are plotted in Fig. (10.2). Note that the mutation rate also does not have any initial variation. This case would correspond to a situation in which some new parameter of evolution had just become available, and so has not yet developed its equilibrium distribution.

Let us analyze the results of this simulation. There are marked differences from our predictions from the scalar phenotype model. First of all, the selection pressure is non-zero even at very small T. This initial selection pressure decreases as we approach an initial fitness of 50%, suggesting that it may be related to the fact that when the average fitness is less than 50%, mutations will on average increase fitness, even before selection operates. This means that there should be an additional term due to the dependence of the selection pressure on the correlation $\langle rF\rangle - \langle r\rangle\langle F\rangle$ from Price's Theorem.

When we look at how the selection pressure scales with time delay, it grows more slowly than the predicted linear dependence on T. Observationally, it seems to scale as $\sqrt{T}$. Furthermore, there is a particular time delay which maximizes the meta-evolutionary pressure which depends on the average fitness of the system.

If we now allow the mutation rate to have an initial variance, but prevent the mutation rate from itself mutating ($\Delta = 0$), we should see a change in the scaling of the selection pressure with T. We now initialize the system such that the rate parameter varies around its mean of -6 with a standard deviation of 1. This corresponds to the continued development of an evolutionary mechanism that is already well-integrated into the population. The selection pressure curves are plotted in Fig. 10.3.

Of note here is that for the case in which the system is initialized with an equal number of zeroes and ones, the selection pressure is initially independendant of T. This matches our prediction from Eq. 10.21, in which the only surviving term when $\Delta = 0$ is is T-independent.

10.2.3 Multi-level Selection

We now turn to the effects of clustering on meta-evolutionary emphases. It is already well-understood that the evolutionary dynamics of groups of organisms give rise to enhanced pressures on collective mechanisms such as altruism[139–142] due to the occurance of multi-level selection. However, something like the single-organism mutation rate is not a priori a collective mechanism. The proper distribution of mutation rates does however dominate the long-term fitness of the group, so in that sense it is a local property that over time gives rise to a collective effect.

There are many possible sources of multi-level selection of this sort. We will be examining the case in which the reproductive fates of a number of cells are all correlated — that is, their either live to reproduce or die as a group, using a collective fitness. This synchronous selection process means that the individual selective pressures on each organism are diluted. In effect, it may be a number of replications before the fitness of a given organism is responsible for whether or not the group survived. As such it may have the same effects on selection pressure as the time delay that we introduced in the previous section.

In this simulation, we divide the total population into a number of subgroups, each with a fixed size n. The total population is set such that there are always the same number of groups (G), so that the effective population upon which selection is operating is held constant. The fitness of a group is simply taken to be the average fitness of organisms in the group. Every iteration, the fraction of the set of groups f with lowest fitness die, and are randomly repopulated from the other groups. There are a few possible ways to do this. One way would be to pick random organisms from the rest of the system in order to assemble the new group. This would tend to destroy correlations within groups over time, and may not be the best test system. Furthermore, biofilms and colonies of cells may repopulate through sporulation[143], which suggests that new groups should be generated from a specific progenitor group. The system we use, therefore, is that the new group is be formed from a random choice of members of a randomly chosen progenitor group. A third possibility would be that each organism of the progenitor group replicates in-place to fill up the new group, exactly maintaining the internal population structure of the group. However, this is somewhat unrealistic for things that are not multicellular organisms and so we do not use it.

For this simulation, we use a genome of length 1024. The number of groups used is $G = 256$, and selection kills off the lowest 25% of groups. The initial mutation rate parameter is, as before, $r = -6$. We look at both the case in which the initial variation in the mutation rate is zero but $\Delta = 0.5$ (Fig. 10.5), and the case in which the initial distribution of the mutation rate is set, but $\Delta = 0$ (Fig. 10.6).

Here we see some of the same features as when we varied the selection timescale.

The selection pressure increases with group size as the group size increases. However there does not seem to be the same sharp dropoff beyond an optimum, but instead the selection pressure asymptotes to a constant value. Additionally, in the case without initial variation in the mutation rate, the selection pressure is zero up until a threshold group size, at which point it suddenly becomes positive. This can be understood by looking at the $\Delta r(\alpha)$ curves (Fig. 10.4). The selection for the mutation rate at these parameters is too weak to measure when the initial fitness is at 50% (that is, the point at which all mutations are neutral on average). This may be because of the details of the selection procedure (killing off the lowest fitness fraction, rather than having random comparisons or some other mechanism) or because of the low initial mutation rate (since as per Eq. 10.21, the pressure on mutation rate is proportional to the mutation rate).

10.3 Conclusions

We have seen that both long timescales and spatial correlations (in the form of grouping) in the process of selection can increase the selection pressure on meta-evolutionary parameters when compared to parameters with direct phenotypic consequence. For sufficiently large time delays, there is a sharp dropoff in the selection pressure on meta-evolutionary parameters which is not seen in the corresponding case of sufficiently large groups. When there is a population with initial variance in the value of the mutation rate, the delay and group size dependences of the selection pressure change.

The actual scalings observed for the case of an evolving string do not seem to correspond with the predictions for the scalar model. Whereas the scalar model will always predict scaling with an integer power of the time delay, we observe a time-delay scaling that appears to go as $\sqrt{T}$ when the initial variance of the mutation rate is zero. This observed power law changes exponent and regime of validity depending on the initial conditions of the system and when variance of the mutation rate is allowed to exist. This makes it unlikely that evolutionary operators can be cleanly classified by their time-dependence as proposed earlier, because the same operator will give many different scalings in different situations.

The increase of meta-evolutionary pressures in the presence of groupings and time-delays suggest places where the emergence of novel evolutionary mechanisms would be expected to have occured. Biofilms and places with low-frequency natural oscillations (tidal pools, for example) are prime candidates for such transitions, as they have the necessary elements to enhance meta-evolutionary selection pressures over immediate phenotype selection.

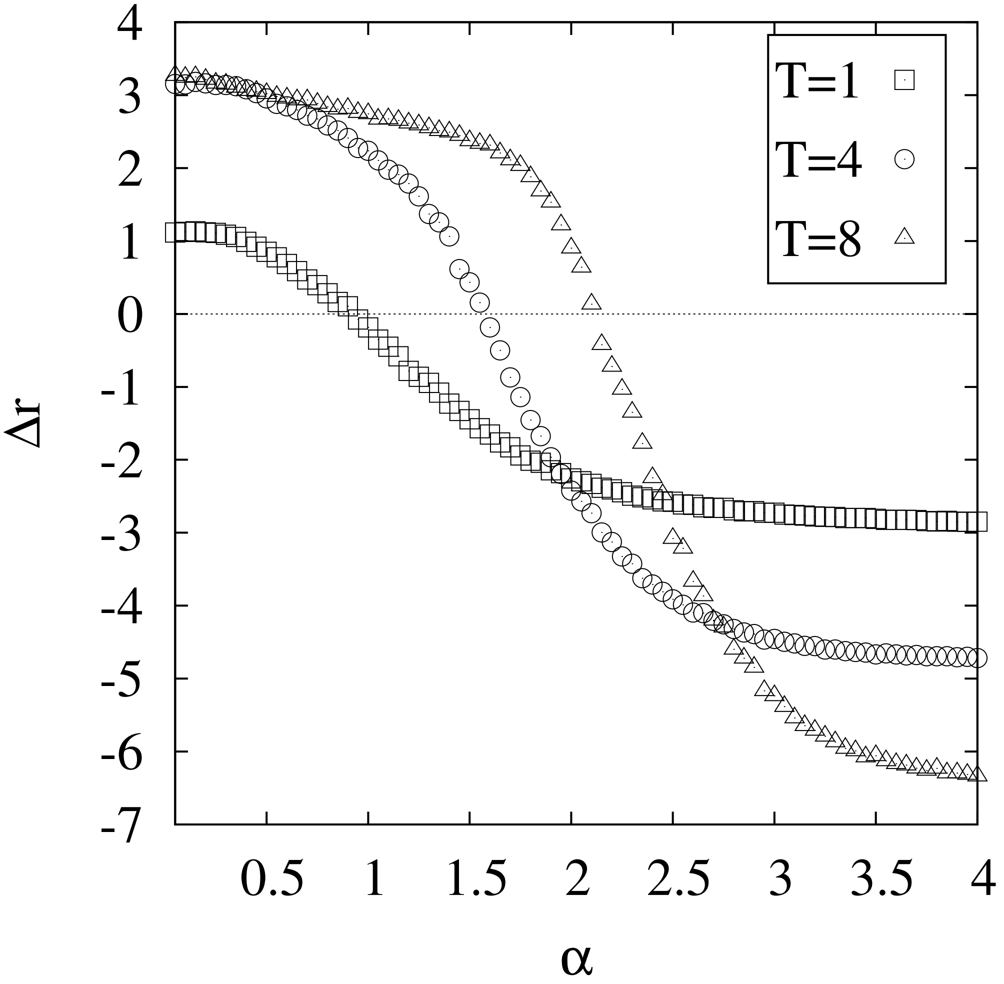

Figure 10.1: Δr versus α for three different time delays in a system with no initial variance in the mutation rate and 25% initial fitness. The locations of the zeroes correspond to selective pressures on mutation rate for the different values of the time delay.

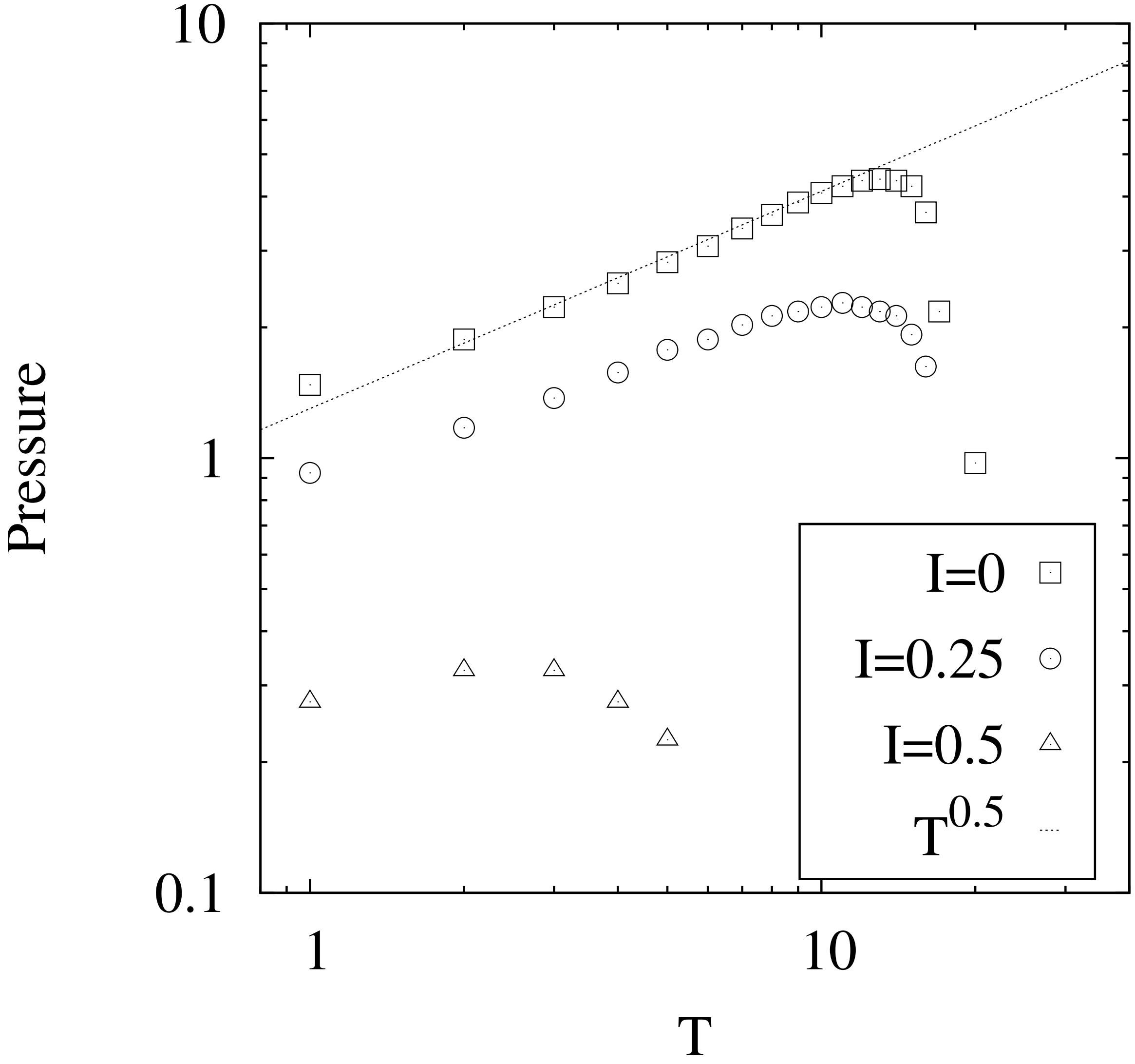

Figure 10.2: Selection pressure α versus time delay T in a system with no initial variation in the mutation rate for different values of the initial fitness I. For $I = 0$, the genome is initially entirely comprised of zeroes, whereas for $I = 50\%$ it is an even mix of zeroes and ones. As such, the automatic benefits of mutation decrease as I increases.

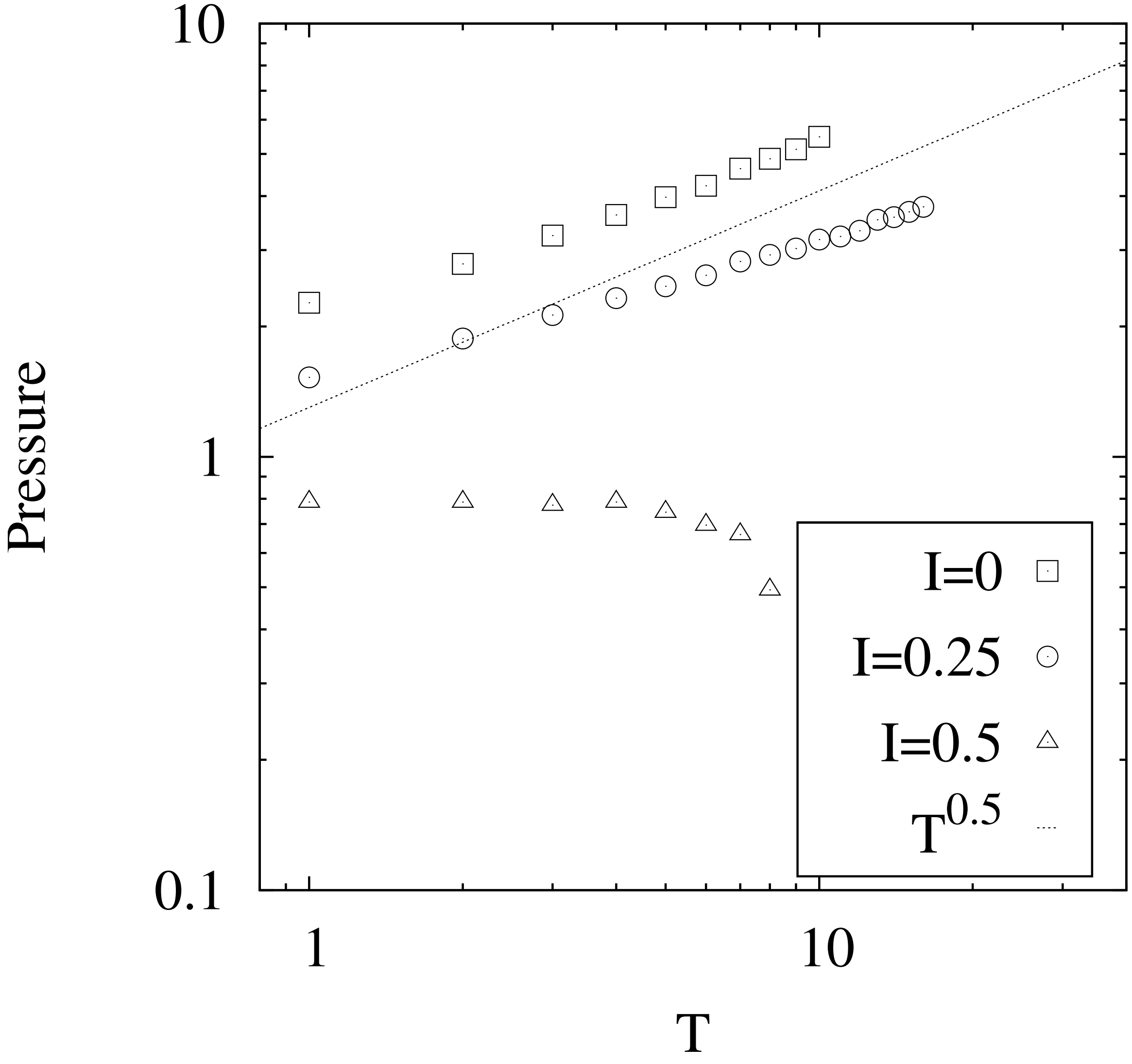

Figure 10.3: Selection pressure α versus time delay T in a system with initial variation in the mutation rate, but no mechanism for changes of the mutation rate aside from selection. In this case, it is observed that at $I = 50\%$, the selection pressure is initially independent of the time delay, as predicted from Eq. 10.21

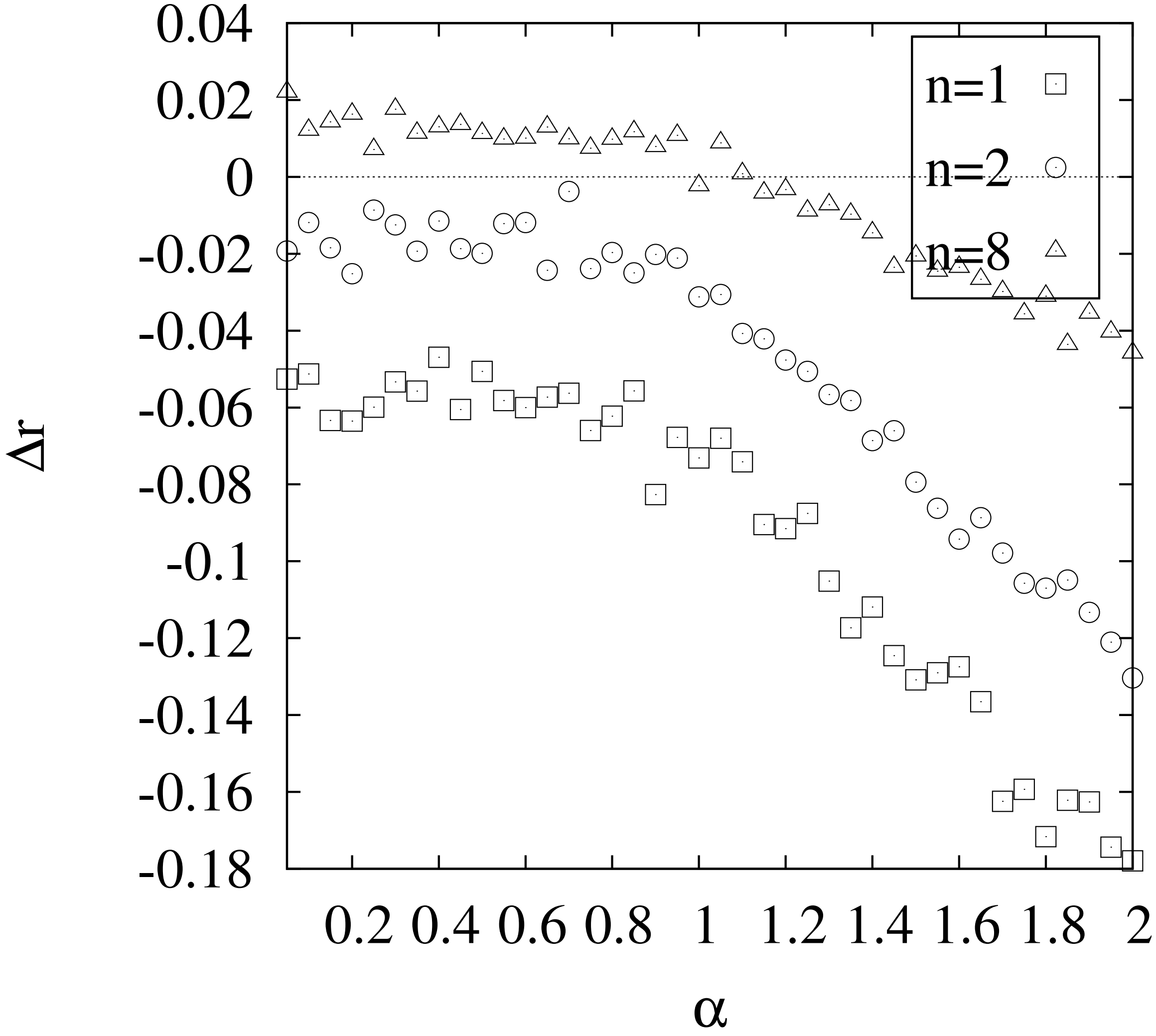

Figure 10.4: Δr versus α for three different group sizes in a system with no initial variance in the mutation rate and 25% initial fitness. For group sizes below a certain threshold size, increases in the mutation rate are selected against. The selection on mutation rate increases discontinuously across this threshold, as evidenced by the plateau in the Δr versus α curves.

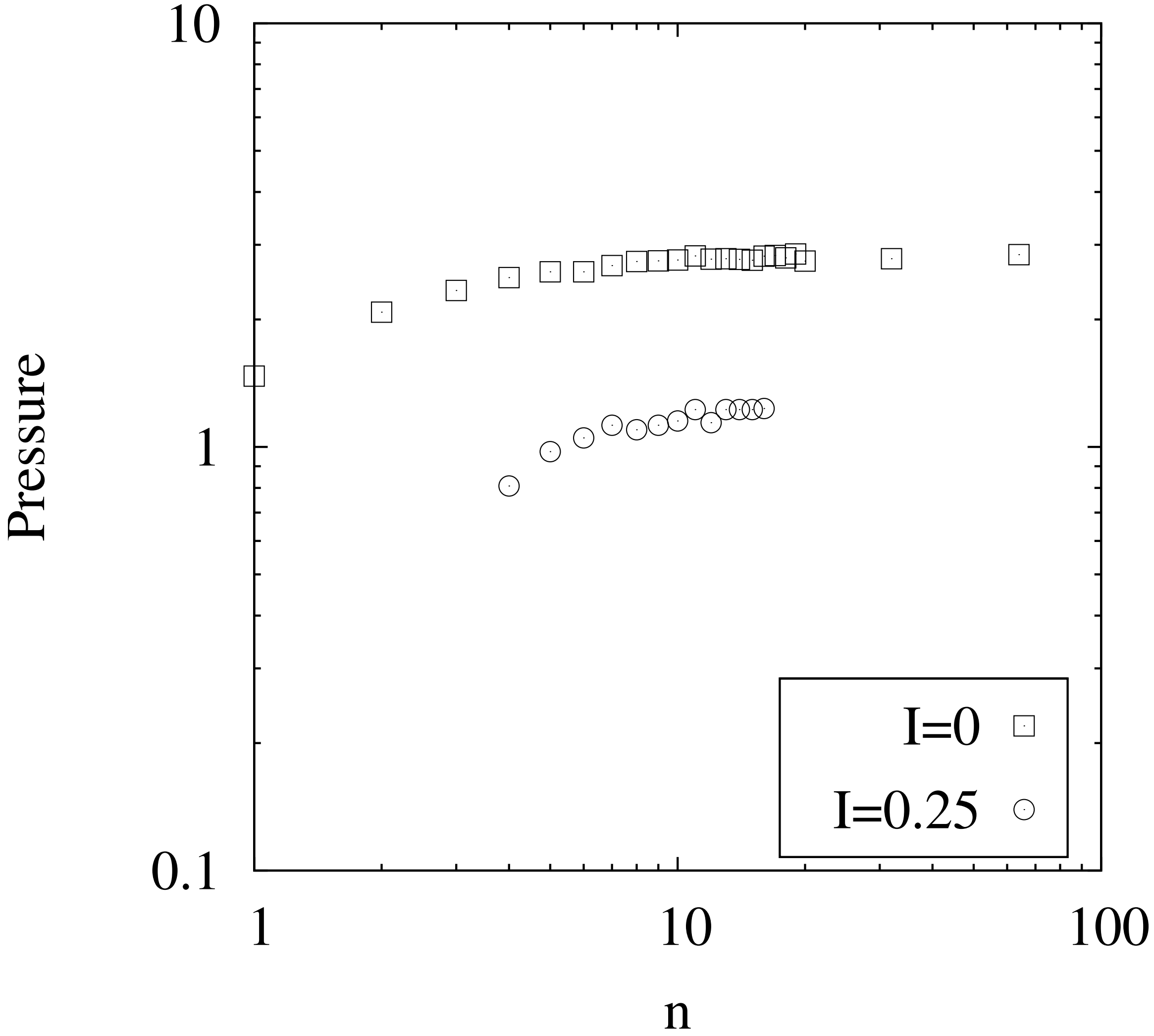

Figure 10.5: Selection pressure α versus group size n in a system with no initial variation in the mutation rate for different values of the initial fitness I. For $I = 0$, mutation rate is selected for even at the minimum group size. For $I = 25\%$ there is a threshold at $n = 4$ below which the mutation rate is selected against, and above which the mutation rate is selected for.

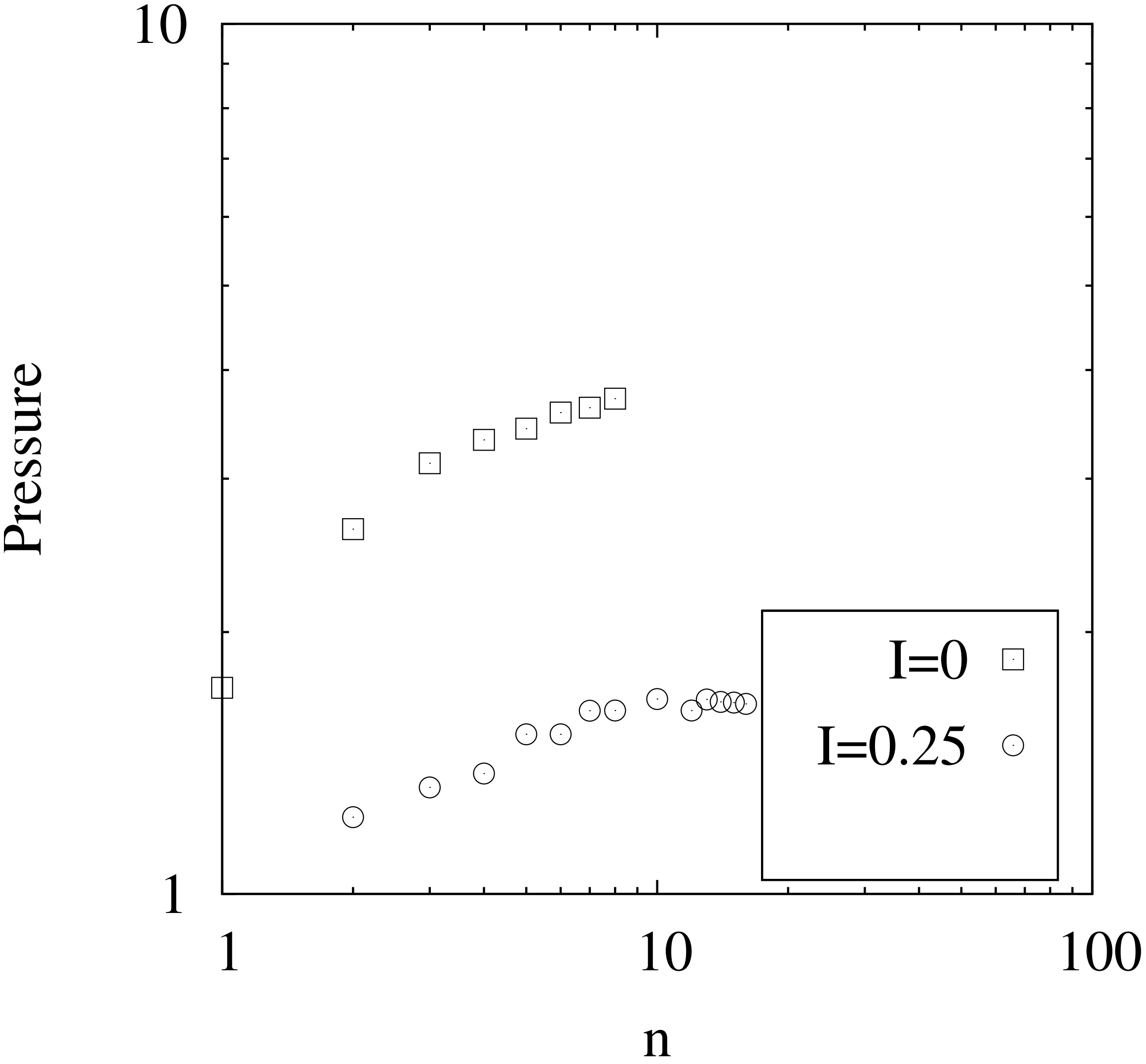

Figure 10.6: Selection pressure α versus group size n in a system with initial variation in the mutation rate but mechanism for drift in the mutation rate, for different values of the initial fitness I.

Chapter 11

Emergence of Politics

To understand the systems of social and political organization in the world, it is tempting to begin by trying to understand the individuals that form them. This approach quickly runs into a problem—the behavior of individuals is very hard to predict. The behavior of any given individual depends upon a large number of factors: their culture, their experiences to date, their genetics, the events they are currently experiencing, their education, their economic status, and so on. It seems as if understanding the behavior of a group is an impossible goal if predicting the behavior a single person is so difficult. However, models of group behavior through agent-based modelling[144] have been reasonably successful despite this, reproducing generic properties of the dynamics of crowds, mobs, and riots[145–147]; collective opinion formation[148–150]; the structure of social groups[151–158]; and financial markets[159]. When large numbers of people interact there exists the possibility for the emergence of collective effects which are surprisingly insensitive to the details of the elements which comprise them.

The purpose of this chapter is to understand the factors at work in setting up and maintaining the large scale structure of societies from the point of view of an abstract model. Other models[160] have analyzed the stability and transitions of an established form of social order. In this paper, we will instead seek to explain how social order emerges from an unstructured state due to collective interactions between individual agents. This must take into account that the connections between individuals may change, leading to a situation in which one has an active network[161].

The emergence of networks of preferred interactions between agents has been observed in [162; 163]. The resultant structure of agents is heterogeneous—a state emerges in which some subset of the population (the leaders) extracts maximal benefit. There is however no explicit flow of information from the leaders to the other agents. We posit that the structure of information exchange in the system is a key element to the form of political organization it possesses. We would like to differentiate between the agent with the greatest payoff and the agent whose decisions hold maximal weight in influencing the decisions of others. In our model, we observe the development of a division of labor from simple selfish behavior and communication between the members of the system. The role of active information is central to achieving this heterogeneous population structure. Information exchange has been studied in various other models. However, in such models it is usually a passive variable, for instance in voting and opinion formation models.

The role of active information—information used to make a decision with either positive or negative consequence—is less well known. In [164], active information played the role of a diffusive field in a spatial prisoner's dilemma model, and in [165], information was given to a subset of members of a swarm to see how informed decisions would propagate to determine the swarm direction. In these cases, the agents had no way of evaluating the quality of the information they received—whether it had in the past led to a good or bad decision. This leads to information acting primarily as a homogenizing agent: it determines the average behavior in [164], and directs the average swarm direction in [165]. In our model, each agent determines the optimal degree of trust to place in information received from another. By giving each agent the ability to tune its trust in the other members of the system, it is possible for clusters to form in which the members of the cluster have voluntarily given over the reins of their decision making to a leader of their choice.

This organization, in its simplest form, arises from uniform information exchange between the individuals in the system, resulting in a homogeneous, shared information pool. This corresponds to communal decision-making by majority vote. In a system in which different agents are better or worse at making decisions, one would expect the emergence of a system of weighting by reputation, simply as a tool to optimize the decision-making process. If, however, resources can be allocated towards making better decisions, it becomes possible for a subset of the individuals to specialize in being an information source. At this point, the majority of agents in the system will be following instructions provided by a minority of agents, without a significant information flow in the reverse direction. These two phases—unstructured and structured respectively—are distinct forms of political organization, and which is achieved depends on the costs and benefits associated with information generation.

A requirement for stability in the structured phase is that the agents which are acting as an information source must either gain from producing information or lose if they fail to produce information, as they dedicate their own resources into providing this information. In modern governments, systems of taxation subsidize the decision-makers, but the emergence of such structures is difficult without a heterogeneous system already being in place. Our results show that in certain circumstances, the decision-making structure of a population may become heterogeneous even without the inclusion of subsidies or resource exchange, due to a collective effect where the refusal to generate information by the majority of the population forces the agents that are the last to act to take on the decision-making role simply to preserve their own benefit. From this phase, the introduction of a resource subsidy would improve the efficiency of the system, and could be done in a continuous manner. A schematic phase diagram that qualitatively exhibits the nature of the phases and transitions between them is illustrated in Fig. (11.1).

11.1 Model

We propose the following model to capture the dynamics of information exchange. The system consists of a set of agents, each which can choose to distribute resources to any other agents. In addition, each agent chooses to allocate its time between producing resources or producing information about the environmental state ('thinking'). Whether or not resource production is successful depends on the accuracy of the agent's guess as to the current nature of its environment, which is randomly in one of O possible states. If the agent guesses the environmental state correctly, it produces a number of resources proportional to the fraction of its time it allocated to production. Furthermore, an agent can look to see what other agents are guessing in order to determine its own guess.

We assume in this first part that each agent has a number of degrees of freedom (how to combine information from other agents, how many resources to distribute to other agents, and how much time to allocate towards producing resources) which are adjusted in order to maximize its average score. The immediate consequence of this is that we may determine trivially what the trust network should be, and thus determine our trust-network order parameters in terms of the distribution of 'thinking' values—the agent with the highest thinking value will have the most trust directed at it, and if the thinking values are distributed homogeneously then trust will also be distributed homogeneously. This treatment neglects dynamical effects and fluctuations. Later, we will analyze the effect of fluctuations and dynamics on the stability of the various phases.

The base accuracy—that due to the agent's own production of information, is a nonlinear function of the fraction of time dedicated towards information production

T. A successful guess then produces $1 - T$ resources. The choice of functional form must satisfy a number of constraints. The accuracy should monotonically increase with the fraction of time dedicated towards it. Additionally, it is bounded above by 1 and below by $1/O$ (the accuracy of a random guess). Given these constraints, we may choose any function of the form $A = (1/O + (1 - 1/O)f(T))$ where $f(T)$ is a monotonically increasing function that maps the interval $[0, 1]$ to itself.

The key character of our choice of function will be its concavity—other details should not strongly influence the results. If $f(T)$ is concave up, then specialization is favored. If concave down, then there are diminishing returns and even an infinitesimal amount of time dedicated towards producing information will be beneficial.

While we could in principle combine an arbitrary number of concave up and concave down regions, it is hard to justify that arbitrary complication. A simple choice of function that allows us to smoothly vary between concave up and concave down behavior with a single parameter is $A = (1/O + (1 - 1/O)T^{\alpha})$, where α is a parameter of the model controlling the nonlinearity. If $\alpha > 1$ then the function is concave up, and specialization is favored. With this basis, we can discuss a number of possible system configurations and evaluate their average score for optimal choices of T.

11.1.1 Disconnected, Homogeneous Phase

In the case that no agent in the system uses information from any other agent, there is an optimal value of T to maximize an agent's score. The average score in this phase $\bar{S}_{DH}$ is perforce independent of N.

$$\bar{S}_{DH}(T) = \frac{1}{O}(1 - T)(1 + (O - 1)T^{\alpha}) \tag{11.1}$$

The optimal value of T satisfies:

$$T^{\alpha-1}(\alpha-(\alpha+1)T)=\frac{1}{O-1} \tag{11.2}$$

If $\alpha = 1$, then this value of T is always less than zero, so $T = 0$ is the optimal choice. At larger values of α a local maximum appears in the curve at a finite value of O, and then becomes a global maximum as O increases. The value of O at which the maximum value of the score is equal to the value at $T = 0$ is $O = 1 + \alpha^{\alpha}/(\alpha-1)^{\alpha-1}$. In the limit of large α, this becomes $O \approx e\alpha + (1 - e/2)$. So in effect, for values of $\alpha > 1$ (representing a nonlinear reward for dedicating resources to 'thinking') there is a first order transition between a 'guessing' phase and a 'thinking' phase, where the more options there are, the more valuable a resource spent on 'thinking' is. The larger α is, the larger O must be for a non-zero thinking phase to be optimal. The score function for various values of α and O is plotted in Fig. 11.2.

An additional consideration is the effect of fluctuations on this phase. If each agent may only specify their actual thinking value to within some standard deviation, then the resulting average score is lower than if fluctuations had been absent. Near the limits of the range of the thinking variable fluctuations are constrained such that they may not take it outside of the range. For fluctuations of magnitude σ around an optimal value of T, we expect that the average score will change by:

$$\Delta S = \sigma^2 \frac{d^2S}{dT^2} = \sigma^2 \alpha T^{\alpha-2}((\alpha-1)-(\alpha+1)T)\frac{O-1}{O} \tag{11.3}$$

For fixed α, as O becomes large the optimal value of T approaches $\alpha/(\alpha+1)$ and so the decrease in the score approaches:

$$\Delta S = -\alpha \left(\frac{\alpha}{\alpha+1}\right)^{\alpha-2} \sigma^2 \tag{11.4}$$

When the optimal solution is $T = 0$, however, the first derivative is non-zero and so fluctuations have a linear effect. The effect of this is that $\Delta S = \sigma \frac{dS}{dT}/\sqrt{\pi}$ assuming Gaussian fluctuations. The slope of the score function around $T = 0$ is:

$$\frac{d\bar{S}_{DH}}{dT} = -1/O \tag{11.5}$$

So we expect $\Delta S = -\sigma/(O\sqrt{\pi})$ to be the leading effect at this point. The consequence of this is that sufficiently large fluctuations will favor the $T = 0$ phase.

11.1.2 Connected, Homogeneous Phase

If communication between agents is permitted, but no resource reallocation takes place, then the resulting accuracy is higher than any of the individual accuracies in the system (so this phase is always favored over the disconnected phase for permitted values of T).

For $O = 2$, the effective accuracy can be solved for in the large N limit. If the initial accuracy is A, then the total number of agents that pick the correct option C_0 is $C_0 = \sum_i^N \eta_i$ where η_i is either 1 (with chance A) or 0 (with chance $1 - A$). In the large N limit, C_0 is described by a Gaussian distribution with mean NA and standard deviation $A(1 - A)\sqrt{N}$.

The probability that the system picks the correct option is thus the probability that $C_0 > N/2$. As the range of permitted values is not infinite, care must be taken to compute the correct normalization factor. So:

$$A_{eff} = \frac{\int_{N/2}^{N} \exp(-(\frac{x-AN}{2NA(1-A)})^2)}{\int_0^N \exp(-(\frac{x-AN}{2NA(1-A)})^2)} \tag{11.6}$$

which evaluates to

$$A_{eff} = \frac{\mathrm{erf}(\frac{(1-A)\sqrt{N}}{\sqrt{2A(1-A)}}) - \mathrm{erf}(\frac{(1/2-A)\sqrt{N}}{\sqrt{2A(1-A)}})}{\mathrm{erf}(\frac{(1-A)\sqrt{N}}{\sqrt{2A(1-A)}}) - \mathrm{erf}(-\frac{A\sqrt{N}}{\sqrt{2A(1-A)}})} \tag{11.7}$$

where $A(T) = (1 + T^\alpha)/2$ in this case.

For $O = 2$ and $\alpha = 2$, benefits from a non-zero value of T do not appear until around $N > 35$. Figure 11.3 shows the score function for the homogeneous, connected phase compared with the isolated phase.

The effects of fluctuations are less obvious in this case, because they must be considered each agent independently, whereas this analysis is done for all agents behaving in the same fashion. In the case of this model, fluctuations may actually increase the effective score, as a fluctuation to higher thinking rate in one agent benefits the guesses of all other agents. Similarly, a decrease in thinking rate in one agent will not significantly decrease his accuracy but may increase his yield. This is a hint that this particular phase is unstable to an inhomogeneous phase.

A rough estimate would suggest that when adding together the effects of fluctuations on each of the individual agents, the effective size of fluctuations is reduced from σ to $\sigma' = \sigma/\sqrt{N}$. This has the consequence that the connected, homogeneous phase is less sensitive to fluctuations than the disconnected phase.

11.1.3 Connected, Inhomogeneous Phase

If the agents become inhomogeneous and divide their labor between thinking and working, then structures in which there is a directional information flow become possible. Given perfect communication and no fluctuations, the optimal configuration will be that of a single agent with high accuracy (T), and $N-1$ agents with minimum accuracy but always picking the action of the 'leader' agent. The average score for this phase is simply:

$$\bar{S}_{CI} = \frac{(N-1) + (1-T)(1/O + (1-1/O)T^{\alpha}}{N} \tag{11.8}$$

This phase in static conditions scores far better than the homogeneous phases, but it is very susceptible to fluctuations lowering the score, compared to the connected homogeneous phase. The result of this is that neither the pure homogeneous nor heterogeneous phases are realized. In a fully-connected population with some form of noise, the system produces a number of leaders L which scales with the population size.

The inhomogeneous phase with a number of leaders can always have a higher average score than the homogeneous phase, but is not generally stable when the individual scores are examined. Each leader agent can improve their score by decreasing the portion of resources they dedicate to thinking to the optimal value for the disconnected phase. When the disconnected phase optimal value is T is greater than zero, the inhomogeneous phase may still occur. This occurs for small α, large O, and small N. If α is too large, the height of the secondary maximum is decreased below that of the $T = 0$ score function maximum and a homogeneous $T = 0$ phase occurs. If N is sufficiently large, the homogeneous connected phase with nonzero T can outperform a phase consisting of a single 'selfish leader'. So there are first order phase transitions in the space of O, α, and N between three phases: $T = 0$, leader, and homogeneous connected (or 'communal' phase). A cross section of the phase diagram at fixed O is plotted in Fig. 11.4, and a cross section at fixed N is plotted in Fig. 11.5.

11.1.4 Resource Subsidy

We have so far shown that for certain values of the parameters, the inhomogeneous 'leader' phase is stable even without the leaders being subsidized. The system has not maximized its resource production in this phase—rather, the limit on resource production is set by the cost to the leader agent, in that even though it might produce a large amount of resources for others by changing its behavior, doing so would decrease its own resource production.

If we allow agents to exchange resources as well as information, then starting from the connected, inhomogeneous phase it is possible to improve or keep constant the scores of all agents. If we have a phase with a single leader agent, then for that agent to dedicate more than the disconnected optimal fraction of resources to thinking, it must be reimbursed by at least the same amount of resources as it loses to increase the resources it spends on thinking. This resource cost may then be absorbed by the remaining $N-1$ agents. In effect, the criterion of selfish optimization becomes one of global optimization. The globally optimal phase in the absence of fluctuations is that with a single leader agent.

This need not be the case in general, as one may posit the existence of cheaters: agents which do not give resources towards the subsidy but still gain its benefits. A system with multiple types of resource or multiple agendas, such as in [166] might also retain a more detailed structure.

When fluctuations are added, it becomes beneficial to have multiple leaders in order to reduce the impact of fluctuations but retain the benefit of increased efficiency. We use the connected, homogeneous solution for L agents to determine the accuracy of the remaining $N-L$ given a known accuracy of the leaders. For simplicity, we will assume that the leader agents have $T = 1$, which the optimal choice converges to as $N \gg L$. For a given level of fluctuations, each leader will have an effective, adjusted accuracy. We evaluate the score function numerically as a function of L and find the location of the maximum as a function of N. The results are plotted in Fig. 11.6. At large N, the optimal number of leaders approaches $L \propto \log(N)$.

For a spatially distributed system, or one in which there is not total connectivity, it is expected that such effects will require a larger number of leaders to cover the system extent. For example, in a two-dimensional system in which agents can only communicate within a radius R, a number of leaders proportional to $\sqrt{N}/\pi R^2$ would be expected to ensure total coverage.

11.2 Applications

The abstract model of emergent political systems that we have outlined is capable of providing a framework in which to analyze real social systems, and in this section we briefly indicate some examples. It is important to emphasize that our model is not required to model all situations in which division of labor occurs—a simpler model with only a nonlinear benefit to specialization and some form of exchange of services would be sufficient to enable division of labor to emerge. Such a process would not need to involve information sharing as a core element. On the other hand, our work shows that the emergence of the leader phase (which corresponds to the occurrence of a division of labor in other pictures) is primarily a consequence of the special property of information when compared to other resources that, once created, it can be duplicated with a much smaller additional cost than the cost to first generate it. This process of information amplification makes the leader phase described here distinct from other scenarios that produce division of labor.

We must also be careful to understand the nature of the relevant optimization being implicitly performed when considering a given system. In human economic and political behavior, one considers that each individual tries to maximize its personal benefit in the context of the greater system. In other systems, such as foraging insects, the net benefit to the colony as a whole is what is likely maximized—this corresponds to the case where resources may be redistributed, which in our model means that the leader phase is always optimal for all parameter values.

Even with these caveats, there are several systems which could potentially be understood in the context of our model: the behavior of social and hierarchical insects compared to asocial insects[167–169], the distribution of information in swarms[170], and innovation-sharing in unicellular organisms via horizontal gene transfer[106; 171]. All of these cases involve some piece of information being discovered by a single individual—a randomly chosen one of a set of similar individuals in the case of swarm behavior (corresponding to the homogeneous phase), or via directed searching by a specialized subset of the population, as is the case in some foraging insects (corresponding to the leader phase). We will now briefly discuss each case.

Different species of insects are socialized to different degrees. On one extreme, there are insects such as the solitary wasps[167], which do not share resources or information. On the other extreme, eusocial insects such as bees, ants, and certain kinds of wasps have highly structured communication channels and vehicles of information discovery. Foragers and scouts use various means to communicate the location of food supplies or nesting sites. The distinction here seems to be that bees and ants reproduce centrally via a queen, and so maximizing their interest corresponds to maximizing the interest of the queen. As a result, resources can be redistributed freely, and so we expect the system to emerge in the leader phase.

In the case of honey bees, the various scouts return with information about potential food locations, after which the swarm comes to a unified decision about which site to pursue. The method of decision making seems to be a weighted average[170], similar to what we use to model the decision making of our agents. Each scout has a certain chance of finding the best site within a given distance—even if they spend 100% of their time searching, they have a limited maximum accuracy. This corresponds to the fluctuating case in our model, so, as the swarm size grows, we can predict that the optimal number of scouts should scale logarithmically with the swarm size.

Microbial organisms[172; 173] and even multicellular eukaryotes[174–177] have the ability to swap genetic material and integrate it into their genomes via several pathways, mediated often by mobile genetic elements such as viruses and plasmids. There is cellular machinery associated with this process, which can be active or inactive in a given cell. In microbes, the state in which such an organism is receptive to external DNA is called genetic competence. The regulatory network associated with competence has been shown to generate a distribution of cells with differing levels of competence[171]. A small subset of the cells at any given time end up being receptive to this information exchange, whereas the rest remain closed. The competent subset changes with time, so eventually, all cells will at some time be able to accept foreign genetic material. This dynamic may be analogous to the leader phase in our model. Here, the information amplification takes place when a subset of cells exchange material and either live or die as a result. The surviving exchanges are then passed on to the local population, amplifying the induced information.

11.3 Conclusions

We have shown that a model of communicating agents that divide their time between information generation and information usage has three distinct phases of organization corresponding to structures identifiable in human political systems. The flow of information between agents in the system is critical to this phase structure. If agents can exchange resources in a way that does not permit cheating, then the optimal structure is to have a small number of leaders that scales logarithmically with the system size, and a larger number of workers. Fluctuations in the reliability of agents tend to emphasize the communal phase over the leader phase.

The phase transitions predicted by this model are all first order in nature. As such, in a situation in which the agents are approaching equilibrium dynamically, the various phases can coexist over much of the parameter space. This makes sense when one looks at the diversity of actual political systems in existence, on both the local and national scales. The transition to the leader phase from a communal phase takes the form of an inhomogeneous decay in the levels of decision making of the agents in the system, leaving one agent in charge by default. In a dynamical version of this model in which the distribution of agents changes with time, the transition between different leader agents could be studied.

This model has a relatively simple phase structure, as only the thinking value and trust levels are allowed to vary. The addition of spatial considerations, information exchange costs, lying, resource exchange with cheating, or other such factors could vastly increase the diversity of phases exhibited by the model.

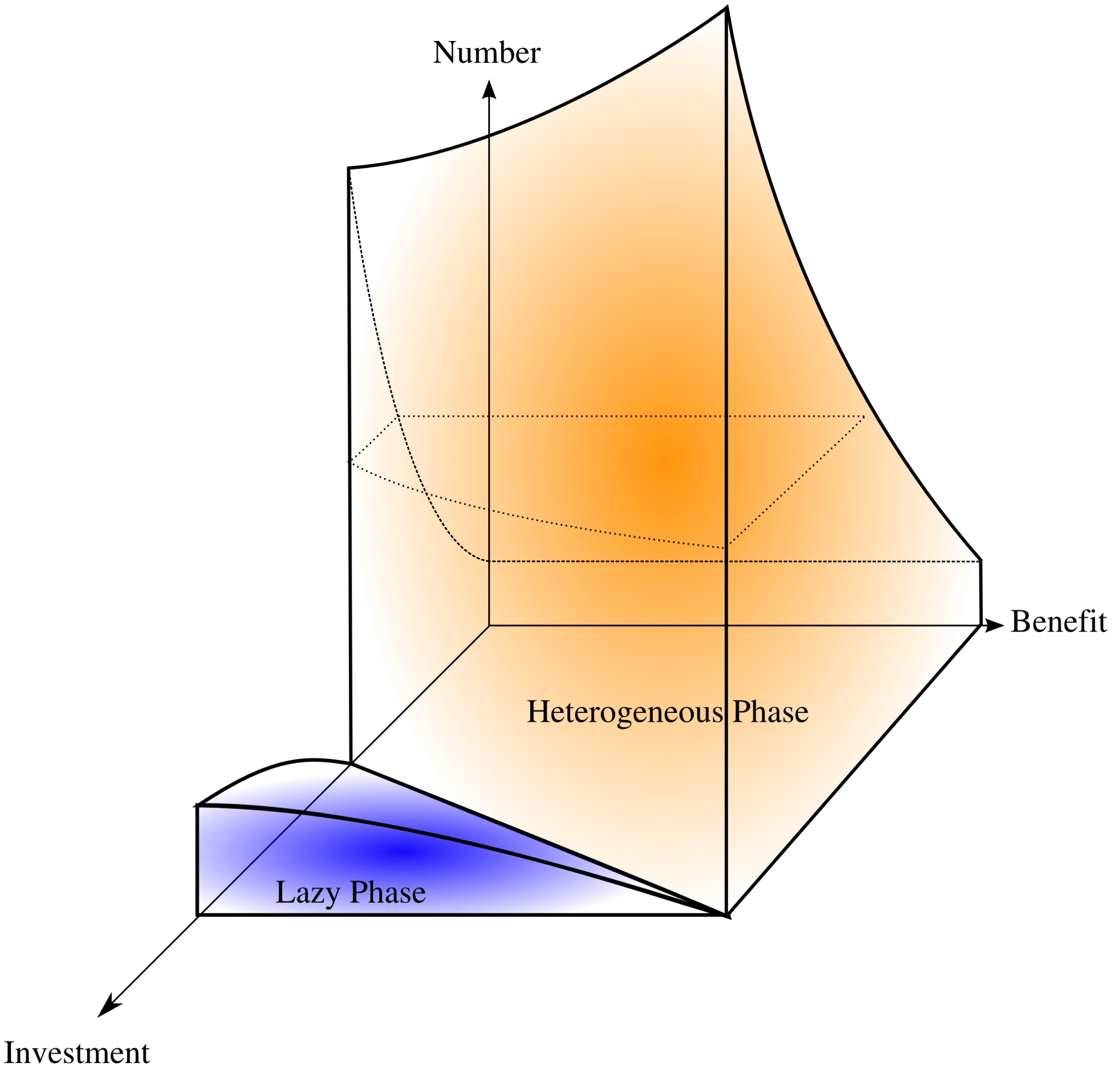

Figure 11.1: Schematic phase diagram for our model. The Investment axis is the degree that a large initial investment of resources is needed to see an improvement in accuracy: this corresponds to the nonlinearity α in the model. The benefit axis is the total difference in accuracy between random guessing and perfect knowledge, which corresponds to the variable O in our model. In the Lazy Phase, random guessing is the optimal behavior. In the Heterogeneous Phase, a subset of agents dedicate their resources to thinking whereas the rest of the agents dedicate their resources to working (division of labor). In the Homogeneous Phase), all the agents dedicate the same non-zero amount of resources to thinking.

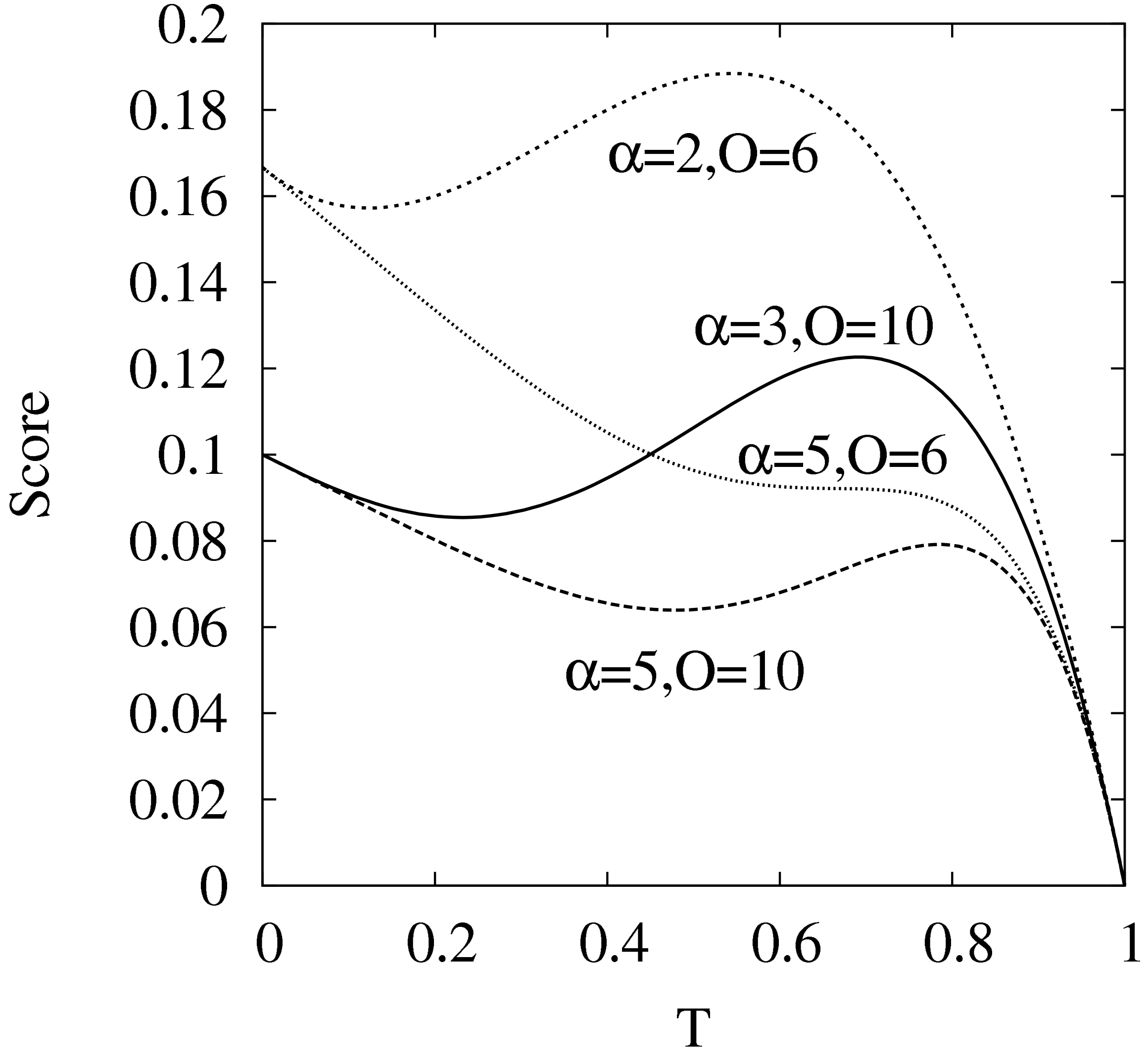

Figure 11.2: Score functions in the Isolated Phase as a function of thinking time T, for four values of O and α. In the Isolated Phase, each agent does not receive information from other agents.

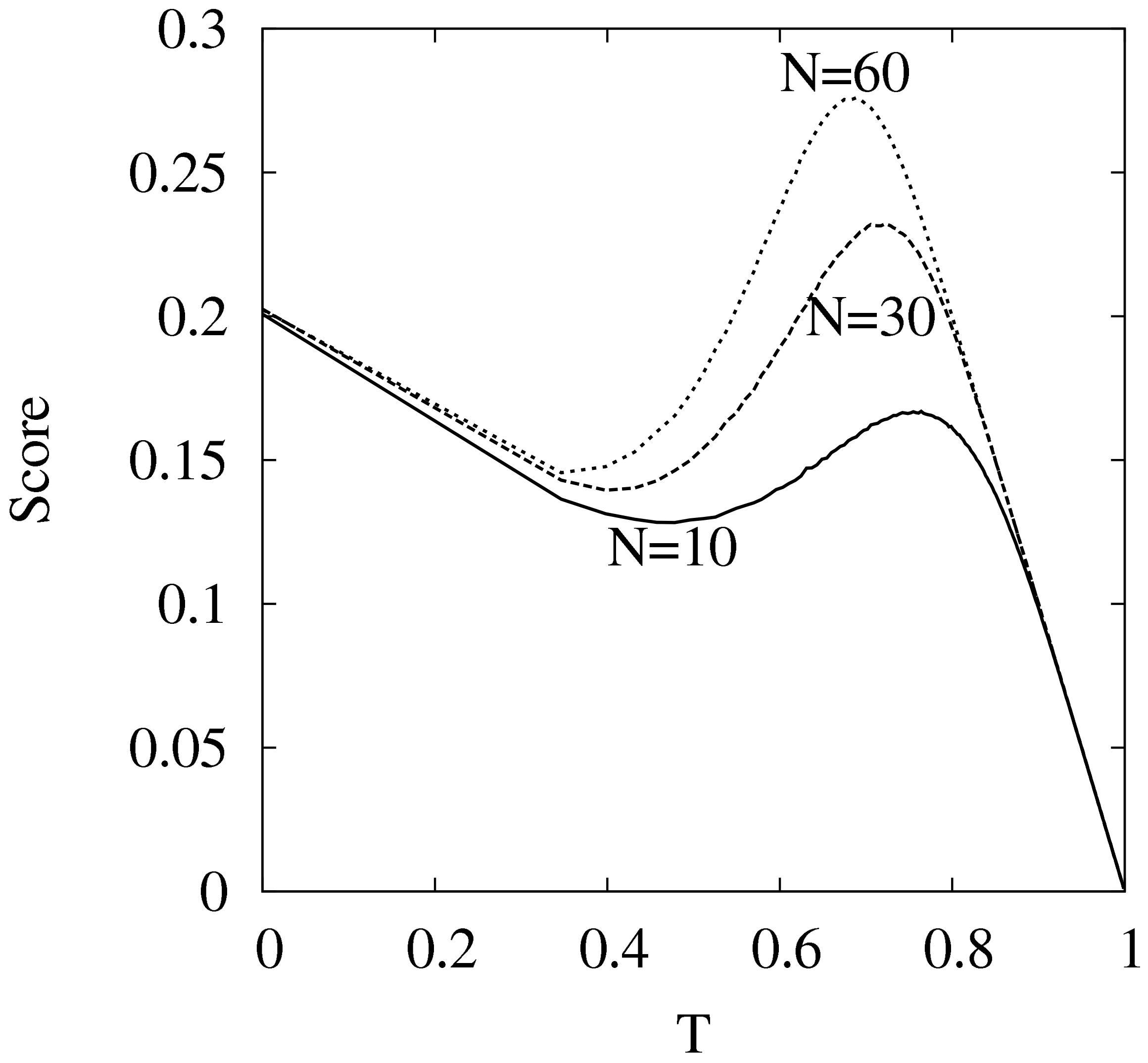

Figure 11.3: Score functions in the Homogeneous Phase as a function of thinking time T, for different numbers of agents N, with $O = 5$ and $\alpha = 5$. In the Homogeneous Phase all agents share information and have the same parameters.

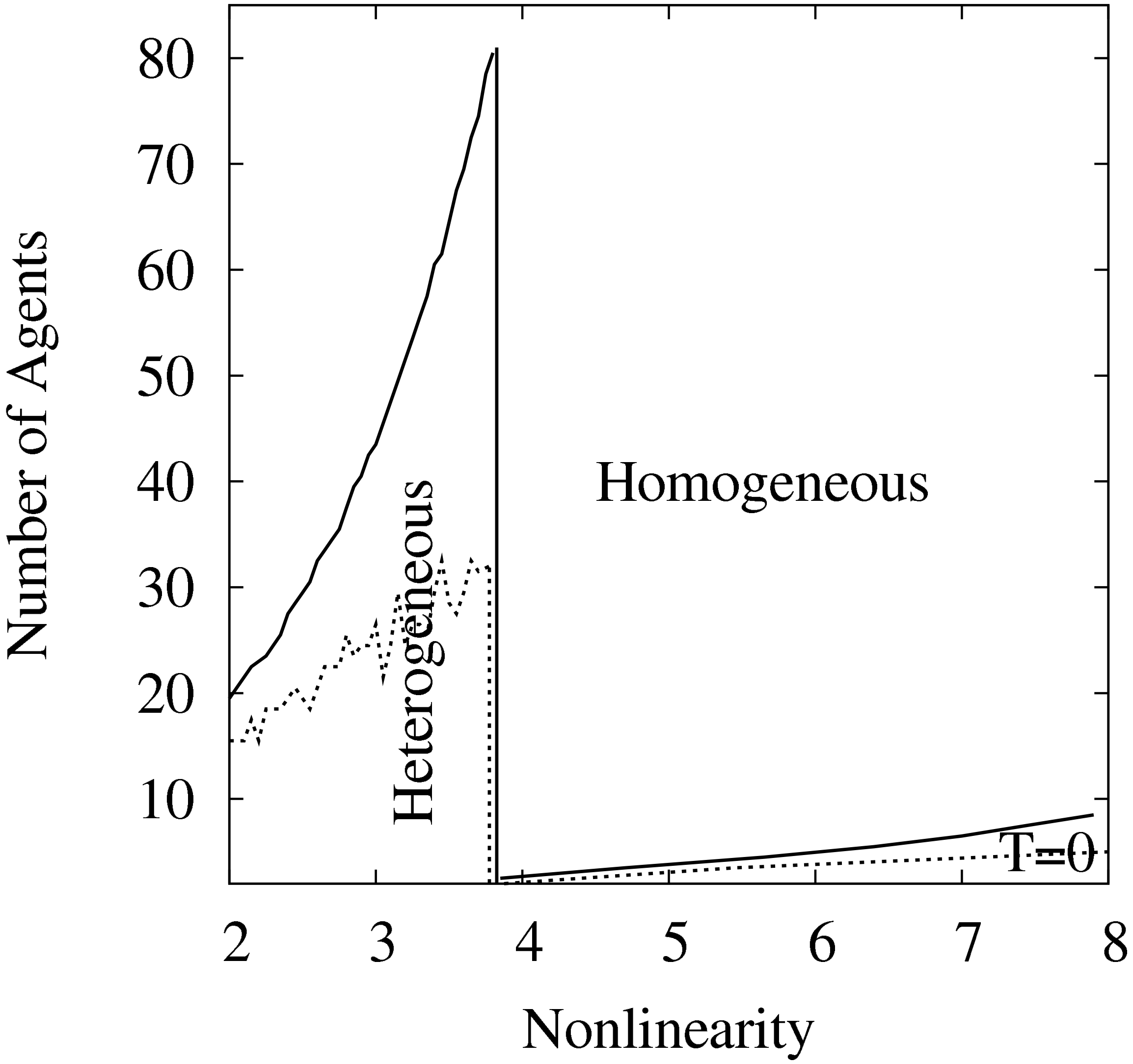

Figure 11.4: Phase diagram for $O = 10$ in the space of the nonlinearity α and number of agents N. The phase transition from the heterogeneous phase as N increases is due to the communal phase being more efficient than a selfish leader phase. The phase transition as α increases is due to the transition of the isolated phase to a $T = 0$ phase. The dotted lines show the phase boundaries when Gaussian fluctuations with a standard deviation of 0.1 are added to the T value of each agent.

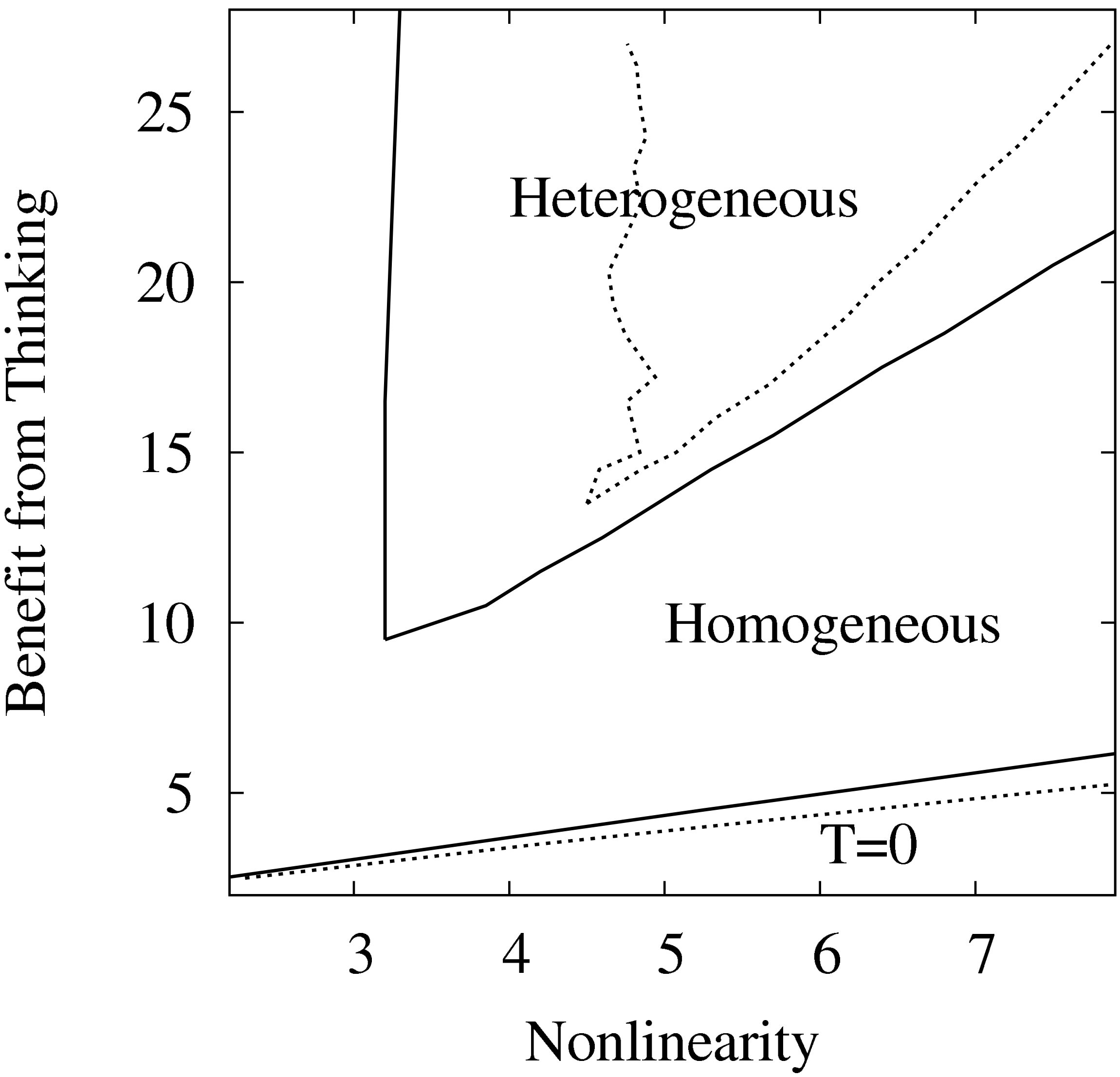

Figure 11.5: Phase diagram for $N = 50$ in the space of the nonlinearity α and thinking benefit O. The dotted lines show the phase boundaries when Gaussian fluctuations with a standard deviation of 0.1 are added to the T value of each agent.

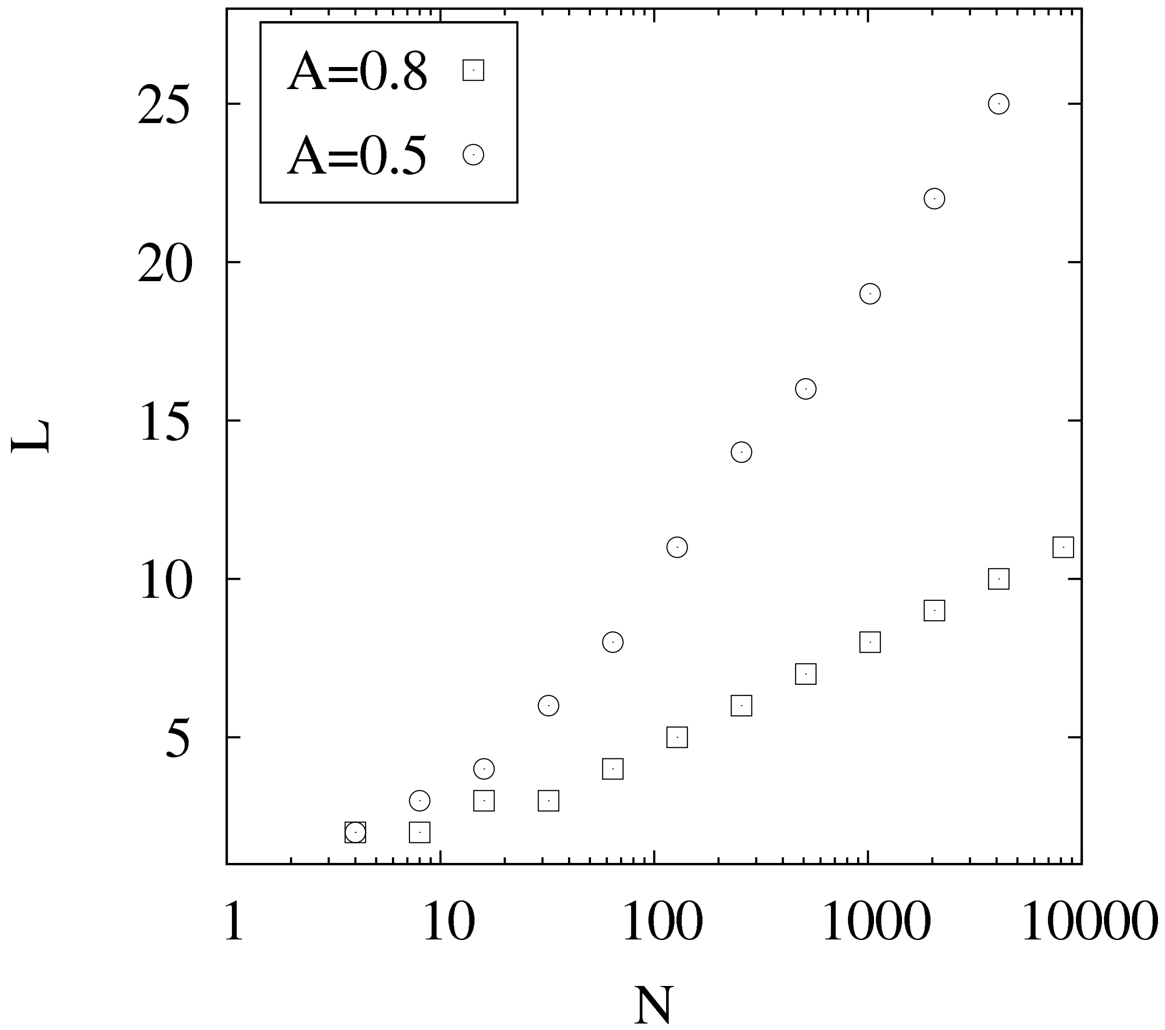

Figure 11.6: Optimal number of leaders as a function of total number of agents for a system with $O = 5$ and two different fluctuation strengths. In this case, the fluctuations are parameterized by the resultant average accuracy A of a leader with $T = 1$.

Chapter 12

Conclusions

In this thesis I have discussed several different systems in which multiscale structure emerges from relatively simple local rules. In turbulence, these rules are the transport of momentum by a flow; in evolution, these rules are the way in which organisms interact with each other and their environment. In both of these cases, the determining factor for the structures which emerged was not some small detail of the rules, but rather the broad symmetries that the rules possessed — Galilean invariance and energy conservation (and enstrophy conservation in 2D) in the case of turbulence, and complexity invariance in the case of the complexity cascade in evolution. The universality of these cascades connects with the ideas of phase transitions and criticality, where only certain dominant degrees of freedom remain relevant to the behavior of the system.

Furthermore, it is important to understand the scale-free structure in each of these cases in order to understand the actual macroscopic behaviors that are more immediately observable. In the case of turbulence, I showed that the drag friction and velocity profiles of pipe flows depends on the particular form of the turbulent cascade present — either energy or enstrophy. Consequently the way in which the turbulence was generated determines its properties in a fashion not predicted by the classical theory. The fluctuations of flow properties are also tied to the scale-free structure of the flow, in a way that informs the transition into the turbulent state.

The consequences of an evolutionary complexity cascade are readily apparent in the richness and complexity of life on Earth. Whether by Red Queen dynamics[131] or a symbiotic cascade or some other invariant pathway, life on Earth has explored and created a hierarchy of niches, something more structured and complex than 'the simplest, fastest replicator'. By understanding the symmetries that underly this complexification, it is possible to begin to ask questions about the distributions of diversity that it engenders, the rate of evolution, and the history of evolution of Earth. Additionally, understanding of the necessary symmetries to achieve a complexity cascade is a useful tool in constructing simulations that exhibit open-ended evolution, either to answer evolutionary questions or as a tool for computer science (that is, as a certain subclass of genetic algorithms).

This work suggests directions for future inquiry. None of the simulations presented in this thesis satisfy the criteria for abstraction transitions. Some physical systems can give rise to quasiparticles on a coarser scale than the underlying dynamics, while others cannot. As an abstraction transition implies a sort of multi-level emergence, understanding the necessary elements for the emergence of evolutionary quasiparticles may allow such a simulation to be constructed. The dynamics of meta-evolution show interesting dependences on system size and timescale, but in none of these cases are the evolutionary pressures on the meta-evolutionary parameters very strong compared to the direct selection pressures. In particular, if the average mutation is neutral or harmful, the meta-evolutionary pressures are incredibly weak. Are there additional elements in real evolving systems that make meta-evolutionary selection stronger, or is real meta-evolution simply a matter of very long times?

References

[1] N. Goldenfeld, "Roughness-Induced Critical Phenomena in a Turbulent Flow," *Physical Review Letters*, vol. 96, pp. 044503–+, Jan. 2006.

[2] G. Gioia and P. Chakraborty, "Turbulent friction in rough pipes and the energy spectrum of the phenomenological theory," *Phys. Rev. Lett*, vol. 96, p. 044502, 2006.

[3] G. Batchelor, *The Theory of Homogeneous Turbulence.* Cambridge University Press, Cambridge, UK, 1982.

[4] R. H. Kraichnan, "Inertial ranges in two-dimensional turbulence," *The Physics of Fluids*, vol. 10, no. 7, pp. 1417–1423, 1967.

[5] J. Nikuradze, "Gesetzm assigkeiten der turbulenten strömung in glatten rohren." VDI Forschungsheft, vol. 356 [In English, in NASA TT F-10, 359 (1966).], 1932.

[6] C. Waters and B. Bassler, "Quorum sensing: cell-to-cell communication in bacteria," 2005.

[7] A. J. Chorin, *Vorticity and turbulence.* Springer, New York, 1997.

[8] K. R. Sreenivasan, "Fluid turbulence," *Rev. Mod. Phys.*, vol. 71, pp. S383–S395, 1999.

[9] A. N. Kolmogorov, "Local structure of turbulence in an incompressible fluid at very high Reynolds numbers," *Dokl Acad Nauk USSR*, vol. 30, pp. 299–303, 1941.

[10] A. M. Obukhov, "Energy distribution in the spectrum of turbulent flow," *Dokl. Akad. Nauk. SSSR*, vol. 32, pp. 22–24, 1941.

[11] L. Richardson, *Weather Prediction by Numerical Process.* Cambridge University Press, Cambridge, UK, 1922.

[12] J. Nikuradze *VDI Forschungsheft*, vol. 361, 1933.

[13] H. Schlichting and K. Gersten, *Boundary-Layer Theory.* Springer, New York, USA, 2000.

[14] B. J. Mckeon, C. J. Swanson, M. V. Zagarola, R. J. Donnelly, and A. J. Smits, "Friction factors for smooth pipe flow," *Journal of Fluid Mechanics*, vol. 511, pp. 41–44, 2004.

[15] M. Shockling, J. Allen, and A. Smits, "Roughness effects in turbulent pipe flow," *Journal of Fluid Mechanics*, vol. 564, pp. 267–285, 2006.

[16] J. J. Allen, M. A. Shockling, G. J. Kunkel, and A. J. Smits, "Turbulent flow in smooth and rough pipes," *Phil. Trans. Roy. Soc. A*, vol. 365, pp. 699–714, 2007.

[17] M. Mehrafarin and N. Pourtolami, "Intermittency and rough-pipe turbulence," *Phys. Rev. E*, vol. 77, p. 055304, 2008.

[18] R. Kraichnan and D. Montgomery, "Two-dimensional turbulence," *Reports on Progress in Physics*, vol. 43, pp. 547–619, 1980.

[19] H. Kellay and W. I. Goldburg, "Two-dimensional turbulence: a review of some recent experiments," *Rep. Prog. Phys.*, vol. 65, pp. 845–894, 2002.

[20] R. Brodkey, *The phenomena of fluid motions.* Gary Tatterson, 2004.

[21] M. Bazant and H. Moffatt, "Exact solutions of the Navier–Stokes equations having steady vortex structures," *Journal of Fluid Mechanics*, vol. 541, pp. 55–64, 2005.

[22] C. Wang, "Exact solutions of the steady-state Navier-Stokes equations," *Annual Review of Fluid Mechanics*, vol. 23, no. 1, pp. 159–177, 1991.

[23] A. Patera and S. Orszag, "Finite-amplitude stability of axisymmetric pipe flow," *Journal of Fluid Mechanics Digital Archive*, vol. 112, pp. 467–474, 2006.

[24] M. Lessen, S. Sadler, and T. Liu, "Stability of pipe Poiseuille flow," *Physics of Fluids*, vol. 11, p. 1404, 1968.

[25] B. Eckhardt, H. Faisst, A. Schmiegel, and T. Schneider, "Dynamical systems and the transition to turbulence in linearly stable shear flows," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 366, no. 1868, p. 1297, 2008.

[26] B. Eckhardt, "Turbulence transition in pipe flow: some open questions," *Nonlinearity*, vol. 21, no. 1, p. 1, 2008.

[27] B. Eckhardt, T. Schneider, B. Hof, and J. Westerweel, "Turbulence Transition in Pipe Flow," *Annual Review of Fluid Mechanics*, vol. 39, no. 1, pp. 447–468, 2007.

[28] S. Grossmann, "The onset of shear flow turbulence," *Reviews of modern physics*, vol. 72, no. 2, pp. 603–618, 2000.

[29] G. Taylor, "The spectrum of turbulence," *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, pp. 476–490, 1938.

[30] R. Brodkey, *The phenomena of fluid motions.* Brodkey Publishing, 2004.

[31] J. Meyers and C. Meneveau, "A functional form for the energy spectrum parametrizing bottleneck and intermittency effects," *Physics of Fluids*, vol. 20, p. 065109, 2008.

[32] S. Pope, *Turbulent flows.* Cambridge Univ. Press, 2008.

[33] Y. Pao, "Structure of turbulent velocity and scalar fields at large wavenumbers," *Physics of Fluids*, vol. 8, p. 1063, 1965.

[34] J. O. Hinze, *Turbulence.* New York, NY: McGraw-Hill Book Company, 2nd ed., 1975.

[35] G. Batchelor, "The Theory of Homogeneous Turbulence," *J. Fluid Mech*, vol. 189, pp. 87–116, 1953.

[36] G. D. Nastrom and K. S. Gage, "A climatology of atmospheric spectra of wind and temperature observed by commercial aircraft," *Journal of the Atmospheric Sciences*, vol. 42, pp. 950–960, 1985.

[37] M. A. Rutgers, "Forced 2D Turbulence: Experimental Evidence of Simultaneous Inverse Energy and Forward Enstrophy Cascades," *Physical Review Letters*, vol. 81, pp. 2244–2247, Sept. 1998.

[38] K. Tung and W. Orlando, "The k- 3 and k- 5/3 energy spectrum of atmospheric turbulence: Quasigeostrophic two-level model simulation," *Journal of the Atmospheric Sciences*, vol. 60, no. 6, pp. 824–835, 2003.

[39] G. Boffetta, "Energy and enstrophy fluxes in the double cascade of two-dimensional turbulence," *J. Fluid Mech.*, vol. 589, pp. 253–260, 2007.

[40] G. Brown, "The history of the Darcy-Weisbach equation for pipe flow resistance," *Environmental and Water Resources History*, vol. 38, no. 7, pp. 34–43, 2002.

[41] H. Blasius *Z. Ver. Dtsch. Ing.*, vol. 56, p. 639, 1912.

[42] A. Strickler, "Beitrage zur frage der geschwindigkeitsformel und der rauhigkeitszahlen fur strome, kanale und geschlossene leitungen." Mitteilungen des Eidgenössischen Amtes für Wasserwirtschaft 16, Bern, Switzerland Translated as "Contributions to the question of a velocity formula and roughness data for streams, channels and closed pipelines." by T. Roesgan and W. R. Brownie, Translation T-10, W. M. Keck Lab of Hydraulics and Water Resources, Calif. Inst. Tech., Pasadena, Calif. January 1981, 1923.

[43] C. Casciola, P. Gualtieri, B. Jacob, and R. Piva, "Scaling Properties in the Production Range of Shear Dominated Flows," *Physical Review Letters*, vol. 95, no. 2, p. 24503, 2005.

[44] L. Prandtl, "Über den Reibungswiderstand strömender Luft," *Ergebnisse der Aerodynamischen Versuchsanstalt zu Göttingen*, no. 1.

[45] R. Panton, "Composite asymptotic expansions and scaling wall turbulence," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 365, no. 1852, pp. 733–754, 2007.

[46] C. Colebrook and C. White, "Experiments with fluid friction in roughened pipes," *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, pp. 367–381, 1937.

[47] M. Rogers and R. Moser, "Direct simulation of a self-similar turbulent mixing layer," *Physics of Fluids*, vol. 6, no. 2, pp. 903–923, 1994.

[48] L. Prandtl, "Zur turbulenten Stromung in Rohren und langs Platten," *Ergebn. Aerodyn. Versuchsanst. Gottingen*, vol. 4, pp. 18–29, 1932.

[49] T. Von Kármán, *Mechanische Ähnlichkeit und Turbulenz.* Weidmannsche Buchh, 1930.

[50] N. Guttenberg and N. Goldenfeld, "Friction factor of two-dimensional rough-boundary turbulent soap film flows," *Physical Review E*, vol. 79, no. 6, p. 65306, 2009.

[51] A. Townsend, "A. 1976 The Structure of Turbulent Shear Flow," 1976.

[52] O. Reynolds, "An Experimental Investigation of the Circumstances which determine whether the Motion of Water shall be Direct or Sinuous and the Law of Resistance in Parallel Channel," *Philos. Trans. R. Soc. London*, vol. 174, p. 935, 1883.

[53] M. Lagha and P. Manneville, "Modeling transitional plane Couette flow," *The European Physical Journal B*, vol. 58, no. 4, pp. 433–447, 2007.

[54] A. Willis and R. Kerswell, "Critical Behavior in the Relaminarization of Localized Turbulence in Pipe Flow," *Physical Review Letters*, vol. 98, no. 1, p. 14501, 2007.

[55] F. Daviaud, J. Hegseth, and P. Berge, "Subcritical transition to turbulence in plane Couette flow," *Physical Review Letters*, vol. 69, pp. 2511–2511, 1992.

[56] S. Bottin and H. Chate, "Statistical analysis of the transition to turbulence in plane Couette flow," *The European Physical Journal B-Condensed Matter*, vol. 6, no. 1, pp. 143–155, 1998.

[57] B. Hof, C. van Doorne, J. Westerweel, F. Nieuwstadt, H. Faisst, B. Eckhardt, H. Wedin, R. Kerswell, and F. Waleffe, "Experimental observation of nonlinear traveling waves in turbulent pipe flow," *Science*, vol. 305, no. 5690, pp. 1594–1598, 2004.

[58] J. Peixinho and T. Mullin, "Decay of turbulence in pipe flow.," *Physical Review Letters*, vol. 96, no. 9, p. 094501, 2006.

[59] B. Hof, A. de Lozar, D. Kuik, and J. Westerweel, "Repeller or attractor? selecting the dynamical model for the onset of turbulence in pipe flow," *Phys. Rev. Lett.*, vol. 101, p. 214501, 2008.

[60] D. Borrero-Echeverry, R. Tagg, and M. Schatz, "Transient Turbulence in Taylor-Couette Flow." Arxiv preprint arXiv:0905.0147, 2009.

[61] B. Hof, J. Westerweel, T. Schneider, and B. Eckhardt, "Finite lifetime of turbulence in shear flows," *Nature*, vol. 443, pp. 59–62, 2006.

[62] H. Faisst and B. Eckhardt, "Sensitive dependence on initial conditions in transition to turbulence in pipe flow," *Journal of Fluid Mechanics*, vol. 504, pp. 343–352, 2004.

[63] A. Meseguer and L. Trefethen, "Linearized pipe flow to Reynolds number 107," *Journal of Computational Physics*, vol. 186, no. 1, pp. 178–197, 2003.

[64] J. Crutchfield and K. Kaneko, "Are attractors relevant to turbulence?," *Physical Review Letters*, vol. 60, pp. 2715–2718, 1988.

[65] H. Kantz and P. Grassberger, "Repellers, semi-attractors, and long-lived chaotic transients," *Physica D: Nonlinear Phenomena*, vol. 17, no. 1, 1985.

[66] T. Tél and Y. Lai, "Chaotic transients in spatially extended systems," *Physics Reports*, vol. 460, no. 6, pp. 245–275, 2008.

[67] R. Fisher and L. Tippett, "Limiting forms of the frequency distribution of the largest or smallest member of a sample," *Proc. Camb. Phil. Soc.*, vol. 24, no. 2, pp. 180–190, 1928.

[68] E. Gumbel, *Statistics of Extremes.* Columbia University Press, New York, NY, 1958.

[69] T. Schneider and B. Eckhardt, "Lifetime statistics in transitional pipe flow," *Physical Review E*, vol. 78, no. 4, p. 46310, 2008.

[70] A. Willis and R. Kerswell, "Turbulent dynamics of pipe flow captured in a reduced model: puff relaminarization and localized edge states," *Journal of Fluid Mechanics*, vol. 619, pp. 213–223, 2009.

[71] T. Tran, P. Chakraborty, G. Gioia, N. Guttenberg, N. Goldenfeld, A. Prescott, W. I. Goldburg, and H. Kellay, "Dynamics of falling soap films," 2009.

[72] M. Rutgers, X. Wu, and W. Daniel, "Conducting fluid dynamics experiments with vertically falling soap films," *Review of Scientific Instruments*, vol. 72, p. 3025, 2001.

[73] D. Georgiev and P. Vorobieff, "The slowest soap-film tunnel in the Southwest," *Review of Scientific Instruments*, vol. 73, p. 1177, 2002.

[74] R. Mittal and G. Iaccarino, "Immersed boundary methods," 2005.

[75] J. Monaghan, "Smoothed particle hydrodynamics," *Annual review of astronomy and astrophysics*, vol. 30, no. 1, pp. 543–574, 1992.

[76] X. He and L. Luo, "Lattice Boltzmann model for the incompressible Navier–Stokes equation," *Journal of Statistical Physics*, vol. 88, no. 3, pp. 927–944, 1997.

[77] S. Chen and G. Doolen, "Lattice Boltzmann method for fluid flows," *Annual Review of Fluid Mechanics*, vol. 30, no. 1, pp. 329–364, 1998.

[78] O. Zienkiewicz and R. Taylor, *The finite element method for solid and structural mechanics.* Butterworth-Heinemann, 2005.

[79] V. Ferreira, M. Tome, N. Mangiavacchi, A. Castelo, J. Cuminato, A. Fortuna, and S. McKee, "High-order upwinding and the hydraulic jump," *International Journal for Numerical Methods in Fluids*, vol. 39, no. 7, 2002.

[80] P. H. Gaskell and A. K. C. Lau, "Curvature-compensated convective transport: Smart, a new boundedness-preserving transport algorithm," *International Journal for Numerical Methods in Fluids*, vol. 8, pp. 617–641, 1988.

[81] J. Paret and P. Tabeling, "Intermittency in the two-dimensional inverse cascade of energy: Experimental observations," *Physics of Fluids*, vol. 10, p. 3126, 1998.

[82] N. Kotey, D. Bergstrom, and M. Tachie, "Power laws for rough wall turbulent boundary layers," *Physics of Fluids*, vol. 15, p. 1396, 2003.

[83] V. Patel, "Perspective: Flow at high Reynolds number and over rough surfaces: Achilles heel of CFD," *Journal of Fluids Engineering*, vol. 120, no. 3, pp. 434–444, 1998.

[84] N. Afzal, A. Seena, and A. Bushra, "Power Law Turbulent Velocity Profile in Transitional Rough Pipes," *Journal of Fluids Engineering*, vol. 128, p. 548, 2006.

[85] M. Davey and G. O'toole, "Microbial biofilms: from ecology to molecular genetics," *Microbiology and molecular biology reviews*, vol. 64, no. 4, pp. 847–867, 2000.

[86] G. West, J. Brown, and B. Enquist, “The fourth dimension of life: fractal geometry and allometric scaling of organisms,” *Science*, vol. 284, no. 5420, p. 1677, 1999.

[87] P. Stewart and J. William Costerton, “Antibiotic resistance of bacteria in biofilms,” *The Lancet*, vol. 358, no. 9276, pp. 135–138, 2001.

[88] W. Allee, A. Emerson, O. Park, T. Park, and K. Schmidt, “Principles of animal ecology,” *Philadelphia and London*, 1949.

[89] P. Stephens, W. Sutherland, and R. Freckleton, “What is the Allee effect?,” *Oikos*, vol. 87, no. 1, p. 185, 1999.

[90] R. Shapiro, “A replicator was not involved in the origin of life,” *IUBMB life*, vol. 49, no. 3, pp. 173–176, 2000.

[91] W. Gilbert, “The RNA world,” *Nature*, vol. 319, no. 6055, p. 618, 1986.

[92] H. Trinks, W. Schröder, and C. Biebricher, “Ice and the origin of life,” *Origins of Life and Evolution of Biospheres*, vol. 35, no. 5, pp. 429–445, 2005.

[93] T. R. Gregory, “Animal genome size database,” 2006. http://www.genomesize.com.

[94] J. Von Neumann, A. Burks, *et al.*, *Theory of self-reproducing automata.* University of Illinois Press Urbana, 1966.

[95] U. Pesavento, “An implementation of von Neumann’s self-reproducing machine,” *Artificial Life*, vol. 2, pp. 337–354, 1995.

[96] B. McMullin, “John von Neumann and the evolutionary growth of complexity: Looking backward, looking forward,” *Artificial Life*, vol. 6, no. 4, pp. 347–361, 2000.

[97] T. S. Ray, “Evolution, ecology, and optimization of digital organisms,” in *Scientific Excellence in Supercomputing: The IBM 1990 Contest Prize Papers* (K. R. Billingsley, E. Derohanes, and I. H Brown, eds.), pp. 489–531, The University of Georgia: The Baldwin Press, 1991.

[98] C. Adami and C. T. Brown, “Evolutionary learning in the 2d artificial life systems: Avida,” in *Proc. Artificial Life IV* (R. Brooks and P. Maes, eds.), (US), pp. 377–381, MIT Press, 1994.

[99] T. Ray, “Evolution, ecology and optimization of digital organisms,” *Santa Fe*, 1992.

[100] R. K. Standish, “Open-ended artificial evolution,” *International Journal of Computational Intelligence and Applications*, vol. 3, 2002.

[101] K. Vetsigian and N. Goldenfeld, “Global divergence of microbial genome sequences mediated by propagating fronts,” *Proc. Natl. Acad. Sci.*, vol. 102, pp. 7332–7337, 2005.

[102] C. R. Woese, “The emergence of genetic organization(Genetic organization emergence, considering pretranslational evolution in nontranslational protein synthesis, nucleic acid evolution and gene origin),” *Exobiology.(A 72-22001 08-04) Amsterdam, North-Holland Publishing Co., 1972,*, pp. 301–341, 1972.

[103] H. Echols and M. Goodman, “Fidelity mechanisms in DNA replication,” *Annual review of biochemistry*, vol. 60, no. 1, pp. 477–511, 1991.

[104] S. Kowalczykowski, D. Dixon, A. Eggleston, S. Lauder, and W. Rehrauer, “Biochemistry of homologous recombination in Escherichia coli.,” *Microbiology and Molecular Biology Reviews*, vol. 58, no. 3, pp. 401–465, 1994.

[105] M. Syvanen, “Horizontal gene transfer: evidence and possible consequences,” *Annual Review of Genetics*, vol. 28, no. 1, pp. 237–261, 1994.

[106] K. Vetsigian, C. Woese, and N. Goldenfeld, “Collective evolution and the genetic code,” *Proceedings of the National Academy of Sciences*, vol. 103, no. 28, pp. 10696–10701, 2006.

[107] R. Jain, M. Rivera, and J. Lake, “Horizontal gene transfer among genomes: the complexity hypothesis,” 1999.

[108] C. Adami, “Digital genetics: unravelling the genetic basis of evolution,” *Nature Reviews Genetics*, vol. 7, pp. 109–118, 2006.

[109] S. Elena and R. Lenski, “Evolution experiments with microorganisms: The dynamics and genetic bases of adaptation,” *Nature Reviews: Genetics*, vol. 4, no. 6, pp. 457–469, 2003.

[110] C. Wilke and C. Adami, “The biology of digital organisms,” *Trends in Ecology & Evolution*, vol. 17, no. 11, pp. 528–532, 2002.

[111] R. E. Lenski, C. Ofria, T. C. Collier, and C. Adami, “Genome complexity, robustness and genetic interactions in digital organisms,” *Nature*, vol. 400, pp. 661–664, 1999.

[112] C. Adami, C. T. Brown, and M. R. Haggerty, “Abundance-distributions in artificial life and stochastic models: “age and area” revisited,” in *Proceedings of the Third European Conference on Advances in Artificial Life*, (London, UK), pp. 503–514, Springer-Verlag, 1995.

[113] E. Ostrowski, C. Ofria, and R. Lenski, “Ecological specialization and adaptive decay in digital organisms.,” *Am. Nat*, vol. 169, pp. E1–E20, 2007.

[114] R. Lenski, C. Ofria, R. T. Pennock, and C. Adami, “The evolutionary origin of complex features,” *Nature*, vol. 423, pp. 139–144, 2003.

[115] S. Wright, "The roles of mutation, inbreeding, crossbreeding and selection in evolution," *Proceedings of the Sixth International Congress on Genetics*, vol. 1, no. 6, pp. 356–366, 1932.

[116] S. Gavrilets, *Fitness landscapes and the origin of species.* Princeton University Press, Princeton, US, 2004.

[117] H. Orr, "The genetic theory of adaptation: a brief history," *Nature Reviews Genetics*, vol. 6, no. 2, pp. 119–127, 2005.

[118] C. Adami, "Sequence complexity in darwinian evolution," *Complexity*, vol. 8, no. 2, pp. 49–56, 2002.

[119] G. Caldarelli, P. Higgs, and A. McKane, "Modelling Coevolution in Multispecies Communities," *Journal of Theoretical Biology*, vol. 193, no. 2, pp. 345–358, 1998.

[120] B. Drossel, P. G. Higgs, and A. J. McKane, "The influence of predator-prey population dynamics on the long-term evolution of food web structure," *J. Theor. Biol.*, vol. 208, pp. 91–107, 2001.

[121] A. McKane, "Evolving complex food webs," *The European Physical Journal B-Condensed Matter*, vol. 38, no. 2, pp. 287–295, 2004.

[122] M. Maron and C. T. Fernando, "Food webs and the evolution of organism complexity," in *Complexity Workshop, Artificial Life X*, 2006.

[123] M. Eigen, "Self-organization of matter andthe evolution of biological macromolecules," *Naturwissenschaften*, vol. 10, pp. 465–527, 1971.

[124] C. Wilke, "Quasispecies theory in the context of population genetics," *BMC Evol Biol*, vol. 5, p. 44, 2005.

[125] S. Wright, "Evolution in mendelian populations," *Genetics*, vol. 16, pp. 97–158, 1931.

[126] M. Kimura, "On the probability of fixation of mutant genes in a population," *Genetics*, vol. 47, no. 6, pp. 713–719, 1962.

[127] H. Orr, "The Rate of Adaptation in Asexuals," *Genetics*, vol. 155, no. 2, pp. 961–968, 2000.

[128] D. Stamper, M. Walch, and R. Jacobs, "Bacterial population changes in a membrane bioreactor for graywater treatment monitored by denaturing gradient gel electrophoretic analysis of 16S rRNA gene fragments," *Applied and environmental microbiology*, vol. 69, no. 2, pp. 852–860, 2003.

[129] A. Fernandez, S. Huang, S. Seston, J. Xing, R. Hickey, C. Criddle, and J. Tiedje, "How stable is stable? Function versus community composition," *Applied and Environmental Microbiology*, vol. 65, no. 8, pp. 3697–3704, 1999.

[130] B. Drossel, A. McKane, and C. Quince, “The impact of nonlinear functional responses on the long-term evolution of food web structure,” *Journal of Theoretical Biology*, vol. 229, no. 4, pp. 539–548, 2004.

[131] L. Valen, “A new evolutionary law,” *Evolutionary theory*, vol. 1, pp. 1–30, 1973.

[132] M. Ebner, A. Grigore, A. Heffner, and J. Albert, “Coevolution produces an arms race among virtual plants,” *Lecture notes in computer science*, pp. 316–325, 2002.

[133] M. Ebner, “Evolution and growth of virtual plants,” *Lecture notes in computer science*, pp. 228–237, 2003.

[134] G. Wagner and L. Altenberg, “Complex adaptations and the evolution of evolvability,” *Evolution*, vol. 50, no. 3, pp. 967–976, 1996.

[135] M. Kirschner and J. Gerhart, “Evolvability,” *Proceedings of the National Academy of Sciences*, vol. 95, no. 15, pp. 8420–8427, 1998.

[136] P. Larranaga, C. Kuijpers, R. Murga, I. Inza, and S. Dizdarevic, “Genetic algorithms for the travelling salesman problem: A review of representations and operators,” *Artificial Intelligence Review*, vol. 13, no. 2, pp. 129–170, 1999.

[137] G. Price, “Selection and Covariance,” *Nature*, vol. 227, no. 5257, pp. 520–521, 1970.

[138] S. Rice, *Evolutionary theory: mathematical and conceptual foundations.* Sinauer Associates, 2004.

[139] W. Hamilton, “The evolution of altruistic behavior,” *American naturalist*, pp. 354–356, 1963.

[140] M. van Baalen and D. Rand, “The unit of selection in viscous populations and the evolution of altruism,” *Journal of Theoretical Biology*, vol. 193, no. 4, pp. 631–648, 1998.

[141] D. Wilson, “A theory of group selection,” *Proceedings of the National Academy of Sciences*, vol. 72, no. 1, pp. 143–146, 1975.

[142] M. Nowak, “Five rules for the evolution of cooperation,” *Science*, vol. 314, no. 5805, pp. 1560–1563, 2006.

[143] P. Stragier and R. Losick, “Molecular genetics of sporulation in Bacillus subtilis,” *Annual review of genetics*, vol. 30, no. 1, pp. 297–341, 1996.

[144] E. Bonabeau, “Agent-based modeling: Methods and techniques for simulating human systems,” *Proceedings of the National Academy of Sciences*, vol. 99, no. 90003, pp. 7280–7287, 2002.

[145] M. Granovetter, “Threshold Models of Collective Behavior,” *American Journal of Sociology*, vol. 83, no. 6, p. 1420, 1978.

[146] C. McPhail and R. Wohlstein, “Individual and Collective Behaviors Within Gatherings, Demonstrations, and Riots,” *Annual Reviews in Sociology*, vol. 9, no. 1, pp. 579–600, 1983.

[147] D. Helbing, I. Farkas, P. Molnar, and T. Vicsek, “Simulation of pedestrian crowds in normal and evacuation situations,” *Pedestrian and Evacuation Dynamics*, pp. 21–58, 2002.

[148] G. Robins, P. Pattison, and P. Elliott, “Network models for social influence processes,” *Psychometrika*, vol. 66, no. 2, pp. 161–189, 2001.

[149] S. Huet and G. Deffuant, “Differential Equation Models Derived from an Individual-Based Model Can Help to Understand Emergent Effects,” *Journal of Artificial Societies and Social Simulation*, vol. 11, no. 2, p. 10, 2008.

[150] C. Nardini, B. Kozma, and A. Barrat, “Whoś Talking First? Consensus or Lack Thereof in Coevolving Opinion Formation Models,” *Physical Review Letters*, vol. 100, no. 15, p. 158701, 2008.

[151] W. Zachary, “An information flow model for conflict and fission in small groups,” *Journal of Anthropological Research*, vol. 33, no. 4, pp. 452–473, 1977.

[152] B. Skyrms and R. Pemantle, “A dynamic model of social network formation.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 16, pp. 9340–9346, 2000.

[153] E. M. Jin, M. Girvan, and M. E. J. Newman, “Structure of growing social networks,” *Phys. Rev. E*, vol. 64, p. 046132, Sep 2001.

[154] M. Newman, D. Watts, and S. Strogatz, “Random graph models of social networks,” *Proceedings of the National Academy of Sciences*, vol. 99, no. 90001, pp. 2566–2572, 2002.

[155] N. L. Geard and S. G. Bullock, “Group formation and social evolution: a computational model,” in *The Eleventh International Conference on the Simulation and Synthesis of Living Systems (Artificial Life XI)* (S. Bullock, J. Noble, R. A. Watson, and M. A. Bedau, eds.), (Cambridge, MA), pp. 197–203, MIT Press, 2008.

[156] J. Pacheco, A. Traulsen, and M. Nowak, “Coevolution of Strategy and Structure in Complex Networks with Dynamical Linking,” *Physical Review Letters*, vol. 97, no. 25, p. 258103, 2006.

[157] Y. Shoham and M. Tennenholtz, “On the emergence of social conventions: modeling, analysis, and simulations,” *Artificial Intelligence*, vol. 94, no. 1-2, pp. 139–166, 1997.

[158] G. Boella and L. van der Torre, “Organizations in Artificial Social Systems,” *Lecture Notes in Computer Science*, vol. 3913, p. 198, 2006.

[159] A. Kirou, B. Ruszczycki, M. Walser, and N. Johnson, “Computational Modeling of Collective Human Behavior: The Example of Financial Markets,” *Lecture Notes in Computer Science*, vol. 5101, pp. 33–41, 2008.

[160] J. Gandhi and A. Przeworski, “Cooperation, cooptation, and rebellion under dictatorships,” *Economics & Politics*, vol. 18, no. 1, pp. 1–26, 2006.

[161] T. Gross and B. Blasius, “Adaptive coevolutionary networks: a review,” *Journal of The Royal Society Interface*, vol. 5, no. 20, pp. 259–271, 2008.

[162] M. Zimmermann and V. Eguíluz, “Cooperation, social networks, and the emergence of leadership in a prisoners dilemma with adaptive local interactions,” *Physical Review E*, vol. 72, no. 5, p. 56118, 2005.

[163] V. Eguiluz, M. Zimmermann, C. Cela-Conde, and M. Miguel, “Cooperation and the Emergence of Role Differentiation in the Dynamics of Social Networks 1,” *American Journal of Sociology*, vol. 110, no. 4, pp. 977–1008, 2005.

[164] F. Schweitzer, J. Zimmermann, and H. Mühlenbein, “Coordination of decisions in a spatial agent model,” *Physica A: Statistical Mechanics and its Applications*, vol. 303, no. 1-2, pp. 189–216, 2002.

[165] I. Couzin, J. Krause, N. Franks, and S. Levin, “Effective leadership and decision-making in animal groups on the move.,” *Nature*, vol. 433, no. 7025, pp. 513–6, 2005.

[166] I. Trofimova and N. Mitin, “Self-Organization and Resource Exchange in EVS Modeling,” *Nonlinear Dynamics, Psychology, and Life Sciences*, vol. 6, no. 4, pp. 351–362, 2002.

[167] H. Evans, “The behavior patterns of solitary wasps,” *Annual Review of Entomology*, vol. 11, no. 1, pp. 123–154, 1966.

[168] M. Richter, “Social wasp (Hymenoptera: Vespidae) foraging behavior,” *Annual Review of Entomology*, vol. 45, no. 1, pp. 121–150, 2000.

[169] G. Robinson, “Regulation of division of labor in insect societies,” *Annual Review of Entomology*, vol. 37, no. 1, pp. 637–665, 1992.

[170] T. Seeley and S. Buhrman, “Group decision making in swarms of honey bees,” *Behavioral Ecology and Sociobiology*, vol. 45, no. 1, pp. 19–31, 1999.

[171] I. Hecht, E. Ben-Jacob, and H. Levine, “Correlated phenotypic transitions to competence in bacterial colonies,” *Nature (London) Phys Rev E*, vol. 76, p. 040901, 2006.

[172] H. Ochman, J. Lawrence, E. Groisman, *et al.*, "Lateral gene transfer and the nature of bacterial innovation," *Nature*, vol. 405, no. 6784, pp. 299–304, 2000.

[173] A. Babic, A. B. Lindner, M. Vulic, E. J. Stewart, and M. Radman, "Direct Visualization of Horizontal Gene Transfer," *Science*, vol. 319, no. 5869, pp. 1533–1536, 2008.

[174] J. Hotopp, M. Clark, D. Oliveira, J. Foster, P. Fischer, M. Torres, J. Giebel, N. Kumar, N. Ishmael, S. Wang, *et al.*, "Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes," *Science*, vol. 317, no. 5845, pp. 1753–1756, 2007.

[175] E. Gladyshev, M. Meselson, and I. Arkhipova, "Massive horizontal gene transfer in bdelloid rotifers.," *Science*, vol. 320, no. 5880, pp. 1210–3, 2008.

[176] P. Keeling and J. Palmer, "Horizontal gene transfer in eukaryotic evolution," *Nature Reviews Genetics*, vol. 9, no. 8, pp. 605–618, 2008.

[177] J. Pace, C. Gilbert, M. Clark, and C. Feschotte, "Repeated horizontal transfer of a DNA transposon in mammals and other tetrapods," *Proceedings of the National Academy of Sciences*, vol. 105, no. 44, pp. 17023–8, 2008.

Author's Biography

Nicholas Guttenberg was born on December 17, 1982 in Boulder, Colorado, though his family moved to Maryland soon after. During highschool he worked at the National Institute of Standards in Technology. He graduated from St. Andrews high school in 2000, after which he enrolled at George Washington University in Washington, DC. One year later he transferred to McGill University in Montreal, QC. At McGill he participated in the Canadian Association of Physicists University Prize Exam twice, earning 2nd place in 2003 and 3rd place in 2004. He also received a Sigma Xi Excellence in Undergraduate Research award in 2004 for work with Martin Grant and Hong Guo. He received a Bachelors of Science in Physics from McGill University in 2004. In August 2004, he started graduate studies at the University of Illinois at Urbana-Champaign and soon after joined the group of Nigel Goldenfeld. He received the Illinois Distinguished Fellowship from 2004 to 2007, the L. S. Edelheit Biological Physics fellowship in 2006, and the Drickamer Research Fellowship in 2009.