STOCHASTIC MANY-BODY PROBLEMS IN ECOLOGY, EVOLUTION, NEUROSCIENCE, AND SYSTEMS BIOLOGY

BY

THOMAS C. BUTLER

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, 2010

Urbana, Illinois

Doctoral Committee:

Professor Karin Dahmen, Chair
Professor Nigel Goldenfeld, Director of Research
Professor Jon Thaler
Assistant Professor Yann Chemla

Abstract

Using the tools of many-body theory, I analyze problems in four different areas of biology dominated by strong fluctuations: The evolutionary history of the genetic code, spatiotemporal pattern formation in ecology, spatiotemporal pattern formation in neuroscience and the robustness of a model circadian rhythm circuit in systems biology.

In the first two research chapters, I demonstrate that the genetic code is extremely optimal (in the sense that it manages the effects of point mutations or mistranslations efficiently), more than an order of magnitude beyond what was previously thought. I further show that the structure of the genetic code implies that early proteins were probably only loosely defined. Both the nature of early proteins and the extreme optimality of the genetic code are interpreted in light of recent theory [1] as evidence that the evolution of the genetic code was driven by evolutionary dynamics that were dominated by horizontal gene transfer. I then explore the optimality of a proposed precursor to the genetic code. The results show that the precursor code has only limited optimality, which is interpreted as evidence that the precursor emerged prior to translation, or else never existed.

In the next part of the dissertation, I introduce a many-body formalism for reaction-diffusion systems described at the mesoscopic scale with master equations. I first apply this formalism to spatially-extended predator-prey ecosystems, resulting in the prediction that many-body correlations and fluctuations drive population cycles in time, called quasi-cycles. Most of these results were previously known, but were derived using the system size expansion [2, 3]. I next apply the analytical techniques developed in the study of quasi-cycles to a simple model of Turing patterns in a predator-prey ecosystem. This analysis shows that

fluctuations drive the formation of a new kind of spatiotemporal pattern formation that I name "quasi-patterns." These quasi-patterns exist over a much larger range of physically accessible parameters than the patterns predicted in mean field theory and therefore account for the apparent observations in ecology of patterns in regimes where Turing patterns do not occur. I further show that quasi-patterns have statistical properties that allow them to be distinguished empirically from mean field Turing patterns.

I next analyze a model of visual cortex in the brain that has striking similarities to the activator-inhibitor model of ecosystem quasi-pattern formation. Through analysis of the resulting phase diagram, I show that the architecture of the neural network in the visual cortex is configured to make the visual cortex robust to unwanted internally generated spatial structure that interferes with normal visual function. I also predict that some geometric visual hallucinations are quasi-patterns and that the visual cortex supports a new phase of spatially scale invariant behavior present far from criticality.

In the final chapter, I explore the effects of fluctuations on cycles in systems biology, specifically the pervasive phenomenon of circadian rhythms. By exploring the behavior of a generic stochastic model of circadian rhythms, I show that the circadian rhythm circuit exploits leaky mRNA production to safeguard the cycle from failure. I also show that this safeguard mechanism is highly robust to changes in the rate of leaky mRNA production. Finally, I explore the failure of the deterministic model in two different contexts, one where the deterministic model predicts cycles where they do not exist, and another context in which cycles are not predicted by the deterministic model.

To Nathanael Butler, my brother and best friend.

Acknowledgments

I clearly remember my first conversation with my advisor, Nigel Goldenfeld. He invited me to his office to interview for a position in his research group. After more than two hours I finally left, not sure I'd followed the whole conversation, but convinced that whatever Nigel worked on, the intellectual energy he brought with him would lead to discovery. When he eventually offered me a position in his research group, I wasn't sure about my research direction. On the other hand, I was certain that if I spent time around him something really interesting would happen, so I accepted anyway. That turned out to be exactly right, and accepting the position is the best professional decision I have yet made. The four years I have spent working with him have been the greatest years of intellectual growth in memory. I hope he will be a collaborator for the rest of our respective careers. I am certain he'll be a friend.

The students and postdocs of the Goldenfeld group have made graduate school much more fun and informative. John Veysey, Patrick Chan, Nicholas Gutenberg, Patricio Jeraldo, Maksim Sipos, Zhenyu Wang, Nicholas Chia, David Reynolds, Luiza Angheluta, Andreas Menzel all have provided me with sounding boards for often crazy ideas, technical help on my research, and friendship. They create a great research environment that I am confident I'll miss no matter where I go in the future. I also am grateful for the fun of doing science with Damien Mathew, and Professors Jack Cowan, Karin Dahmen, Lee DeVille, Damien Mathew and Zaida Luthey-Schulten.

I never would have seriously considered a scientific career without the encouragement and support of Professor Bill Evenson, who advised my undergraduate research at Brigham

Young University. When I ran out of tuition money, he found me grants and scholarships. When I ran out of confidence, he encouraged me. I still don't know what possessed Bill to spend so much time at the end of a distinguished career at BYU on making the scientific launch of a single undergraduate student possible, but he deserves substantial credit in everything I do as a scientist.

Don and Dianne Shepard have fed me countless times, and have reminded me that even if my biological parents live far away, I have family right here in town as well. Thanks especially to them for housing me as I complete my PhD. Alex Valencic and Noah Ruggieri shared meals, movies, and life as I have moved through graduate school, and they have started their careers. Christina Allan and the entire Allan family from oldest (?) to youngest (10) are all convinced that every person they meet is immediate family and should be welcomed into their home, fed, and loved unconditionally. I am lucky to be a beneficiary of that. Brady and Erin Gibson and their baby, Daisy Gibson, are family even if the genealogical charts don't show it. Hayley Smith is too. Ben Hertzberg, Brittany Schilling, Matt Thorum, Beth Cummings, Andrew Mills, Nick Prince, and many others deserve special thanks for support and friendship.

My mother and father have supported me in studying far into what may have otherwise been seen as the time to start a career. They are full of support and love. My siblings Tim, Connie, Nathan, and Matt are my closest friends. No one in my family is a scientist, but each of them has kindly listened as I explained my research to them and has provided support, material and moral, beyond what one can reasonably expect even from family. While I am hesitant to single out any member of my family, I do need to thank my father for raising me in almost continual contact with wilderness. The spectacular wild places my father introduced me to and still visits with me today drive the sense of wonder behind the best things in my science. I also need to thank my mother for being Mom. No one deserves more thanks than her.

The work in this dissertation was supported by the National Science Foundation under

grant No. NSF-EF-0526747 and by the Drickhamer Fellowship of the University of Illinois.

Table of Contents

List of Tables . . . . . **xi**

List of Figures . . . . . **xii**

Chapter 1 Introduction **1**
1.1 Evolution of the genetic code 4
1.2 Fluctuations and the emergence of quasi-cycles in predator-prey systems . . 5
1.3 Fluctuations and the emergence of quasi-patterns in Turing systems 6
1.4 Quasi-patterns and the visual cortex 6
1.5 Leaky mRNA production and circadian rhythms 7
1.6 My contributions 7

Chapter 2 Extreme Optimality of the Genetic Code **9**
2.1 The evolutionary dynamics of early life and the origin of the genetic code . . 9
2.2 Evolution of the genetic code 10
2.3 Basic properties of the genetic code 13
2.4 Molecular Dynamics of the Polar Requirement 15
2.5 Optimality analysis of the canonical genetic code 16
2.6 Evidence for statistical proteins in early life 21
2.7 Optimality analysis of alternative codes and measures 23

Chapter 3 Optimality of a Proposed Precursor to the Genetic Code . . **25**
3.1 Precursor code proposals 25
3.2 The Copley Smith Morowitz precursor code 26

Chapter 4 Field Theory of Fluctuating Predator-Prey Populations in Space 33
4.1 Ecology as a many-body problem 33
4.2 Failure of traditional predator-prey theory 34
4.2.1 Mathematical models of predator-prey interactions 36
4.3 Individual level modeling of predator-prey population dynamics 38
4.4 Master equation representation of individual level population dynamics . . . 43
4.5 Master equation for predator-prey interactions 45
4.5.1 Mapping to path integral formulation 47
4.6 Derivation of quasi-cycles from Large V expansion 51

Chapter 5 The statistical mechanics of Turing patterns: The Levin-Segel model . . . 56
5.1 The Turing mechanism . . . 58
5.2 Heuristic analysis of the Levin-Segel model . . . 60
5.3 Predator-prey model with intrinsic noise . . . 62
5.3.1 Field theory representation of the model . . . 63
5.3.2 System size expansion . . . 67
5.3.3 The power spectrum . . . 69
5.4 Analysis of the power spectrum . . . 72
5.4.1 Phase diagram for quasi-patterns . . . 72
5.4.2 Wavelength of fluctuation driven patterns . . . 74
5.4.3 Period of quasi-cycles . . . 75
5.5 Distinguishing quasi-patterns and quasi-cycles from other spatiotemporal patterns . . . 76
5.5.1 Distinguishing quasi-cycles from limit cycles . . . 77
5.5.2 Distinguishing quasi-patterns from mean field patterns . . . 78
5.6 Thermodynamic limit . . . 79
5.7 Validity of the large V expansion and the scale of quasi-patterns . . . 80
5.8 Explaining the failure of mean field theory . . . 81
5.9 Application to field data and experiments . . . 82
5.10 Conclusions and prospects for future research . . . 84

Chapter 6 Emergent spatiotemporal pattern formation in the neocortex 86
6.1 Modeling the visual and neocortex . . . 91
6.1.1 General anatomical features of the neocortex . . . 91
6.1.2 A simplified canonical microcircuit for cortical modules . . . 92
6.1.3 The statistical dynamics of a single microcircuit . . . 93
6.1.4 Neocortex as a two dimensional lattice . . . 95
6.1.5 A Lagrangian for neocortical networks . . . 98
6.1.6 System-size expansion of the Lagrangian . . . 99
6.2 Variational Derivatives of the Lagrangians . . . 101
6.3 Mean-field Wilson-Cowan equations . . . 102
6.4 Langevin equations . . . 103
6.5 The emergence of quasi-cycles . . . 107
6.6 Mean field conditions for pattern formation . . . 108
6.7 Pattern formation beyond mean field . . . 112
6.8 Pattern formation with inhibition forbidden in the intermediate length scale 114
6.9 Scale-free spatial effects . . . 115
6.10 Phase diagram when inhibition is introduced at the intermediate length scale 116
6.11 Conclusion . . . 118

Chapter 7 Sloppiness and robustness in a model circadian rhythm circuit 120
7.1 The problem of biological robustness . . . 120
7.2 Sloppy models . . . 122
7.3 A simple model of circadian rhythms . . . 123

7.4 Robustness in the model circadian rhythm 126
7.5 Comparison of stochastic and deterministic approaches 129
7.6 Conclusions 131

Chapter 8 Conclusions 132
8.1 Final thoughts on the role of theory in biology 132

Appendix A MSR response function formalism 135

Appendix B Continuum limit for reaction-diffusion master equations . . 138

Appendix C Some explicit calculations for the Levin-Segel model 140
C.1 Predator competition term in second quantized Hamiltonian 140
C.2 System size expansions 141
C.3 Mean field analysis 142

References 145

Vita 156

List of Tables

2.1 The canonical genetic code. The first base in the codon is given by the row, the second base is given by the column, and the third base is given by the subsections within the rows. For example, UUA maps to Leu. 14
2.2 Computational and experimental polar requirements 17
2.3 The matrix $W_{c,c'}$ of transition/transversion biases taken from [4]. 19
2.4 P_b for several naturally occurring variant codes 23

3.1 Proposed precursor code from Ref.[5]. Row is first base, column is second base. 27

7.1 Reactants for circadian rhythm model from [6] 125

List of Figures

2.1 Predictions of the VWG model for code evolution. The vertical axis indicates the level of optimality a given code has, with lower values corresponding to higher optimality. The horizontal axis is time. Red trajectories correspond to evolution without horizontal gene transfer, and blue trajectories describe code evolution with horizontal gene transfer. The inset describes how much the evolved codes differ from each other, and the histogram on the right hand side describes the probability distribution of optimality for random codes. Figure taken from [1] 12

2.2 Scatter plot showing the relationship between radial distribution function (RDF) peak slope and experimental polar requirement for all amino acids. The straight line is a guide to the eye [7]. 18

2.3 Histogram of codes with different average mutational impacts generated through the monte carlo algorithm described above. The extreme optimality of the code can be seen clearly. 21

2.4 P_b as a function of the exponent q in the amino acid metric. 22

3.1 Histogram of average impact I per point mutation for randomly-generated codes with the same degeneracy structure as the canonical genetic code. . . . 30

3.2 Histogram of average impact I per point mutation for randomly-generated codes with the same degeneracy structure as the proposed precursor. There is more noise relative to the canonical code case due to the smaller ensemble of random codes required to calculate P_b for the precursor. 31

4.1 The mechanism of predator prey oscillations. Predators are red dots, and prey are black dots. Step one is in the upper left. 35

4.2 Comparison of mean field theory (red line) and the individual level model. Note the persistent noisy cycles in the individual level model and the stable spiral dynamics of the mean field theory. 39

5.1 Illustration of the steps of the Turing mechanism as described in the text. The figure should be viewed from top to bottom. The prey (activators) are represented by black dots, and the predators (inhibitors) are represented by red dots. 59

5.2 Phase diagram over stable parameter region in p/d. Region I is MFT level pattern formation, the shaded region II is the portion of the phase diagram where fluctuation driven quasi-patterns are present, region III is a spatially homogeneous phase. 73
5.3 Power spectrum with p=1, ν/μ=15 74
5.4 The phase diagram, with shading removed, is on the left. Region I is mean field pattern formation, region II is fluctuation driven pattern formation, and region III is homogeneous. The red arrow has its tail on the approximate location in parameter space simulated to produce the spatial patterns shown on the right. The image on the right hand is a heat map of population density in two dimensions. Note that the number of organisms is highly variable, even though mean field predicts a homogeneous state. The fluctuation effects are very large and variable, with patch populations ranging from 1200 to 0. The axes are the lattice index from simulation. 81
5.5 Sample trajectories of the Markov process for predator-prey dynamics. Note that while each is roughly oscillatory, a mean field theory derived from the average of many such trajectories would not contain oscillations. 83

6.1 Two examples of form constants in visual hallucinations. Figure taken from [8]. 88
6.2 Shamanic cave art from the Texas-Mexico border. The shaman, on the left, is diving into to a spirit world. Note the spiral images in the upper right, probably interpreted as gateways to the spirit world. The spiral image is also a reasonable representation of typical spiral hallucinations seen at the onset of the effects of hallucinogenic drugs. Figure taken from [9]. 89
6.3 A contemporary piece of Acoma pottery from the American Southwest. Figure taken from an online display of southwestern native art at `http://faculty.vassar.edu/lucic/formsofexchange/pottery.html`. 90
6.4 Block diagram of a reduced neocortical microcircuit. Py =pyramidal neurons, St = spiny stellate neurons. Unfilled triangles = excitatory synapses, Filled triangles = inhibitory synapses. 92
6.5 Block diagram of intra-modular coupling. 96
6.6 Block diagram of inter-modular coupling. 97
6.7 Spatial heterogeneity of neuronal excitation from the Wilson-Cowan equations. 111
6.8 Mean field patterns of neural activity in visual field coordinates. Note the close resemblance to reported geometric visual hallucinations. 112
6.9 The power spectrum as a function of k and ω in the quasi-pattern regime, without quasi-cycles. The scaled parameters are $g^1_{IE} = 3.0$ and $g^1_I = 0.2$. . . . 114
6.10 The real part of the largest eigenvalue of the stability matrix as a function of k in phases V and I respectively. 117

6.11 Phase diagram for pattern formation. This figure shows the conditions for the emergence of spatial structure. The x axis is g_I^1/g_{EE}^1 and the y axis is g_{IE}^1/g_{EE}^1. Region I is scale invariant structure at the mean field level, region II is a mean field pattern formation phase, region III is a homogeneous steady state phase, region IV is the quasi-pattern phase and region V is a mean field pattern forming phase. 118

7.1 24 hour circadian cycles in A and R from the reactions in Eq. 7.2. Figure taken from [6]. 126

7.2 Change in the period of the circadian rhythm as a function of α_A, or unpromoted mRNA production for the gene for A. The y axis is the normalized period, meaning that all periods have been divided by 24 hours. Note that the period stays roughly constant over three orders of magnitude in the stochastic description, and over at least two more orders of magnitude in the deterministic description. Not shown is the failure of the deterministic description at $\alpha_A = 0$, where the period diverges to infinity. 128

7.3 Period of stochastic and deterministic trajectories near a bifurcation to a stable steady state in the deterministic system. Note that the stochastic version continues to have oscillations over a much larger range of the bifurcation parameter. Note also that the period is much too short (around 5, versus the needed 24) near the bifurcation point. 130

Chapter 1

Introduction

This dissertation contains my contributions to several problems at the interface of statistical physics and biology. It can be broadly divided into two main parts. The first part, consisting of chapters two and three, is a summary of my research on the origins of the universal genetic code. The second part, which is the core of the thesis, applies ideas from non-equilibrium statistical mechanics to explain emergent spatiotemporal patterns in ecosystems and visual hallucinations, elucidate the architecture of the neural network in the visual cortex, and probe the mechanisms that maintain the robustness of circadian rhythms in the cells of many species.

The daunting (to me) breadth of topics discussed in this thesis evaporates when the names of the parts of the systems are ignored in favor of the fundamental feedbacks acting on the systems. Biochemical circadian clocks share with the population biology of predator-prey systems oscillatory behavior driven primarily by negative feedback [6, 10, 2, 11]. Spatial patterns that form in the predator-prey Tussock moth-wasp system and the geometric visual hallucinations generated in the visual cortex are both generated by the interplay of positive feedback coupled to slow diffusion and negative feedback coupled to fast diffusion [12, 13, 8]. Detached in this way from biological context, the summary of part two of my dissertation above is narrow and dry: "The second part, which is the core of the thesis, uses the master equation to study interactions that generate simple negative feedback. I then extend this work by adding space and interactions that generate simple positive feedback."

Far from showing that my thesis is trivial, the second version of my thesis summary is a biological application of a deep result in statistical physics that shows that the long

wavelength physics of a system often only depends on the "microscopic" laws governing the system through the most basic features of the system and its interactions such as range, symmetry, sign, dimension, and coupling to different fields, such as external magnetic fields [14]. The other details of the microscopic physics can simply be absorbed into the phenomenological parameters of the theory. While the enormous simplifications afforded by this claim should not be taken too far — it is still a very challenging project to work out the relationship between the parts of a biological system and the mechanisms that drive the collective behavior of that system at long wavelengths — the success of the extremely simplified models of biological phenomena that I present in this dissertation relies on the claim that the long wavelength emergent behavior only depends on the basic features of the system.

The detail independence of long wavelength emergent behavior in complex systems is vividly illustrated by the example of fluid flow. To quantitatively describe the macroscopic flow of water it is not necessary to know that water is at the microscopic level a collection of molecules consisting of two hydrogens and an oxygen atom. Such details are irrelevant to the macroscopic flow. That this is the case is clear from the fact that the equations that describe fluid flow, the Navier-Stokes equations, are *quantitatively* applicable to most fluids with only the modification of phenomenological parameters, such as the viscosity and density. This includes Newtonian fluids with much more complex microscopic physics, such as oils. The emergence of the collective macroscopic state of "fluid" relegates the often extremely complicated microscopic physics of the system to the parameters of the equations describing the macroscopic physics.

As illustrated by the example of fluids above, understanding the quantitative applicability of comparatively simple equations to a variety of complex systems is intimately tied to the concept of emergence in statistical physics. Emergence describes how the interactions between the components of systems with many degrees of freedom can lead to new states of the system with properties that are not subsumed in the properties of the parts. While

in principle the flow of liquid water could be described by a detailed molecule by molecule simulation, the liquid state has its own laws that, provided the microscopic physics leads to an emergent liquid state, are otherwise independent of the microscopic constituents of the fluid except through the values of viscosity and density. The remarkable reduction of complex microscopic detail into a handful of phenomenological parameters can be understood in principle by renormalization group ideas, which show how the parameters governing the long wavelength collective degrees of freedom absorb the details of the microscopic physics into their values (for the best introduction, see [14]).

After the reassurance that long length and timescale predictions can be made that don't depend on the details of the microscale physics, the most important implication of emergence and renormalization group for our purposes is that if the key features that govern the behavior of the emergent state can be identified heuristically, a theory can be constructed for the emergent phase without recourse to a detailed description of the parts of the system being modeled. Such a model is called a "minimal model," and the primary occupation of Goldenfeld group members is the construction, identification, and analysis of such models for systems ranging from superconductors and turbulent fluids to my own work on neuroscience and ecology.

In biology the power of minimal models to capture long wavelength phenomena is essential. This is because biology is incredibly diverse and particular. Writing down a theory of the brain that incorporates even a significant fraction of all of the details of neurons and synapses is virtually impossible, and would be so incomprehensible as to render it useless for insight even if such a model could be made. The fact that coarse grained, effective theories of biological systems free of overwhelming detail can make sufficiently sharp predictions to compare with experiment is what allows meaningful statistical physics theory in biology to progress, including the work in this thesis.

To close this portion of the introduction, I think it is essential to note that biology is different from physics in some very important ways. At risk of over-generalizing, for biolo-

gists, unlike physicists, it is often the underlying parts that make up a system, rather than the emergent states, that make a particular system of interest. Physicists who study high temperature superconductivity never say "I am really only interested in high T superconductivity because there is copper and oxygen in most of the compounds that have a high T superconducting state. If they were made of say, nickel, I wouldn't study them." But an ecologist is likely to say that her primary motivation for studying pattern formation is that it happens in the stripe forests she is attempting to preserve and study in the Colorado highlands [15].

1.1 Evolution of the genetic code

In the second chapter, we focus on the properties of the canonical genetic code. The primary question we address is best illustrated by a rough analogy: If the genome of an organism is a book, then the genetic code is the language the book is written in. How effective is the language the genome (book) is written in? In some conceivable languages, a printing mistake or misspelling will cause a major disruption of the intended message. In others, the intended message is hardly affected. Since the genetic code could have taken at least 20! other forms, some insight can be gained into its evolution by studying how resistant it is to misspellings. To address this question, we used molecular dynamics simulations (carried out by Damien Mathew and Zaida Luthey-Schulten) of the properties of amino acids to develop a metric for the severity of point mutation errors (misspellings) and compared the canonical genetic code to 100 million randomly generated genetic codes. We found that the canonical genetic code is far better than all but a handful (around 20) of the randomly generated codes at resisting the deleterious effects of point mutations, an order of magnitude better than had been thought previously [4].

These results have a natural interpretation. Recent theory [1] shows that such remarkable optimality is only attainable if the evolutionary dynamics of early life had an enormous

amount of horizontal gene transfer (the transfer of genetic material between two organisms without one being a descendant of the other).

Through further analysis of the genetic code, we also were able to show evidence that early proteins were only loosely defined. This allowed the code to evolve much more easily because genetic information, including genes for the code itself, could be transferred between organisms with slightly differing genetic codes.

In the third chapter, we apply the same analysis we applied in the second chapter to the canonical genetic code to a proposed precursor to the genetic code. We find that the proposed precursor is not significantly optimal, and thus either did not exist, or evolved prior to the emergence of translation.

1.2 Fluctuations and the emergence of quasi-cycles in predator-prey systems

In the fourth chapter, we show that predator-prey abundance cycles, which are extremely common in real ecosystems, are primarily driven by fluctuations due to the discrete nature of stochastic interactions of individual organisms with each other. These fluctuations are called intrinsic noise (or demographic noise in ecology) and are best captured analytically through the master equation formalism of non-equilibrium statistical mechanics. Though these results were previously known [2, 3], our formulation used field theory to substantially simplify the calculations, paving the way for extensions to more complex nonlinear spatially extended stochastic systems. The following two chapters carry out such extensions to discover qualitatively new phenomena in ecology and neuroscience.

1.3 Fluctuations and the emergence of quasi-patterns in Turing systems

In the fifth chapter, we resolve an outstanding problem in the theory of pattern formation. While many pattern forming systems seem to be described qualitatively by a mechanism for pattern formation called the Turing mechanism [16], the mathematical descriptions of those systems nearly always requires fine tuning of the system parameters, or an asymptotically large diffusion constant for one species to predict pattern formation. Such conditions are rarely physical. We show, using a simple model of ecological pattern formation, that the requirement for unphysical parameters vanishes when intrinsic noise is included.

This is surprising because noise is not usually seen as supporting the formation of highly structured states such as spatial patterns. To explain this strange result, we outline a complete theory of the new pattern forming states supported by intrinsic noise and name them "quasi-patterns." We also outline several possible ways to detect quasi-patterns experimentally. Quasi-patterns are a new class of spatial patterns that had not previously been known.

1.4 Quasi-patterns and the visual cortex

In the sixth chapter, we apply the insights of the previous chapter to an entirely different system: the visual cortex. Unlike ecosystems, however, pattern formation on the visual cortex is highly deleterious, corresponding to geometric hallucinations like those seen by shamans and LSD users. We show that the highly unusual network architecture of the visual cortex suppresses the emergence of spatial patterns under most circumstances. We also point to the possibility that under some circumstances, this same odd network structure might lead to avalanches of neuronal activity [17].

1.5 Leaky mRNA production and circadian rhythms

In the seventh chapter, we investigate the role of intrinsic fluctuations on a well known model from systems biology of the key feedbacks in the circadian rhythm circuit. Using stochastic algorithms to explore the behavior of the system, we show that the robustness of the circadian rhythm cycle is strengthened by an unlikely source: the leaky production of mRNA by an unpromoted gene. This finding supports the idea that certain aspects of biological robustness are emergent rather than evolved.

1.6 My contributions

In the work on the genetic code, I computed the extreme optimality of the genetic code using the results of Mathews' and Schulten's molecular dynamics simulations of amino acid properties. I also wrote the Monte Carlo code and carried out the optimality analysis. I developed in collaboration with Nigel a bootstrap error analysis that we used to assess the significance of the results. I also carried out the calculations that provided evidence for the statistical nature of early proteins. Damien Mathew aided in the writing of the section on the molecular dynamics results. In the analysis of the precursor code I wrote the code, carried out the optimality analysis and helped develop the interpretation of the results.

In chapter four on predator-prey oscillations, I carried out the calculations in collaboration with David Reynolds. The final structure of the calculations is mine. In the work in the following chapter on quasi-patterns and Turing instabilities, I suggested the model and carried out the calculation that led to the discovery and characterization of the quasi-pattern state. In the work on the visual cortex, Professor Jack Cowan carried out the mapping of the model to a field theory and I derived the conditions for pattern formation as well as the phase diagram. I also carried out the numerical simulations. The final form of the detailed calculational portions of the chapter on the visual cortex contains many contributions from Jack Cowan.

In the work on circadian rhythms, I suggested the calculation, carried out the simulations and helped develop the interpretation of the results.

Chapter 2

Extreme Optimality of the Genetic Code

2.1 The evolutionary dynamics of early life and the origin of the genetic code

Among the deepest problems in biology is the nature and dynamics of the earliest evolution. For reasons to be discussed more fully below, the most direct way to probe these deep evolutionary questions is to explore life's universal evolved features. The purpose of this chapter and the next is to apply simple tools from condensed matter physics to extract from biological data (the genetic code table and the physics of the amino acids) evidence regarding the nature of the earliest evolution. The result of these investigations is to provide rare empirical evidence for an emerging picture of the nature of the earliest life and its accompanying ecological and evolutionary dynamics

One of the most effective techniques we have for probing evolutionary dynamics is to compare the genome sequences of different organisms. Shared sequences correspond in general to shared ancestry, allowing, in principle, the reconstruction of the major evolutionary divergences in the history of life. Such a reconstruction has been carried out in rough outline, resulting in the discovery that all known life can be divided into three distinct domains, Archaea, Bacteria, and Eukarya [18], that emerged very early in life's history from a common ancestral community [19]. The features of life, including sequences, which are universal, are likely retained from the last common ancestral community. This justifies the claim made above that investigating the origin and evolution of life's universal features is equivalent

to exploring the nature and evolutionary dynamics of early life. In recent years, a picture of some of the basic features of this common ancestral community has begun to emerge [19, 1, 20, 21]. The emerging picture is of a diverse community of organisms evolving rapidly by sharing genes (and innovations encoded in them) not only with their descendants, but with other contemporaneous organisms. This is called horizontal gene transfer. This notion is supported by the observation that horizontal gene transfer is also extremely common in modern life with documented examples ranging from the recent horizontal gene transfer of genes for antibiotic resistance in drug resistant staph infections[22] to the transfer of genes among much larger organisms and across even different domains of life [23, 24]. As is intuitively expected from the additional channels of evolution it provides, horizontal gene transfer allows evolution to proceed at much more rapid rates than strictly vertical transmission of genes from organisms to their descendants [25].

One of the fundamental ingredients for the evolutionary dynamics described above is a common language for the expression of genetic information, that is, an essentially universal genetic code. Successful horizontal gene transfer of functioning genes requires that the genes be read in nearly the same way by both the donor organism and the receiving organisms. This points to at least two ways that the genetic code can shed light on early life. First, the dynamics that gave rise to the evolution of the code are the same dynamics that dominated evolution during early life. Second, because of the need for a nearly universal code in the horizontal gene transfer dominated dynamics of early life, understanding the emergence of that universality is likely to greatly expand our knowledge of early evolution.

2.2 Evolution of the genetic code

One very promising theory of the evolution of the genetic code was put forth recently in a paper by Vetsigian, Woese and Goldenfeld [1] (referred to hereafter as the VWG theory). In this paper, they claim that two key observed features of the genetic code, its universality

and its extraordinary ability to compensate for errors in translation [7, 4, 26] (discussed in detail below) can be explained only if two conditions were met in the universal ancestral community:

1. Proteins were only statistical, meaning that their function does not depend on a precise sequence of their amino acid building blocks, but that they function so long as roughly the right amino acids are used. This is also called ambiguity tolerance and has been conjectured to be a feature of early proteins [20, 27]. It should also be noted that this is very different than the observation that modern proteins are robust to small changes, as early proteins were presumably much shorter.
2. There was a significant amount of horizontal gene transfer.

Condition one is needed because it allows communities with slightly different genetic codes to exchange functional genes, including genes associated with the genetic code and translation. Without this requirement, different communities with different codes would be isolated from each other. Condition two, when combined with condition one, prevents the system from getting stuck in a local error tolerance minimum. If both of these conditions are met, then theory indicates that the genetic code will evolve to be both universal and extremely error tolerant [1]. Without both of these conditions satisfied, theory indicates that different variants of the genetic code will coexist with varying, but non-optimal levels of error tolerance (See fig. 2.1).

While the conditions and predictions of the VWG theory are intuitive and consistent with data, to draw stronger conclusions, more stringent tests should be devised. The conditions and predictions of the theory suggest at least three possible lines of inquiry that could lend support to the VWG theory:

1. The genetic code can be tested for extreme optimality. While the evidence that existed at the writing of the VWG paper [4, 26] suggested that the genetic code is highly optimal, if further studies of the genetic code were to show using carefully controlled statistics that the code is extremely optimal, those studies would support the VWG theory.

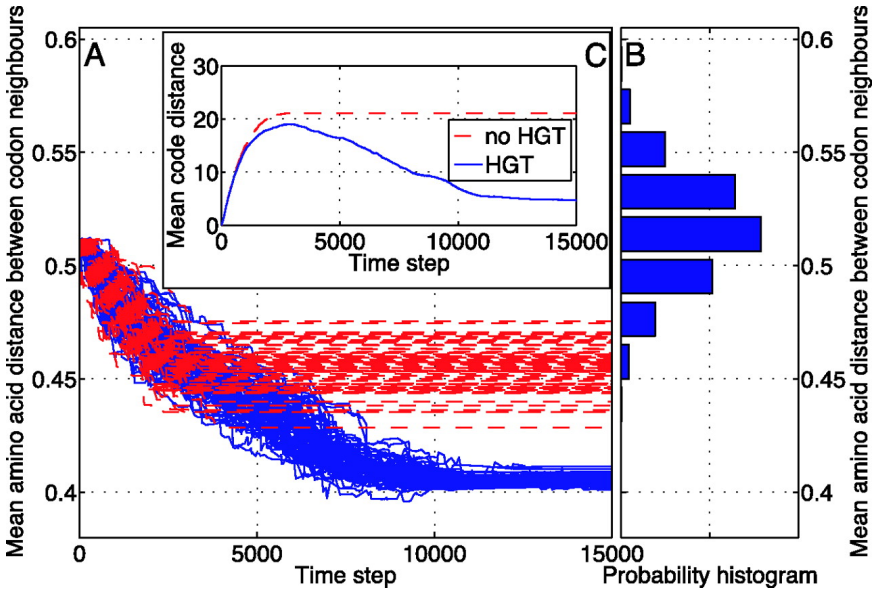

Figure 2.1: Predictions of the VWG model for code evolution. The vertical axis indicates the level of optimality a given code has, with lower values corresponding to higher optimality. The horizontal axis is time. Red trajectories correspond to evolution without horizontal gene transfer, and blue trajectories describe code evolution with horizontal gene transfer. The inset describes how much the evolved codes differ from each other, and the histogram on the right hand side describes the probability distribution of optimality for random codes. Figure taken from [1]

.

2. Evidence for pervasive horizontal gene transfer in early life would provide evidence for condition two of the theory.

3. Evidence for the statistical nature of early proteins would provide evidence for condition one of the theory.

In the following sections, we pursue the first and third of these lines of inquiry through careful statistical analysis of the genetic code. First, we review the basic properties of the genetic code and introduce some important aspects of the physics of amino acids from an evolutionary perspective. Second, we show that the genetic code is an order of magnitude more optimal than previously thought. We will then show that the genetic code carries evidence in its internal structure for the statistical nature of early proteins.

2.3 Basic properties of the genetic code

. To quantitatively analyze the genetic code, its basic features need to be described. The genetic code summarizes how RNA transcripts are translated into amino acids to form proteins, and, as noted above, is shared across the three domains of life, with only a few very minor variations[28, 29]. It represents a complex series of biochemical steps that comprise all known cell's translation apparatus. The code translates sets of three nucleotide bases (codons) which can take four possible values (adenine, guanine, cytosine and uracil, indicated by the first letter of their names) into one of twenty possible amino acids. The genome of an organism consists of strings of these bases that are translated into strings of amino acids, forming functional proteins that ultimately give rise to each organism's phenotype. Since there are four bases, and each codon consists of three bases, there are $4^3 = 64$ possible codons (see table 2.1). Given that the actual genetic code only represents 20 amino acids, there is considerable freedom in the arrangement of the code, as well as considerable redundancy. Thus almost immediately after its elucidation, attempts were made to explain the assignment of codons to amino acids. It was noticed that amino acids with related properties were grouped together, which would have the effect of minimizing translation errors[30, 31, 20]. In order to determine whether or not this was a genuine correlation or simply a fluctuation reflecting the limited size of the codon table, the canonical genetic code was compared to samples of randomly-generated synthetic codes, starting with the early but inconclusive Monte Carlo work of Alff-Steinberger[32], and compellingly revisited with larger sample sizes by Haig and Hurst[26]. Depending on the measure used to characterize or score the sampled codes, high degrees of optimality have been reported. For example, using an empirical measure of amino acid differences referred to below as the "experimental polar requirement" (EPR) [33, 34], Freeland and Hurst calculated that the genetic code is "One in a million" [4, 26, 35] (albeit with poorly controlled statistics). More recently, it has been shown that when coupled to known patterns of codon usage, the canonical code (and the codon usage)

The canonical genetic code					
	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Stop	Stop	A
	Leu	Ser	Stop	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

Table 2.1: The canonical genetic code. The first base in the codon is given by the row, the second base is given by the column, and the third base is given by the subsections within the rows. For example, UUA maps to Leu.

is simultaneously optimized with respect to point mutations and to the rapid termination of peptides that are generated with frame shift errors [36].

As noted above, the results of the VWG theory for code evolution show clearly that vertically-dominated evolution is only capable of a relatively weak degree of optimization, failing to find global extrema, and neither strongly optimized nor converged to a unique code. On the other hand, only if the evolutionary dynamics are horizontally-dominated, the observed modularity of structures such as the translation apparatus and the genome emerge naturally[37], and optimization is strong, rapid and convergent to a universal genetic code [1]. Thus, the known structure of the genetic code and translation apparatus reflects the evolutionary dynamics from which the code emerged as presented in the VWG theory.

In the next two sections, we will provide two new pieces of evidence for the VWG theory of the collective evolution of the genetic code. First, we set an extremely high lower bound on

the level of optimality of the canonical genetic code by using molecular dynamics to construct a measure of code optimality, the "computational polar requirement" (CPR) without any input from experiment. We then use Monte Carlo simulation to determine the level of code optimality, and find that the level is so high that a new and detailed error analysis is required to ensure statistically significant assessment of very small probabilities. Second, we explore the dependence of our results on the scale of code variations. Our results indicate a level of optimization that would only be attainable from collective dynamics[1], and a dependence on scale that indicates that the dynamics involved the refinement over evolutionary time of an ambiguous primitive translation machinery. Ambiguous translation generates a statistical ensemble of related proteins ("statistical proteins")[20, 27] rather than a unique protein, as is now the case, and is the first requirement for the coevolutionary mechanism proposed in the VWG theory for code evolution [38, 1].

2.4 Molecular Dynamics of the Polar Requirement

The experimental polar requirement is a chromotagraphic measure of amino acid affinity to a water-pyridine solution that was originally motivated by a simple stereochemical theory of the origin of the genetic code [39, 20, 33, 34]. This measure is related to, and strongly correlated with, several other amino acid measures, such as hydrophobicity and Grantham polarity [40]. In the EPR experiments, water/dimethylpyridine (DMP) ratios ranging from 40-80% mole fraction water were used for chromotographic separations of each amino acid measured. When the chromatographic factor, R_m was plotted as a function of mole fraction water in log-log scale, a linear trend was observed for each amino acid. The slope of the corresponding best fit line was taken to be the amino acid's EPR.

The methods used for obtaining the computational polar requirement numbers (CPR) are reported elsewhere [41] and are summarized here. The distribution of solute molecules across the water/DMP interface is related to the equilibrium solvent environment surrounding the

molecules in a binary solution similar to that used in the experiments. Trends in the local water density of a solvated amino acid in water/DMP solutions were found to be linear functions of mole fraction water. The slopes of these linear trends were used to obtain a set of computed CPR values. To quantitatively measure the differences in local solvent environment, molecular dynamics (MD) calculations were performed using NAMD2 software with an NPT (number, pressure, temperature) ensemble [42] and the Charmm 27 force field [43, 44]. Standard pressure and temperature were maintained for the simulations. The systems consisted of a single amino acid molecule in a box of water and randomly placed DMP molecules of a determined water/DMP ratio. For each amino acid at least four systems, each with a different water/DMP ratio, were simulated. Radial distribution functions (RDFs) of water relative to the amino acid side chains were calculated from the equilibrated MD trajectories using VMD (Visual Molecular Dynamics) [45]. The RDFs were calculated by a time average over the equilibrated portion of a trajectory [46]. The most distant atom of the amino acid side chain was used as a reference atom, and the oxygen or hydrogens (as appropriate) from the water molecules were used as a selection in calculating the RDFs. Calculated in this manner, the maximum value of the first peak in an RDF is related to the relative density of water in the first solvation layer of the amino acid side chain. It was found that these maxima varied linearly with water/DMP ratios for each amino acid, and that the slopes of the corresponding lines was strongly correlated with the EPR ($R^2 = 0.92$) (Fig. 2.2). We confirmed that tyrosine's large deviation from the experimental value was not due to a weak signal in the RDF.

2.5 Optimality analysis of the canonical genetic code

To analyze the CPR, we used the point mutation code analysis algorithms described in [26] and [4] along with an analytical realization of bootstrap error analysis to assess the statistical significance of the results. The algorithms treat the genetic code as a mapping

Polar requirements		
Amino Acid	Computational Polar requirement	Experimental Polar Requirement
Ala	6.63	7.0
Arg	8.63	9.1
Asn	9.63	10.0
Asp	12.18	13.0
Cys	4.56	4.8
Gln	8.96	8.6
Glu	13.47	12.5
Gly	9.02	7.9
His	8.04	8.4
Ile	5.20	4.9
Leu	4.63	4.9
Lys	10.21	10.1
Met	5.26	5.3
Phe	4.78	5.0
Pro	6.28	6.6
Ser	7.61	7.5
Thr	6.41	6.6
Trp	5.16	5.2
Tyr	7.77	5.4
Val	6.36	5.6

Table 2.2: Computational and experimental polar requirements

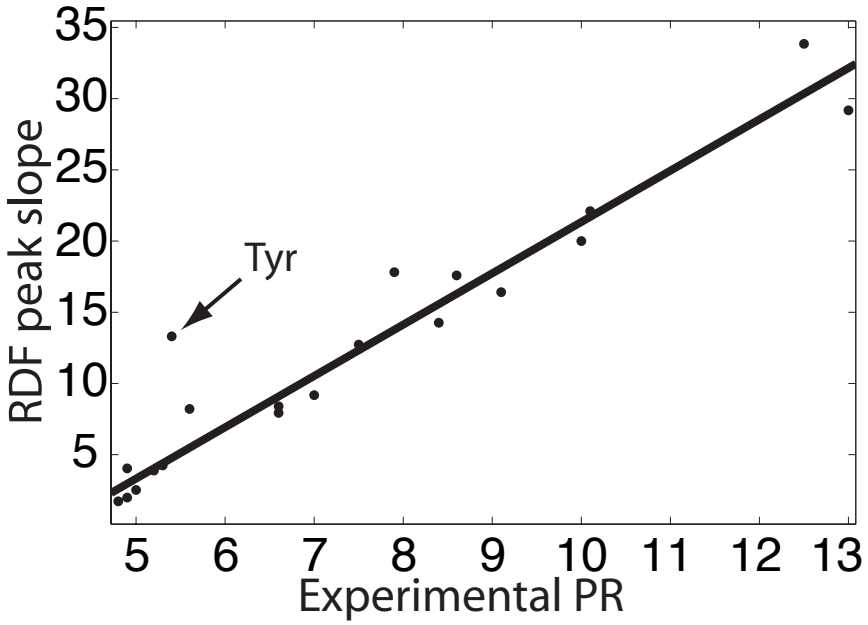

Figure 2.2: Scatter plot showing the relationship between radial distribution function (RDF) peak slope and experimental polar requirement for all amino acids. The straight line is a guide to the eye [7].

$GC^i : Codons \rightarrow Amino\ Acids$, where i indexes a particular set of assignments of codons to amino acids, with GC^1 as the canonical code. *Codons* is the set of codons excepting the termination codons, and *Amino Acids* is the set of amino acids, i.e. $GC^1(UUU) = Phe$. New versions $GC^{i\neq 1}$ of the mapping are generated by randomly permuting amino acid labels, leaving termination codons fixed. This preserves the degeneracy structure of the genetic code. The average impact I_i of mutations for a given realization of the genetic code GC^i is assessed by evaluating the sum

$$\begin{aligned} I_i &= \sum_{\langle c,c'\rangle \neq Ter} W_{c,c'}\ d^q(GC^i(c), GC^i(c')) \\ &= \sum_{\langle c,c'\rangle \neq Ter} I_{c,c'} \end{aligned} \tag{2.1}$$

Table 2.3: The matrix $W_{c,c'}$ of transition/transversion biases taken from [4].

	First Base	Second Base	Third base
Transitions	1	0.5	1
Transversions	0.5	0.1	1

where $\langle c, c' \rangle \neq Ter$ denotes a sum over nearest neighbor codons with the nearest neighbors of a codon defined by its single point mutations, with all mutations to or from a termination codon excluded. The matrix $W_{c,c'}$ weights transition mutations (U $\leftrightarrow$ C and G $\leftrightarrow$ A) and transversion mutations (U/C $\leftrightarrow$ G/A), according to a toy model of typical transversion/transition biases in real translation. In our calculations, we used the values from [4] as listed in table 2.3. Finally, $d^q(x, y)$ is a metric on the space of amino acids. For the polar requirement, the metric is taken to be $d^q(x, y) = |x - y|^q$ over the polar requirement values corresponding to the given amino acids.

The appropriate quantity to compute is the probability $P_b = Pr(I < I_1)$ that a random realization is more optimal than the canonical code. To compute P_b, we count the number of randomly generated codes that are more optimal than the canonical code and divide by the total number of random codes generated. P_b is invariant to uniform linear rescaling of the amino acid polar requirement data, and is smaller for more optimal codes while including the effects of the large number of codes that can be explored, rather than the simple linear scale provided by the bare optimality score.

The error in the computed P_b can be estimated using an analytical realization of bootstrap resampling. Simulated data sets for bootstrap are created by randomly sampling optimality scores from the original data set. When the samples are drawn from the original set, there are only two alternatives: a more, or less optimal code can be sampled, with probability $P_b = N_{I<I_1}/N_{total}$ of drawing a random code better than the canonical code. Since the number of better codes in a sample is the number whose error we wish to estimate, we can regard drawing a better code as a step to the right with probability P_b in a one dimensional random walk. The known formulas for the asymmetric one dimensional random walk allow

us to compute the bootstrap error estimate in the limit of infinitely many resampled sets, i.e. the exact bootstrap estimate. For metrics under which $P_b \ll 1$ holds, we obtain the variance in P_b to be

$$var[P_b] = var[\frac{N_{I<I_1}}{N_{total}}] = \frac{P_b(1-P_b)}{N_{total}} \approx \frac{N_{I<I_1}}{N_{total}^2} \tag{2.2}$$

To obtain a reasonable estimate of error, or to compare the results of different metrics on the space of amino acids, the number of more optimal codes, $N_{I<I_1}$ from the random sample must be sufficiently large ($\sqrt{N_{I<I_1}} \ll N_{I<I_1}$, or about $N_{I<I_1} = 10$ as a reasonable minimum).

When the computational polar requirement difference squared is used in the amino acid metric, $P_b = (19 \pm 4.36) \times 10^{-8}$. In contrast, with the experimental polar requirement, $P_b = (26.5 \pm 1.63) \times 10^{-7}$, an order of magnitude improvement. To assess the impact of tyrosine (which had the largest variation between the CPR and EPR values) on these results we redid the calculation of P_b for the CPR, but with tyrosine replaced with the value from the EPR. The result is ($P_b = (9.3 \pm 1.0) \times 10^{-7}$). To test the sensitivity of the results for the CPR, we varied each element of $W_{c,c'}$ independently by $\pm 0.1 \times W_{c,c'}$ and repeated the calculation of P_b. This led to results that were statistically indistinguishable from the results reported above. Shorter computations (justified by the faster convergence due to decreased optimality) for the EPR indicate a similar level of robustness. With a $W_{c,c'}$ uniform among nearest neighbors we saw substantial increases in P_b in agreement with [26]. However, the CPR continued to be superior to the EPR, with the CPR yielding $P_b = (3.7 \pm .61) \times 10^{-5}$ and the EPR yielding $P_b = (11.8 \pm 1.1) \times 10^{-5}$. This order of magnitude improvement in the measured optimality of the genetic code lends strong support to the VWG theory of code evolution, and in particular, supports the notion that horizontal gene transfer enabled the evolution of an extremely optimal universal genetic code

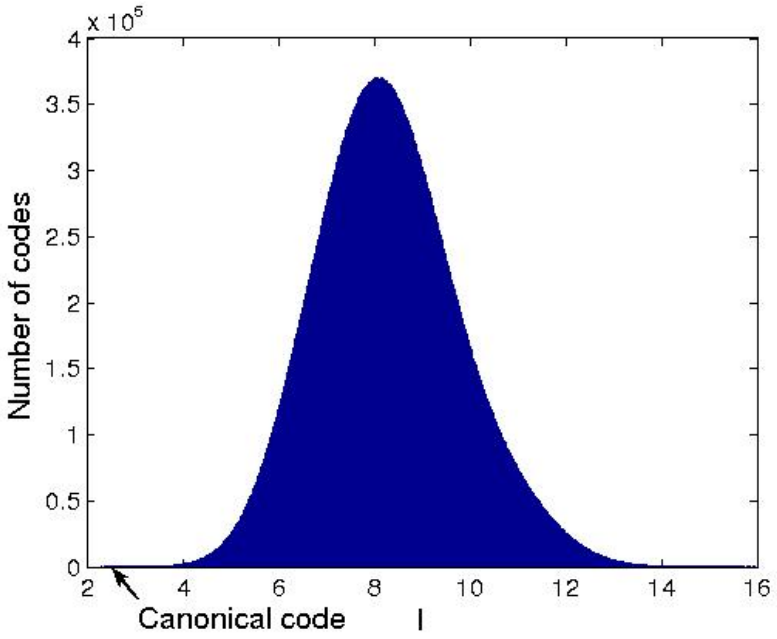

Figure 2.3: Histogram of codes with different average mutational impacts generated through the monte carlo algorithm described above. The extreme optimality of the code can be seen clearly.

2.6 Evidence for statistical proteins in early life

Varying the value of q in the metric [47] provides a further probe to explore the optimization of the genetic code. Increasing the value of q is equivalent to emphasizing the role of larger and larger differences between the amino acid intended, and the one generated by point mutation. Thus, if P_b reduces for increased values of q, the code (along with $W_{c,c'}$) evolved to suppress the effects of rarer, possibly catastrophic errors that may be generated by point mutations. This may happen primarily by evolving small elements of $W_{c,c'}$ where $c \to c'$ is catastrophic. Conversely, if P_b reduces for smaller values of q, the code evolved to both mitigate the possibility of these catastrophic errors, and to minimize the effects of frequent, small errors. Varying q we find that the canonical genetic code is most optimal for q between one and two with significant increases outside this regime in either direction (Fig. 2.4). This indicates that the genetic code is optimized for minimizing errors according to their size

with no emphasis given to larger or smaller errors. Given the relative weakness of the code when emphasizing large errors, evolution must have favored organisms that discarded or edited fatally flawed proteins over evolving the code to make them less likely at the cost of reducing its ability to minimize the more frequent moderate and minor errors. The weakness of the canonical code when minor errors are emphasized ($q < 1$) suggests that while the code was still evolving, minor errors were on the whole less important biologically, as would be expected in evolutionary dynamics [1, 38] that utilized ambiguity tolerance in early proteins [20, 27].

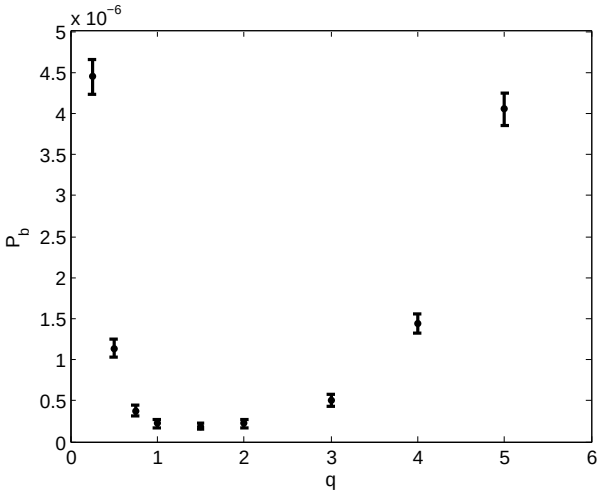

Figure 2.4: P_b as a function of the exponent q in the amino acid metric.

As stated above, ambiguity tolerance in early proteins is the first condition in the VWG theory for the successful evolution of an extremely optimal universal genetic code. As far as we are aware, the evidence presented in this analysis is the first empirical evidence of early statistical proteins or equivalently, ambiguity tolerance.

Table 2.4: P_b for several naturally occurring variant codes

Code	P_b
Canonical	$(19 \pm 4.36) \times 10^{-8}$
Yeast Mitochondrial	$(11 \pm 3.32) \times 10^{-8}$
CDH Nuclear Code	$(21 \pm 4.58) \times 10^{-8}$
Ascidian Mitochondrial	$(583 \pm 24.15) \times 10^{-8}$
Echinoderm Mitochondrial	$(51 \pm 7.14) \times 10^{-8}$

2.7 Optimality analysis of alternative codes and measures

To further explore the evolution of the genetic code, we analyzed the genetic code using other measures of amino acid physics and also analyzed a selection of variant codes using the CPR. Our findings, displayed in table 2.4 were consistent with the previous findings of Knight in that the alternative codes did not show marked improvements in optimality over the canonical code [35]. This is consistent with our expectation that evolutionary pressure to optimize the code with respect to the polar requirement was eased after the last universal ancestral state as proteins became longer, and other more accurate proofreading mechanisms evolved.

We also tested Grantham polarity [40], which has been argued in a survey of genetic code optimality under different amino acid measures to be the amino acid measure most optimized by the genetic code [35]. The results yield $P_b = (285 \pm 16.88) \times 10^{-8}$, or an order of magnitude higher than with the CPR metric, leading to the conclusion that the CPR is the most effective known metric for optimization of the genetic code by an order of magnitude. Previous computations evaluated P_b by generating $100,000$ random codes [35]. Scaling our results to the size of these original simulations, we see that the optimality of the genetic code measured from the EPR and the Grantham polarity are virtually identical. Scaling the errors for the CPR and the Grantham polarity to errors assessed from only $100,000$ codes, we get for the CPR, $P_b = (0.19 \pm 0.44) \times 10^{-5}$ and for the Grantham polarity,

$P_b = (2.85 \pm 1.69) \times 10^{-5}$. These results are within a standard deviation and a half of each other, and are therefore not different in a statistically meaningful way. Thus the conclusion from earlier studies [35] that the Grantham Polarity maximizes the optimality of the genetic code was based on sample sizes that were too small to distinguish between different results from different amino acid measures. Our updated calculations utilize sufficiently large data sets and the bootstrap error analysis described above to conclude that the CPR measure of amino acid properties exhibits the highest degree of optimality.

In conclusion, earlier estimates of code optimality were understated by a statistically significant amount. The extent of optimality and its dependence on metric revealed here further support the notion that the genetic code must have evolved during an early communal state of life[1].

Chapter 3

Optimality of a Proposed Precursor to the Genetic Code

3.1 Precursor code proposals

In the previous chapter, we exhibited new evidence that the canonical genetic code is not a frozen accident, but exhibits a pattern of amino acid-codon correspondences that have the effect of making the code insensitive to certain classes of point mutation or translation error [31, 39, 32, 26, 4, 48, 49]. A variety of schemes [50], including the VWG theory of horizontal gene transfer dynamics outlined and supported in the previous chapter [1] and stereochemistry [51, 52], have been put forward to explain this pattern and others[53] in the genetic code (for recent reviews, see [54, 55]). It is important to stress that while the code exhibits some optimality with respect to several measures, such as hydrophobicity [26], the code exhibits extreme optimality with respect to only one particular class of amino acid attributes, related to the free amino acid polar requirement [33, 34], and this suggests the code is a very ancient part of the cell's machinery, functioning either in its present role of translation, or in some earlier unknown function. This result lends strong support to the suggestion that the code's evolutionary dynamics was dominated by collective mechanisms arising from horizontal gene transfer [1]. Computational evidence shows that core chemical affinities in the genetic code are fully compatible with, and independent from, evolutionary dynamics that lead to error minimizing optimality [56], suggesting that error-minimizing optimality is not a by-product of chemistry but arises from the evolutionary dynamics.

To extend the work of the previous chapter, we attempt to ascertain to what extent, if any, error-minimizing optimality can be used to constrain a proposed scenario for a precursor

to the genetic code. If the optimality with respect to polar requirement was a feature of the code from very early times, then precursors to the canonical code must respect error-minimizing optimality to a significant degree. Alternatively, proposed precursor codes may claim to date prior to any code evolution, and to be the product of other factors alone. Such precursors would not be expected to display a significant level of error-minimizing optimality, assuming that it is indeed the case that optimality is primarily a reflection of evolutionary dynamics. Here we show that a specific biochemically-motivated precursor code does not show evidence for significant error-minimizing optimality, even though it is a projection of the canonical code; these results support the notion that error-minimizing optimality primarily reflects evolutionary dynamics, and imply that this type of precursor code, if it ever existed, would have arisen prior to the emergence of translation.

3.2 The Copley Smith Morowitz precursor code

Copley, Smith and Morowitz have suggested that first and, to a lesser extent, second base assignments in the canonical code would arise if the code has its origin in amino acid synthesis channels embedded in dinucleotide complexes prior to the emergence of translation [5]. The proposal exploits the strong constraints such a theory imposes on the first two bases of the genetic code to generate a specific precursor doublet code based on a projection of the canonical genetic code to a doublet code. For most of the projection, the third codon is sufficiently redundant that the first two bases are sufficient to define the amino acid coded for by doublet. In the event that the third bases associated with a doublet codon code for multiple amino acids, the proposal favors the simpler of the amino acids (table 3.1). They further refine the proposal by incorporating possible precursor amino acids motivated by their study of the biosynthetic pathways for amino acids (not shown) [5].

To further assess and characterize the proposed precursor code in [5] we analyze the degree to which it contains error-minimizing optimality. As noted above, the proposed precursor

Proposed Precursor Code				
	G	**C**	**A**	**U**
G	Gly	Ala	Asp	Val
C	Arg	Pro	Gln	Leu
A	Ser	Thr	Asn	Ile
U	Cys	Ser	Tyr	Leu

Table 3.1: Proposed precursor code from Ref.[5]. Row is first base, column is second base.

code is based primarily on arguments about biosynthetic pathways rather than evolutionary considerations. Additionally, it explicitly dates to prior to translation [5]. All mechanisms of which we are aware for code evolvability explicitly require translation machinery (see for example [57, 58, 29, 59, 60, 1]). Thus we anticipate that the proposed precursor code should contain little, if any, evidence for optimality.

We have analyzed the former of these proposed precursor doublet codes (see table) for error-minimizing optimality using the "experimental polar requirement" (EPR) [39, 20, 33, 34] derived originally by Woese and co-workers. We have also analyzed the precursor using a modern computational update of the polar requirement (CPR) [41]. Analysis with the CPR is of particular interest, because it is the measure of amino acid difference that when applied in code optimality analysis algorithms to the canonical genetic code gives rise to the extreme optimality cited above [49]. Thus the CPR can be considered to capture some essential aspect of amino acid chemistry of particular relevance during the evolution of the genetic code. Analysis of the more refined version of the proposed precursor code is difficult due to the fact that the polar requirements for the proposed precursor amino acids are unknown. This problem can be partially solved by sensitivity analysis, and is discussed in greater detail below.

To analyze the error-minimizing optimality in the proposed precursor code, we used the point mutation code analysis algorithm described in [26] and [4] to calculate a measure of the average impact of a point mutation of a given code indexed by i. Calculating the impact per point mutation allows direct comparison of the optimality of the canonical and precursor

codes, because the different size of the set of point mutations for the doublet versus the canonical code has been divided out. With this convention, the optimality distributions for doublet codes and triplet codes are similar (see figs. 3.1,3.2).

As in the previous chapter, the algorithm is to consider an ensemble of random genetic codes that are mappings from the set of codons (minus the termination codons) to the set of amino acids, $GC^i : Codons \rightarrow Amino\ Acids$, where i indexes a particular set of assignments of codons to amino acids, with GC^1 as the precursor code. Versions $GC^{i\neq 1}$ are generated by randomly permuting amino acid labels, again excluding termination codons. A measure of the average impact I per point mutation for a given code i, can then be calculated as

$$I_i = \frac{\sum_{\langle c,c'\rangle \neq Ter} (GC^i(c) - GC^i(c'))^2}{\sum_{\langle c,c'\rangle \neq Ter} 1} \tag{3.1}$$

where $\langle c, c'\rangle \neq Ter$ denotes a sum over nearest neighbor codons with the nearest neighbors of a codon defined by its single point mutations, with all mutations to or from a termination codon excluded.

To extract a measure of optimality that restricts optimality comparisons of the precursor codes to other doublet codes, we compute the probability $P_b = Pr(I < I_1)$ that a random realization is less impacted by point mutations (more optimal) than the proposed precursor code. This can be achieved by calculating the percentage of random doublet codes that are more optimal than the precursor code. If we are computing the optimality of the canonical code, P_b is calculated strictly from an ensemble of triplet codes. The fact that P_b is based on strict comparison to the appropriate ensemble of random codes will allow us to compare P_b from the proposed precursor to that of the canonical code.

The error in the computed P_b can be estimated using an analytical realization of bootstrap resampling derived from an exact correspondence with the statistics of the asymmetric one dimensional random walk [7]. This correspondence shows that if N codes are sampled, and $N_{I<I_1}$ are more optimal than the code being tested, then P_b with standard error is given by

the expression

$$P_b = (N_{I<I_1} \pm \sqrt{N_{I<I_1}})/N \tag{3.2}$$

While this is in line with naïve expectations for the form of error, the problem of sampling more optimal random codes is a problem of rare event sampling, which is frequently unstable and prone to nonstandard large errors. This makes a rigorous derivation of the exact error a key result essential for robust interpretation of optimality calculation results. The form of the error also informs the computations. It is clear from Eq. 3.2 that the relevant sample size for a statistically sound analysis is not N, but the number of more optimal codes sampled, $N_{I<I_1}$ [7]. A reasonable minimum is, perhaps, 20 more optimal codes sampled to get a statistical estimate. Much larger samples would be preferable, but in many applications may be hard to obtain due to computational limitations encountered when analyzing highly optimized codes.

When applied to the proposed precursor code, we calculated $P_b = (1.44 \pm 0.038) \times 10^{-2}$ with the experimental polar requirement, or $P_b = (7.95 \pm 0.282) \times 10^{-3}$ with the computational polar requirement [41]. To compare, we applied this simplified code analysis algorithm to the canonical genetic code. The canonical genetic code has optimality of $P_b = (1.18 \pm 0.109) \times 10^{-4}$ or $P_b = (4.7 \pm 0.686) \times 10^{-5}$ with the EPR and CPR respectively (the extreme optimality discussed above included transition and transversion biases for each base position in the calculation [4, 7]). Thus the optimality of the precursor is, with either the EPR and the CPR, two orders of magnitude less optimal than the canonical genetic code evaluated with the equivalent algorithm. The absolute I for both codes can also be compared because they are calculated per point mutation (see discussion above). For the canonical code, $I = 5.293$, which we know from P_b to highly optimal. For the precursor, $I = 7.498$. Given that the mean of the I distribution is near 10 for both the doublet and the triplet case (figs. 3.1,3.1), the optimality is substantially reduced for the doublet, consistent

with the results from P_b.

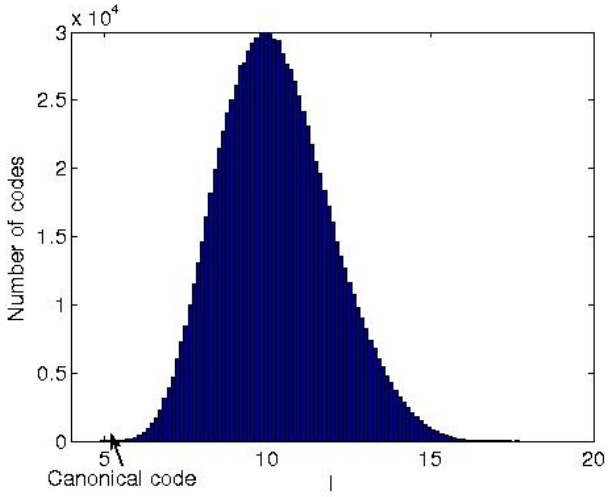

Figure 3.1: Histogram of average impact I per point mutation for randomly-generated codes with the same degeneracy structure as the canonical genetic code.

As discussed above, the derivation of the doublet code in table 3.1 depended on projecting the third base onto the doublets by favoring the simplest amino acid coded for by the triplet codons associated with a given doublet. We repeated the optimality analysis for versions of the doublet code that favored more complex amino acids at individual doublets (such as substituting Arg for Ser at the AG position). None of the modified doublet codes displayed a significant increase in optimality over the version in table 3.1.

We also note that the version of the precursor code we studied used some amino acids that are regarded as late additions [61]. While it seems unlikely that the later amino acids would have substantially different polar requirements than their predecessors in the same synthesis path, to assess the impact of possible changes in polar requirement values as these amino acids (Arg, Gln, Asn, Ile, Cys) were introduced, we varied their polar requirement values $\pm 20\%$, and redid the optimality calculation. In all cases, the optimality of the precursor declined, or showed such small improvement that the error bars overlapped with the primary calculation, leaving our basic conclusions about the optimality of the precursor code unchanged. This analysis shows that our results are unlikely to be changed when analyzed

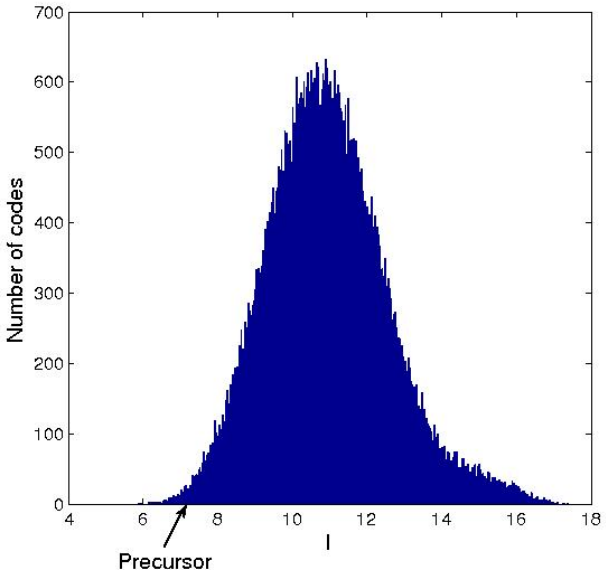

Figure 3.2: Histogram of average impact I per point mutation for randomly-generated codes with the same degeneracy structure as the proposed precursor. There is more noise relative to the canonical code case due to the smaller ensemble of random codes required to calculate P_b for the precursor.

with all of the polar requirements for precursor amino acids proposed in [5]. Since varying individual amino acid polar requirement values did not enhance the optimality properties of the precursor, a version of the precursor code which is highly optimal and respects the underlying biosynthesis theory would differ in several positions from the proposal by Copley et al.[5].

Our results show that the proposed precursor code has weak error-minimizing optimality with respect to the polar requirement, compared to the canonical genetic code. This result is surprising in one respect, because the doublet code is a projection of the canonical code. A number of interpretations are possible. (1) The doublet precursor code is not an intermediate evolutionary stage from some earlier precursor code; this is consistent with the basis for the original proposal of this code as a biosynthetic pathway, but is puzzling because the later canonical triplet code is optimized with respect to the free amino acid polar requirement. (2) The precursor has no biological significance at all, and did not evolve from an earlier

precursor, which exhibits free amino acid polar requirement optimality. (3) The precursor doublet code predates evolution for error minimization, and if the amino acid synthesis scheme is correct, then modifications to the doublet code during its evolution to today's canonical code are responsible for its observed error-minimizing optimality. The relatively large P_b value (i.e. small amount of observed optimality) in the precursor is an artifact of deriving the doublet code from the highly evolved canonical code.

Our analysis does not address the question of whether or not the detailed biochemical theory proposed is correct, because presumably optimal precursor codes that are both consistent with the biochemical theory and uncorrupted by evolution could be constructed.

Chapter 4

Field Theory of Fluctuating Predator-Prey Populations in Space

4.1 Ecology as a many-body problem

The previous chapters focused primarily on the problem of early evolution. The focus of this chapter is to introduce ecology as a complex many-body problem through the nontrivial example of predator-prey population oscillations. We will elucidate the origin of many predator-prey oscillations by applying the techniques of many-body theory to show that the oscillations have their origin in fluctuations and correlations that cannot be captured with less sophisticated tools. The following chapters will generalize these insights to provide new insights into spatiotemporal pattern formation in ecosystems and in the brain.

Ecosystems, while very complex and containing an enormous amount of particularity, also exhibit many common, if not universal, features. Some of these features include spatial patterns of population abundance, a distance of three or fewer links between pairs of species in a food web, Levy flight foraging by predators in areas of low prey abundance, etc. [15, 62, 63, 64]. While each of the examples cited above, and many others, can be fruitfully studied with statistical physics techniques, in this chapter we focus on the collective oscillatory dynamics of predator and prey species pairs, which are extremely common in real ecosystems [65].

4.2 Failure of traditional predator-prey theory

As stated above, it is common for predator-prey species pairs to exhibit population oscillations [65]. The mechanism (see fig. 4.1) for these persistent oscillations is easily understood:

1. Consider initial conditions with a large abundance of prey, and low abundance of predators. In these conditions, the predators increase in abundance through predation.

2. As the predator abundance increases, the prey abundance decreases due to increased predation.

3. As the prey abundance decreases, the food supply for the predators is decreased, leading to a decline in predator abundance.

4. With reduced predator abundance, the prey abundance increases. This restores the conditions of step one, restarting the cycle.

To model these population cycles, we will initially adopt the traditional techniques of theoretical ecology to capture the feedbacks in the above steps. We will then show that the resulting models are inadequate, and briefly review the response of traditional theoretical ecology to this challenge and show its inadequacy [2]. Finally, the disagreement between observations and ecological theory will be resolved through a rather surprising source – the oscillations are driven by the intrinsic fluctuations and noise that accompanies the many-body nature of interactions between organisms [2]. We will demonstrate this claim by systematically analyzing a full many-body theory of predator-prey interactions. In the chapters that follow, we will apply the formalism and insights developed in this chapter to related problems in ecology and neuroscience while pointing out opportunities for further

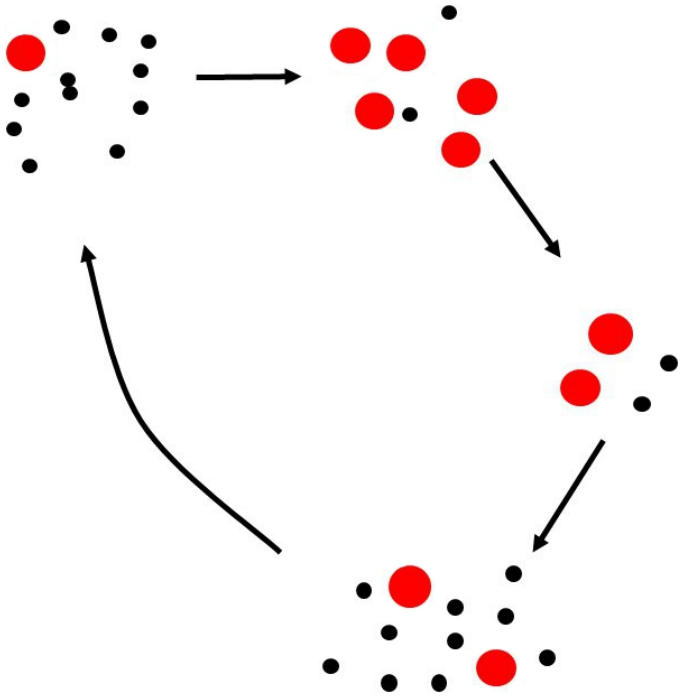

Figure 4.1: The mechanism of predator prey oscillations. Predators are red dots, and prey are black dots. Step one is in the upper left.

extensions.

The traditional modeling approach to population dynamics such as the predator-prey dynamics presented above is to model the population of each species as a single number, the concentration [66], and to then phenomenologically model the basic interactions of the concentrations in a manner similar to Landau theory in condensed matter systems [67, 11, 14]. As in mean field theory in statistical mechanics where large collections of degrees of freedom are also reduced to one degree of freedom, using a concentration instead of representing each organism independently decreases the mathematical complexity of the model by decreasing the number of degrees of freedom to one for each species rather than one for each organism. In fact, we will show later that the approximation of a population by a concentration is a mean field theory in exactly the same sense as in statistical mechanics, with the same advantages and disadvantages [66, 2, 49, 13]. While clearly unrealistic in their neglect of fluctuations, the standard mean field models of ecology are easily analyzed using

the methods of nonlinear dynamics and have been successful in many cases at effectively capturing the effects of interactions between organisms [68, 67].

4.2.1 Mathematical models of predator-prey interactions

To model the predator-prey feedbacks outlined above, we start with a population φ of predators and a population ψ of prey. The simplest equations for the feedbacks are then [69, 67]

$$\frac{d\varphi}{dt} = p_1\varphi\psi - d_1\varphi \tag{4.1}$$

$$\frac{d\psi}{dt} = b\psi - p_2\varphi\psi \tag{4.2}$$

where p_1 is the rate of predator reproduction through predation, d_1 is the predator death rate, p_2 is the predation rate and b is the birth rate of the prey. As can be verified by straightforward linear stability analysis, this equation does predict cyclic behavior, as required by the intuitive argument above as well as by observation. But the cycles that result are not structurally stable (meaning they are not robust to small changes in the underlying equations of motion) and are therefore unsatisfactory as a model [70]. The dynamical manifestation of this structural instability is that the amplitude of the cycles depends on initial conditions. However, real population cycles are limit cycles with robust periods and amplitudes. Additionally, this model makes the troubling prediction that if the predator population goes to zero, the prey population will grow exponentially without bound.

To correct these defects, the next simplest theory adds a term to limit the unbounded growth of the prey, and hopefully correct the structural instability of the solutions. To do this, we adopt for the prey "logistic growth"

$$\frac{d\psi}{dt} = b\psi - c\psi^2 \tag{4.3}$$

This equation adds a population limiting term to the prey only dynamics, so that when there are no predators in the system, there is a fixed point at $\psi = b/c$. The quadratic term is usually interpreted to correspond to an increase in death rate due to competition for resources once the population becomes large enough. This is the logistic equation, which is as central to population dynamics as the Schroedinger equation is to quantum mechanics. With diffusion added, it is the Fisher equation, which supports traveling wave fronts [71]. Using these dynamics for prey will correct the prediction of the predator-prey Eqs. 4.2 that prey population is unbounded when predators are absent. Incorporating these dynamics for the prey results in the predator-prey equations

$$\frac{d\varphi}{dt} = p_2\varphi\psi - d_1\varphi \tag{4.4}$$

$$\frac{d\psi}{dt} = b\psi - c\psi^2 - p_1\varphi\psi \tag{4.5}$$

These equations, known as the Lotka-Volterra equations, are structurally stable and capture all of the feedbacks in the heuristic argument presented above. However, they no longer predict persistent periodic behavior. As can be shown from simple linear stability analysis, the solutions to Eq. 4.5 are stable spirals that converge on a stable coexistence fixed point. Further elaborations such as satiation are required to generate stable limit cycles [72]. This is very troubling, because the cycles can be understood through only the mechanisms presented in the verbal arguments, so requiring the models to have further mechanisms to explain the cycles in every case is awkward at best. In the following section, we will outline a more satisfactory solution and verify it through detailed calculation.

4.3 Individual level modeling of predator-prey population dynamics

As noted in the previous section, among the simplifying assumptions in the Lotka-Volterra equations for predator-prey dynamics is that the populations can be modeled as concentrations represented by a continuous variable. This assumption neglects all of the correlations and noise that arise from the many-body interactions of organisms on the individual level. Equivalently, the Lotka-Volterra equations represent continuously varying concentrations, when biological reality allows only integer jumps in population number through discrete and noisy interactions. The noise and correlations from these mechanisms is known as demographic or intrinsic noise [73]. As will be shown below, the Lotka-Volterra equations are systematically derived from a theory that includes demographic noise as a mean field limit [13, 49, 66, 2]. Thus intrinsic noise is not a model elaboration, but fundamental to the mathematical structure of predator-prey interactions, which are studied in mean field by the Lotka-Volterra equations.

Since the Lotka-Volterra equations fail to capture intrinsic population fluctuations due to the mean field limit, it is possible that they also fail to capture oscillations, which after all, are simply a more structured kind of population fluctuation. While this claim, which amounts to stating that population cycles in predator-prey populations are driven by noise, seems unusual because noise typically believed to blunt a signal, the rest of the chapter will show that this is exactly what happens. This was first observed by Alan McKane and Tim Newman in their landmark 2005 PRL [2]. In this chapter, I rederive these results and extend them to space using a more flexible field theory formalism.

Heuristically, the effects of intrinsic noise can be included in the Lotka-Volterra equation by adding noise (denoted as γ)

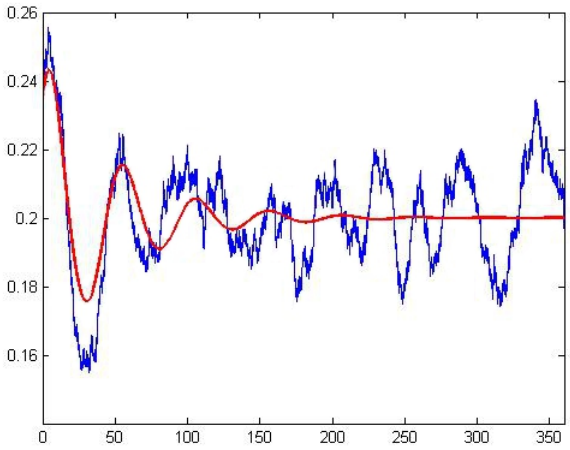

Figure 4.2: Comparison of mean field theory (red line) and the individual level model. Note the persistent noisy cycles in the individual level model and the stable spiral dynamics of the mean field theory.

$$\frac{d\varphi}{dt} = D_1\nabla^2\varphi + p_2\varphi\psi - d_1\varphi + \gamma_1 \tag{4.6}$$

$$\frac{d\psi}{dt} = D_2\nabla^2\psi + b\psi - c\psi^2 - p_2\varphi\psi + \gamma_2 \tag{4.7}$$

with the correlations generated from expansion [74] as

$$\begin{aligned}
\langle\gamma_1\rangle &= \lambda_1\varphi + D_1'\nabla^2\varphi - p_1'\varphi\psi \\
\langle\gamma_2\rangle &= \lambda_1\psi - \lambda_2\psi + D_2'\nabla^2\psi - p_2'\varphi\psi \\
\langle\gamma_1\gamma_1\rangle - \langle\gamma_1\rangle^2 &= c_1\varphi \\
\langle\gamma_1\gamma_2\rangle - \langle\gamma_1\rangle\langle\gamma_2\rangle &= -c_2\varphi\psi \\
\langle\gamma_2\gamma_2\rangle - \langle\gamma_2\rangle^2 &= c_3\psi
\end{aligned} \tag{4.8}$$

The values for the mean of noise can be absorbed into the parameters of the deterministic

portion of the equation 4.7 leading to the correlations (with the constants $c_i > 0$ and of similar order to the kinetic parameters of equal dimension in the deterministic portion of the equation)

$$\begin{aligned} \langle \gamma_1 \rangle &= 0 \\ \langle \gamma_2 \rangle &= 0 \\ \langle \gamma_1 \gamma_1 \rangle &= c_1 \varphi \\ \langle \gamma_1 \gamma_2 \rangle &= -c_2 \varphi \psi \\ \langle \gamma_2 \gamma_2 \rangle &= c_3 \psi \end{aligned} \tag{4.9}$$

These correlation functions can be understood physically as implying that the noise vanishes when the population goes to zero, and that the standard deviation of the populations from their mean values scales as the square root of population size, as would be expected from Gaussian statistics. The cross correlation $\langle \gamma_1 \gamma_2 \rangle$ is negative because fluctuations in predator and prey should be anti correlated (i.e. a fluctuation that increases predator abundance correlates with lower prey abundance).

To solve the model approximately in the limit of long times, we first solve the mean field model for long times, yielding stable fixed point values

$$\begin{aligned} \varphi^* &= \frac{bp_2 - cd_1}{p_2 p_1} \\ \psi^* &= \frac{d_1}{p_1} \end{aligned} \tag{4.10}$$

We can then look for small deviations from the long time solution in mean field, $\varphi = \varphi^* + \eta$ $\psi = \psi^* + \xi$ by discarding all terms beyond linear. Defining the vector quantities

$$\boldsymbol{x} = \begin{pmatrix} \eta \\ \xi \end{pmatrix} \tag{4.11}$$

and Fourier transforming the equations of motion for small fluctuations from equilibrium, leads to the equations of motion (with matrices indicated by upper case bold letters, and vectors by lower case bold letters)

$$-i\omega\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\gamma} \tag{4.12}$$

The matrix A, since it only consists of linear departures form the mean field equilibrium is simply the stability matrix of the mean field theory and is given after some simplification by

$$\boldsymbol{A} = \begin{pmatrix} D_1 k^2 & p_2\varphi \\ -p_1\phi & D_2 k^2 - c\phi \end{pmatrix} \tag{4.13}$$

The correlations are still given by Eq. 4.9, with the predator and prey concentrations in the correlations given by the fixed point values in Eq. 4.10. We can now solve the equations 4.12 for the power spectrum of the linear corrections to the mean field theory with the heuristic inclusion of intrinsic noise. With the correlations defined as

$$\langle \gamma_i(\omega)\gamma_j(-\omega)\rangle = B_{ij} \tag{4.14}$$

These equations are of the same form as the equations reported in [2, 75] and are easily solved using simple linear algebra manipulations [75]

$$\mathbf{x} = -(\mathbf{A} + \mathbf{i}\omega)^{-\mathbf{1}}\gamma(\omega) \equiv \mathbf{D}(\omega)^{-\mathbf{1}}\gamma(\omega)$$

$$\rightarrow x_1 = \eta = -\det(\mathbf{D})^{-\mathbf{1}}(\mathbf{D_{11}}\gamma_\mathbf{1} - \mathbf{D_{12}}\gamma_\mathbf{2})$$

$$x_2 = \xi = -\det(\mathbf{D})^{-\mathbf{1}}(\mathbf{D_{21}}\gamma_\mathbf{1} - \mathbf{D_{22}}\gamma_\mathbf{2}) \tag{4.15}$$

The appropriate statistic to analyze is the average power spectrum, because the power spectrum preserves information about noisy cycles. This is because it sums over the modulus of the trajectories, so that the cycles aren't canceled. The resulting expression is

$$P(k,\omega) = \frac{|D_{22}|^2 B_{11} - 2D_{12}Re(D_{22})B_{21} + |D_{12}|^2 B_{22}}{|det(D)|^2} \tag{4.16}$$

In this form, the expression is not very illuminating. To make further progress, we ignore space (i.e. set $k = 0$) and seek an approximate peak of the power spectrum by examining only the denominator of the rhs of Eq. 4.16. This is

$$|det(D)|^2 = \omega^2 c^2 \psi^2 + \left(p_1 p_2 \psi\varphi - \omega^2\right)^2 \tag{4.17}$$

There will be oscillations in the populations of predator and prey if there is a peak in the power spectrum Eq. 4.16 for $\omega > 0$, or to a good approximation, a minimum in Eq. 4.17 for $\omega > 0$. Conditions for this can be easily obtained by rewriting the equation as

$$|det(D)|^2 = (p_1 p_2 \psi\varphi)^2 + (c^2\psi^2 - 2p_1 p_2 \psi\varphi)\omega^2 + \omega^4 \tag{4.18}$$

If the quadratic in ω term is less than 0, then the predator-prey populations oscillate because there is a nontrivial maximum in the power spectrum. The approximate condition for predator-prey oscillations is therefore

$$c^2\psi < 2p_1 p_2 \varphi \tag{4.19}$$

The existence and consistency of the condition 4.19 indicates that considering demographic noise does indeed imply the existence of the observed predator-prey cycles as first pointed out by McKane and Newman [2]. However, the calculation just carried out was extremely heuristic. It did not start from an individual level model and build up to a description of fluctuating populations with cycles. While the argument is very plausible, to be confident in the predictions, a systematic calculation on a full many-body theory of predator-prey interactions must be carried out. Such a calculation would illuminate the reasons why the standard Lotka-Volterra model fails and would start from a rigorously defined model defined at the level of individual organisms. It is the task of the next several sections to develop the necessary formalism for such a calculation and to then carry the calculation out.

4.4 Master equation representation of individual level population dynamics

Among the primary techniques in non equilibrium statistical mechanics is to develop a master equation description for the time evolution of the probability that the system is in a given state s_i [73]

$$\partial_t P(s_i) = \sum_{s_j} \left[W_{s_j \to s_i} P(s_j) - W_{s_i \to s_j} P(s_i) \right] \tag{4.20}$$

This equation describes the flux of probability between states. In population dynamics, the states are configurations of organisms in space at a given time t. In the following sections, we will derive a master equation for a model of predator-prey interactions defined at the level of the interactions between individual organisms, an individual level model (ILM). The model will be equivalent to the Lotka-Volterra equations in the mean field limit and will include intrinsic noise automatically.

Many analyses of predator-prey at the individual level have been carried out over the

last couple of decades, especially if computer simulations are considered. In all of these analyzes, the discrete nature of predator prey populations and finite size effects lead to persistent oscillations in time, but spatial patterns fail to form. Many of these authors have been motivated by the observation made above that the simple differential equation (mean field) models of predator prey dynamics do not exhibit limit cycles [68, 67]. Several authors have addressed this difficulty by developing spatial individual level models (ILMs) that incorporate the stochastic effects of individual predator-prey interactions as in, for example, [76, 77, 78, 79]. These models yield limit cycles [79] or stochastically induced cycles dependent on space for their existence [78, 76, 77]. The present work was motivated by recent work on a 0 dimensional model that has shown that intrinsic noise without space is sufficient to generate temporal oscillations in predator-prey populations [2]. Generalization of this work to space shows oscillations in time, but fails to exhibit oscillations in space [3].

The analysis in the present chapter goes beyond previous work by developing a modified version of the spatial ILM of predator-prey interactions in [3] and analyzing the oscillatory fluctuations using path integral techniques. This allows simple analysis of spatial fluctuations, scaling analysis, and the possibility to study phase transitions and the effects of disorder using the techniques of field theory. The model we will develop includes the motion of both predator and prey in space and will be shown to predict oscillations at the global scale consistent with previous results [3]. We map the master equation to a bosonic field theory [80, 81, 82, 83] to obtain a simple derivation of coupled Langevin equations for the fluctuations of predator-prey populations analogous to those obtained heuristically in Eq. 4.7. The path integral approach has the formal advantage of directly manipulating the population variables themselves rather than their probability distributions as in the master equation approach.

4.5 Master equation for predator-prey interactions

Consider a single, well-mixed patch of volume V. Species A is a predator for species B. We then have the following reactions:

$$\begin{aligned} B &\xrightarrow{b_1} BB \\ B &\xrightarrow{d_1} \emptyset \\ AB &\xrightarrow{p_1/V} A \\ AB &\xrightarrow{p_2/V} AA \\ A &\xrightarrow{d_2} \emptyset \end{aligned} \tag{4.21}$$

We give the rates of the two body reactions an inverse V dependence, which is interpreted as the volume scaling of the probability in a volume V that the two organisms will be close enough to interact. We also assume V is much greater than the size of an organism.

The above model is analogous to the first predator-prey model presented above in Eq. 4.2 and inspection verifies that it shares with Eq. 4.2 a serious defect: in the absence of predators, the prey population diverges to infinity (in mean field). To overcome this defect, there exist a variety of options to induce a finite "carrying capacity" for prey as in Eq. 4.5. Each option has advantages depending on the specifics of the predator-prey system being described, though many of the predictions end up being generic [84]. One option is to restrict the total patch population to some number N, including empty space (i.e. $N_A + N_B + N_E = N$). This is the "urn model" description [66]. In spatial models, N is often chosen to be 1, which is equivalent to a coarse graining scheme which takes a patch to be the space required for one organism. When $N > 1$ models are generalized to space, a patch is a locally well mixed area. Space is added as diffusion between such patches. In our model, we adopt the perspective that a patch is a well mixed region with many organisms, but do not constrain the population to a given N, choosing instead to obtain a finite carrying

capacity by allowing the death rate to increase with concentration. Equivalently, we could have simply included an intra-species competition reaction for the prey. The urn models are preferable in extremely crowded environments, but cause artifactual corrections to diffusion when the system is more dilute [66, 3]. The comparison between urn models and the present approach is analogous to the relationship between the nonlinear sigma model in field theory where the local spin is constrained to be exactly one and the ϕ^4 model of magnetism where the constraint on the local magnetism is soft. Our model is analogous to the ϕ^4 model in that it only includes a soft constraint on population in a small area. An advantage of the current approach is that it avoids nonlinear diffusive cross terms in spatial urn models that do not seem to change the dynamics substantially [3]. Additionally, urn models lead to complications in the interpretation of model parameters at the mean field level and in the master equation due to the fact that reaction rates in urn models must be combined with the joint probability for drawing the reactants from the urn prior to use in the master equation or mean field description [66]. With the soft constraint we apply, the reaction rates have similar, predictable meanings at every level of description from master equation to mean field.

Formally, we include the concentration dependence of the death rate by noting that $n_A = N_A/V$ is small

$$d_1(n_A) = d_1(0) + cn_A + O(n_A^2), \;\; c = d'(0) > 0 \tag{4.22}$$

Using the reactions defined in Eq. 4.21 as rules for transitions between different population states, we can now write a master equation for the patch

$$
\begin{aligned}
\partial_t P(m,n) = {} & d_1(-nP(m,n) + (n+1)P(m,n+1)) \\
& + c(-n^2 P(m,n) + (n+1)^2 P(m,n+1)) \\
& + b_1(-nP(m,n) + (n-1)P(m,n-1)) \\
& + p_1(-mnP(m,n) + (n+1)mP(m,n+1)) \\
& + p_2(-mnP(m,n) + (m-1)(n+1)P(m-1,n+1)) \\
& + d_2(-mP(m,n) + (m+1)P(m+1,n))
\end{aligned} \tag{4.23}
$$

Where m denotes the number of predators, and n denotes the number of prey. This master equation defines the time evolution of the probability distribution of population states.

4.5.1 Mapping to path integral formulation

To analyze the predator prey dynamics, we map Eq. 4.23 to a field theory. This is done using the standard Doi formalism to obtain a second quantized Hamiltonian [80] and bosonic coherent states to map the resulting theory to a path integral. For our approach and helpful reviews, see [85, 86]. The mapping is achieved by introducing the state vector

$$
|\psi\rangle = \sum_{m,n} P(m,n)|m,n\rangle \tag{4.24}
$$

and the operator pairs $a, \hat{a}$, $b, \hat{b}$ such that

$$\begin{aligned} a|m,n\rangle &= m|m-1,n\rangle \\ \hat{a}|m,n\rangle &= |m+1,n\rangle \\ [a,\hat{a}] &= 1 \\ b|m,n\rangle &= n|m,n-1\rangle \\ \hat{b}|m,n\rangle &= |m,n+1\rangle \\ \left[b,\hat{b}\right] &= 1 \end{aligned} \tag{4.25}$$

Finally, all other commutators are zero. We can then rewrite the dynamics given by the master equation (Eq. 4.23) as a Schroedinger like equation.

$$\partial_t|\psi\rangle = -\hat{H}(a,\hat{a},b,\hat{b})|\psi\rangle \tag{4.26}$$

We now specify the Hamiltonian (more accurately Liouvillian [83]) operator by multiplying the master equation by the state vector $|m,n\rangle$, summing over m and n, and applying the algebra of Eq. 4.26 to replace m and n by various combinations of the operators a, $\hat{a}$ and b, $\hat{b}$. From this algebra, working out the structure of the Hamiltonian is direct and simple. As an example, we work out the term corresponding to prey birth explicitly

$$\begin{aligned} & b_1 \sum_{m,n} (-nP(m,n) + (n-1)P(m,n-1))|m,n\rangle \\ & = b_1 \sum_{m,n} (-\hat{b}bP(m,n) + (n-1)P(m,n-1))|m,n\rangle \\ & = -b_1\hat{b}b|\psi\rangle + \sum_{m,n} nP(m,n)|m,n+1\rangle \\ & = -b_1\hat{b}b|\psi\rangle + b_1\hat{b}\hat{b}b|\psi\rangle \end{aligned} \tag{4.27}$$

Other terms are treated analogously. With normal ordering, this leads to the Hamiltonian

$$\begin{aligned}\hat{H} &= b_1(\hat{b}b - \hat{b}^2 b) + d_1(\hat{b}b - b) + \frac{c}{V}(\hat{b}^2 b^2 - \hat{b}b^2) \\ &+ \frac{p_1}{V}(\hat{a}a\hat{b}b - \hat{a}ab) + \frac{p_2}{V}(\hat{a}a\hat{b}b - \hat{a}^2 ab) \\ &+ d_2(\hat{a}a - a)\end{aligned} \tag{4.28}$$

Expectation values of functions of the random variables m and n are given by

$$\langle f \rangle = \langle 0, 0 | e^{a+b} f(\hat{a}, a, \hat{b}, b) e^{-H(\hat{a},a,\hat{b},b)t} | \psi(0) \rangle \tag{4.29}$$

Using bosonic coherent states, we write Eq. 4.29 as a path integral resulting in a Lagrangian description of the dynamics with generalization to space [81, 82]. Since we are interested in persistent oscillations around the only stable fixed point in the system, our choice of initial conditions is irrelevant and can be ignored. To link patches together for a spatial description, we define a lattice of patches and demand that each organism carry out a random walk on the lattice with given hopping probabilities for predator and prey. The continuum limit of a random walk is well known to be diffusion. We thus define diffusion rates D_1 and D_2 for predator and prey respectively and add diffusion operators to the Lagrangian. Careful manipulation of the field operators leads to the same results, provided the hopping probability for a species τ scales as $\tau \sim 1/a^2$ where a is the lattice constant taken to 0 in the continuum limit. Then $D = \lim_{a\to 0} a^2\tau$ (See appendix 2 for details). The resulting Lagrangian density is given by

$$\mathcal{L} = a^*\partial_t a + b^*\partial_t b - D_1 a^*\nabla^2 a - D_2 b^*\nabla^2 b$$
$$+H(\hat{b}, \hat{a}, b, a) \tag{4.30}$$

With fields derived from boson operators, the Lagrangian form of the master equation is difficult to interpret. This is because the field variables in the Lagrangian are not simply related to the physical variables of population number. This proves to be the source of difficulties in deriving correlation functions that are physically meaningful. To address this difficulty, we use a canonical Cole-Hopf transformation [87] to transform the field variables to number variables

$$a = ze^{-\hat{z}},\ \hat{a} = e^{\hat{z}} \tag{4.31}$$

$$b = \rho e^{-\hat{\rho}},\ \hat{b} = e^{\hat{\rho}} \tag{4.32}$$

This formulation has the advantage that z and ρ can be directly interpreted as the density variables for predator and prey respectively, while $\hat{\rho}$ and $\hat{z}$ generate noise terms. The transformed Lagrangian takes the form

$$\begin{aligned}\mathcal{L} = {} & \hat{z}\partial_t z + \hat{\rho}\partial_t \rho - D_1\hat{z}\nabla^2 z - D_1 z(\nabla\hat{z})^2 \\ & - D_2\rho(\nabla\hat{\rho})^2 - D_2\hat{\rho}\nabla^2\rho - b_1\rho(1-e^{\hat{\rho}}) \\ & + d_1\rho(1-e^{-\hat{\rho}}) + \frac{c}{V}\rho^2(1-e^{-\hat{\rho}}) \\ & + \frac{p_1}{V}z\rho(1-e^{-\hat{\rho}}) + \frac{p_2}{V}z\rho(1-e^{\hat{z}-\hat{\rho}}) \\ & + d_2 z(1-e^{-\hat{z}})\end{aligned} \tag{4.33}$$

In this form, the Lagrangian has diffusive noise, and difficult to handle exponential terms. In the following section, we exploit the small parameter $1/V$ to resolve these difficulties and analyze the theory.

4.6 Derivation of quasi-cycles from Large V expansion

To obtain predictions from the Lagrangian in Eq. 4.33, we can proceed directly by rewriting the fields as

$$\begin{aligned} \hat{z} \to \frac{\hat{z}}{\sqrt{V}}, \;\; \hat{\rho} \to \frac{\hat{\rho}}{\sqrt{V}} \\ z = V\varphi + \sqrt{V}\eta, \;\; \rho = V\phi + \sqrt{V}\xi \end{aligned} \tag{4.34}$$

and inserting them into the Lagrangian. These forms are intended to capture Gaussian fluctuations in the spirit of the traditional system size expansion of the master equation [73] while directly manipulating the population variables. The fields $\hat{z}$ and $\hat{\rho}$ have a mean field value of 0 due to conservation of probability [86]. This means that within the Gaussian approximation, the leading order term in those fields is a small correction of order $1/\sqrt{V}$ as above.

To derive the mean field theory and the fluctuations, we then insert the rhs forms of the fields in Eq. 4.34 into the Lagrangian Eq. 4.33 and retain only leading and next to leading order, resulting in an effective Lagrangian of the form

$$\mathcal{L} = \sqrt{V}\mathcal{L}_1 + \mathcal{L}_2 + O(1/\sqrt{V}) \tag{4.35}$$

Deriving each of these terms is straightforward. For purposes of illustration, we will carry out the expansion for the prey birth term explicitly

$$
\begin{aligned}
& b_1\rho(1-e^{\hat{\rho}}) \\
&= b_1(V\phi+\sqrt{V}\xi)(-\frac{\hat{\rho}}{V}-\frac{\hat{\rho}^2}{2V}) \\
&= b_1(-\sqrt{V}\hat{\rho}\phi-\frac{\hat{\rho}^2\phi}{2}-\hat{\rho}\eta)
\end{aligned} \tag{4.36}
$$

Carrying this out for each term in the Lagrangian and collecting terms yields at order $\sqrt{V}$

$$
\begin{aligned}
\mathcal{L}_1 = \hat{\rho}\partial_t\phi + \hat{z}\partial_t\varphi - D_1\hat{z}\nabla^2\varphi - D_2\hat{\rho}\nabla^2\phi \\
-b_1\phi\hat{\rho} + d_1\varphi\hat{\rho} + c\hat{\rho}\phi^2 + p_1\hat{\rho}\varphi\phi + p_2\hat{\rho}\phi\varphi \\
-p_2\hat{z}\phi\varphi + d_2\hat{z}\varphi
\end{aligned} \tag{4.37}
$$

Minimizing this term provides the mean field theory. For $V \to \infty$, this minimum is exact. The Euler-Lagrange equations are:

$$
\begin{aligned}
\frac{\delta\mathcal{L}_1}{\delta\hat{z}} &= \partial_t\varphi - D_1\nabla^2\varphi - p_2\phi\varphi + d_2\varphi = 0 \\
\frac{\delta\mathcal{L}_1}{\delta\hat{\rho}} &= \partial_t\phi - D_2\nabla^2\phi - b_1\phi + d_1\phi + c\phi^2 \\
&\quad + p_1\varphi\phi + p_2\phi\varphi = 0
\end{aligned} \tag{4.38}
$$

These are the standard Lotka-Volterra equations generalized to include space. They do not satisfy the criteria for pattern formation in predator-prey equations (reviewed in [88]). The long time dynamics relax to spatially uniform predator-prey populations with magnitudes given by the fixed points of the ordinary differential equations obtained by dropping the diffusion operator in Eqs. 4.38 above.

At next to leading order, we Fourier transform and switch to matrix notation, defining

$$\mathbf{x} = \begin{pmatrix} \eta \\ \xi \end{pmatrix}, \ \mathbf{y} = \begin{pmatrix} \hat{z} \\ \hat{\rho} \end{pmatrix} \tag{4.39}$$

By collecting terms as in Eq. 4.34 we can write down $\mathcal{L}_2$ as

$$\mathcal{L}_2 = i\omega \mathbf{y^T x} + \mathbf{y^T A x} - \frac{\mathbf{1}}{\mathbf{2}} \mathbf{y^T B y} \tag{4.40}$$

The matrices are given by

$$\mathbf{A} = \begin{pmatrix} D_1 k^2 & -p_2 \varphi \\ (p_1 + p_2)\phi & D_2 k^2 + c\phi \end{pmatrix} \tag{4.41}$$

and

$$\mathbf{B} = \begin{pmatrix} 2(d_2 + D_1 k^2)\varphi & -p_2 \varphi \phi \\ -p_2 \varphi \phi & 2(b_1 + D_2 k^2)\phi \end{pmatrix} \tag{4.42}$$

We now note that the vector $\mathbf{y}$ is a response field in the Martin Siggia Rose response function formalism for Langevin equations (reviewed in appendix 1) [89, 90]. Thus the fluctuations around mean field in the path integral are coupled Langevin equations. The resulting Langevin equations with the appropriate noise and correlations are

$$\begin{aligned} -i\omega \mathbf{x} &= \mathbf{A x} + \gamma(\omega) \\ \langle \gamma_i(\omega) \gamma_j(-\omega) \rangle &= B_{ij} \end{aligned} \tag{4.43}$$

These equations are of the same form as the equations reported in [2, 75] and are easily solved using simple linear algebra manipulations [75]

$$\begin{aligned} \mathbf{x} &= -(\mathbf{A} + \mathbf{i}\omega)^{-\mathbf{1}}\gamma(\omega) \equiv \mathbf{D}(\omega)^{-\mathbf{1}}\gamma(\omega) \\ \rightarrow x_1 = \eta &= -det(\mathbf{D})^{-\mathbf{1}}(\mathbf{D_{11}}\gamma_\mathbf{1} - \mathbf{D_{12}}\gamma_\mathbf{2}) \\ x_2 = \xi &= -det(\mathbf{D})^{-\mathbf{1}}(\mathbf{D_{21}}\gamma_\mathbf{1} - \mathbf{D_{22}}\gamma_\mathbf{2}) \end{aligned} \tag{4.44}$$

To obtain information from these solutions, we calculate the average power spectrum which captures oscillations but is free of phase cancellations [2]. The average power spectrum is obtained by taking the amplitude squared and averaging. For predator fluctuations this gives

$$\langle x_1 x_1^* \rangle = \frac{\alpha_k + \beta_k \omega^2}{(\omega^2 - \Omega_k^2)^2 + \Gamma_k^2 \omega^2} \tag{4.45}$$

with

$$\begin{aligned} \alpha_k &= B_{11}(k)A_{22}^2 + B_{22}(k)A_{12}^2 \\ \beta_k &= B_{11}(k) \\ \Omega_k^2 &= D_1 k^2(D_2 k^2 + c\phi) + p_2(p_1 + p_2)\phi\varphi > 0 \\ \Gamma &= -A_{11} - A_{22} \end{aligned} \tag{4.46}$$

The power spectrum contains a nontrivial peak in ω corresponding to the expected temporal oscillations. The peak in k is at 0 wavenumber as can be seen from the strictly

increasing functions of k present in the spectrum. This rules out spatial pattern formation. These results are in qualitative agreement with results from expansion of the master equation for analogous models with the hard constraints of urn models [2, 3]. Additional work will investigate the scaling of population fluctuations near extinction transitions and in disordered environments. These applications are of clear ecological interest and are difficult to study with system size expansions. However, they can be studied using well known methods from the functional integral formalism.

Chapter 5

The statistical mechanics of Turing patterns: The Levin-Segel model

The study of the emergent spatiotemporal patterns in physical or biological systems is an exciting and fruitful line of research in physics and in many other disciplines such as chemistry, ecology, animal biology, and neuroscience [91, 92, 16, 8, 93]. Examples include patterns on animal coats [94], engineered bacterial systems [95], chemical pattern formation [96], mussel population densities [97], and Rayleigh-Benard convection in fluids [98].

One particularly satisfying aspect of these studies is that insight into the origins of one kind of pattern often yields insight into the origins of patterns in entirely different systems. A key example is the Turing mechanism [16]. Turing's argument, which will be described in detail below, showed how diffusion, which is typically thought of as a randomizing influence, can give rise to spatial pattern formation when there are two or more classes of degrees of freedom (species) with "activator" and "inhibitor" dynamics. This mechanism has been proposed as an explanation for an enormous variety of systems including short (< 10m) length scale patchiness in planktonic ecosystems [99, 100, 101, 102], patterning in plant-resource systems [15], patchiness in insect abundance [103], stripe and spot patterns on the coats of animals [94], patterns in mussel beds [97] and even the geometric visual hallucinations experienced by shamans and users of hallucinogenic drugs [8, 104].

However, in spite of the seeming success of the Turing mechanism in explaining patterns across many disciplines, the partial differential equations representing the dynamics of systems with Turing patterns typically require unphysical fine tuning of parameters or separation of scales in the diffusivities of the different species in order to predict pattern formation [96, 99, 105, 12, 106, 13, 93]. The requirement that the system either have fine tuning

of kinetic parameters or a separation of scales in diffusivities in order to predict patterns, is unphysical for many applications and will be referred to below as the "fine tuning problem". To resolve the fine tuning problem for Turing patterns we show that a full statistical mechanical treatment of Turing patterns, where fluctuations due to the discrete nature of the degrees of freedom in the system – intrinsic noise – are included, the fine tuning problem is resolved [13].

It may seem counterintuitive to claim that including fluctuations resolves the fine tuning problem for Turing patterns because fluctuations are generally expected to *destabilize* ordered states such as spatial patterns. This is the rule in standard statistical mechanics [14] and many statistical mechanical models in ecology [107, 108]. However, exceptions exist in systems out of equilibrium. Careful experiments on Rayleigh-Benard convection have shown that fluctuations can drive the formation of convection rolls in fluid dynamics that would not form in the absence of fluctuations [109]. In ecology, recent theoretical work and careful data analysis have shown that the frequently observed cyclic population dynamics of predator-prey systems can be explained in many cases by fluctuation driven cycles in time [2, 3, 49, 110]. Similar phenomena have been predicted in evolutionary game theory as well as systems biology [111, 75]. Thus it seems possible that a full many body treatment of the Turing mechanism that incorporates the fluctuations, intrinsic noise, generated by the stochastic interaction of the constituents of the system will resolve the fine tuning problem.

The purpose of this chapter is to present a careful analysis of the Turing mechanism with intrinsic noise included. This chapter is an expansion and elaboration of our paper [13] which originally reported the resolution of the fine tuning problem of Turing instabilities through the incorporation of intrinsic noise. We will first review the Turing mechanism, and then present an extremely simple model of the Turing mechanism for planktonic predator-prey populations that we then analyze in detail. The results of the analysis show that in large regions of parameter space predicted by deterministic modeling to have only trivial spatial states a new kind of spatial pattern emerges that we call a "quasi-pattern." The

quasi-pattern state is analogous to intrinsic noise driven "quasi-cycles" recently discovered in the time domain [2]. Quasi-patterns are recognizable immediately as spatial patterns, but with a few important, experimentally relevant, differences from patterns predicted with deterministic analysis. The final sections of the chapter will focus on possible experimental tests and extensions of the theory developed in the body of the chapter. We focus on a model of planktonic predator-prey interactions throughout the paper for simplicity and also because predator-prey systems have been extensively analyzed theoretically [99, 105, 106, 100, 76] and there is beginning to be an experimental literature [103, 12]. However, we emphasize that the goal of this chapter is insight into the general interactions of intrinsic fluctuations with the Turing mechanism for pattern formation and that the results should be valid for most models of Turing instabilities. Evidence for this assertion is provided by the recent replication of our results on the Brusselator model of chemical pattern formation [112] and our own forthcoming results on pattern formation on the visual cortex [104].

5.1 The Turing mechanism

The Turing mechanism in its most basic form requires two different species that react and diffuse. One species, the "activator," diffuses relatively slowly, and catalyzes (activates) both its own production and the production of the second species. The second species, the "inhibitor," diffuses faster, and reduces (inhibits) the concentration of both the activator species and itself. These combined mechanisms lead to pattern formation from random initial conditions. We illustrate these steps with the example of predator-prey dynamics with random initial conditions

1. Random regions of activator (prey) with higher local concentrations reproduce rapidly, leading to dense clumps of activator species that then begin to diffuse.

2. Rapidly diffusing inhibitors (predators) are produced in the neighborhood of the high density autocatalyzing clumps of prey.

3. The predators inhibit the spread of the prey clumps through their production in the neighborhood of prey clumps. The autoinhibitory nature of predators prevents them from overwhelming the prey population.

These steps, summarized in fig. 5.1, show how activator-inhibitor dynamics can lead to spontaneous pattern formation [16]. As was noted above, formalizing this argument into standard partial differential equation models results in models that only exhibit Turing patterns if the predator (inhibitor) diffusivity is much larger than the prey (activator) diffusivity or the parameters are fine tuned [99, 105, 12, 106, 96]. Note that consistent with the existence of pattern forming systems which do not apparently display very large separation of diffusivities [103, 15] the qualitative argument made above for pattern formation does not depend on very large differences in diffusivities, nor on additional kinetic details.

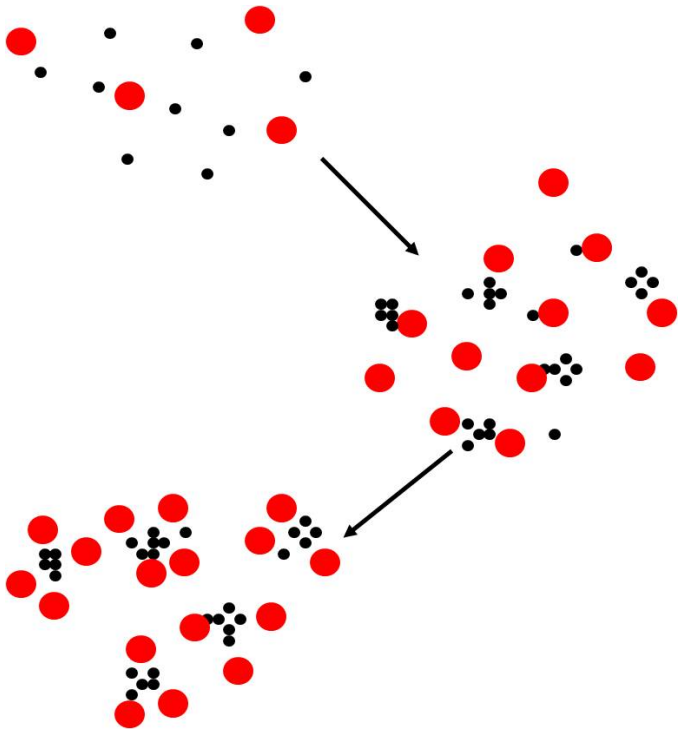

Figure 5.1: Illustration of the steps of the Turing mechanism as described in the text. The figure should be viewed from top to bottom. The prey (activators) are represented by black dots, and the predators (inhibitors) are represented by red dots.

5.2 Heuristic analysis of the Levin-Segel model

One of the simplest models of ecological pattern formation was originally introduced to model plankton-herbivore dynamics [99]. This model takes the form

$$\begin{aligned} \partial_t \psi &= \mu \nabla^2 \psi + b\psi + e\psi^2 - (p_1 + p_2)\psi\varphi \\ \partial_t \varphi &= \nu \nabla^2 \varphi + p_2 \varphi\psi - d\varphi^2 \end{aligned} \tag{5.1}$$

where the plankton population density is given by ψ, the herbivore population density is given by φ, b is birthrate for the plankton, p_1 and p_2 are predation, d is competition-driven death of the predators and e corresponds to an enhancement of the birthrate of plankton due when the concentration of plankton increases. In the original presentation of this model, this term was intended to be a proxy for reduced predator efficiency at higher prey concentrations [99]. It can also be interpreted as an Allee effect, wherein many species have enhanced reproduction at higher concentrations (for a review, see [113]). The parameters e and $p,\ d$, identify the prey as the activator and the predator as the inhibitor in the mechanism for pattern formation above and distinguish this model from the standard Lotka-Volterra based individual level models recently analyzed and demonstrated not to contain patterns in [3, 49].

Setting $p_1 = 0$ and $p_2 = p$, the model contains a stable homogeneous coexistence state when

$$p > e \text{ and } p^2 > de \tag{5.2}$$

with stationary fixed point populations given by

$$\psi_s = \frac{bd}{p^2 - de}, \quad \varphi_s = \frac{bp}{p^2 - de} \tag{5.3}$$

It contains a Turing instability if [99]

$$\frac{\nu}{\mu} > \left(\frac{1}{\left(\sqrt{p/d} - \sqrt{p/d - e/p} \right)} \right)^2 \tag{5.4}$$

When the model violates the stability conditions in Eq. 5.2, the plankton population diverges and a plankton regulation term (i.e. $-f\psi^3$) is required to make the model valid. Such a term would only materially affect the outcomes of this analysis near the instability, where it would decrease the set of parameters for which pattern formation occurs. To examine the behavior of the model, we take the generic set of $O(1)$ kinetic parameters $b = 1/2,\ e = 1/2,\ d = 1/2$ and $p = 1$. With these parameters Eq. 5.4 shows that non-generic diffusivities, $\nu/\mu > 27.8$, are required for pattern formation. Similar results are obtained for other generic parameter sets.

Demographic noise may change this picture [114] by inhibiting the decay of transient patterns leading to quasi-patterns. Turing instabilities occur when, for some specific set of wave vectors, small perturbations no longer decay. However, even when the parameters are tuned away from the Turing instability, perturbations with some wavelengths may decay more slowly than others, leading to transient patterns. Demographic noise maintains these patterns by generating continual perturbations. This is reminiscent of extrinsic noise driven patterns reported in other contexts [115, 116, 117].

To quantify this heuristic argument, we look at the Fourier transformed dynamics of the fluctuations from the coexistence fixed point with added white noise ξ, variance 1. These dynamics are given by

$$-i\omega \boldsymbol{x} = \boldsymbol{A}\boldsymbol{x} + \xi \tag{5.5}$$

The matrix $\boldsymbol{A}$ is the Fourier transformed stability matrix and $\boldsymbol{x}$ is the vector of deviations

from equilibrium of predator and prey populations respectively,

$$\boldsymbol{A} = \begin{pmatrix} -\nu k^2 - p\psi_s & p\varphi_s \\ -p\psi_s & -\mu k^2 + e\psi_s \end{pmatrix} \tag{5.6}$$

Simple manipulations yield the average power spectrum

$$\begin{aligned} P(k,\omega) = \left[p^2\varphi_s^2 + (e\psi_s - \mu k^2)^2\right] \times \Bigg[\big(pb\psi_s + \mu\nu k^4 - \omega^2 \\ - \psi_s k^2 e\nu \left(1 - \frac{p\mu}{e\nu}\right)\big)^2 + \omega^2((e-p)\psi_s - (\mu+\nu)k^2)\Bigg]^{-2} \end{aligned} \tag{5.7}$$

The numerator is proportional to the variance of the noise, which is in this case one. Very approximately, Eq. 5.7 predicts that patterns (indicated by peaks in the power spectrum) form whenever $e\nu > p\mu$, and that without noise and away from a classical Turing instability the power spectrum is zero. The condition $e\nu > p\mu$ is much less stringent than Eq. 5.4 and can be satisfied for generic sets of parameters. However, to reliably demonstrate our hypotheses and extract precise experimental predictions, we must perform a much more systematic calculation. Carrying out such a calculation is the goal of the next section.

5.3 Predator-prey model with intrinsic noise

To systematically include the effects of intrinsic noise requires a model defined at the level of individual organisms, since intrinsic noise is generated by the stochastic nature of individual birth and death events as well as the stochastic interactions between individual organisms. Such a description of the dynamics at the individual level is called an individual level model (ILM). One simple way to define an ILM is to specify the reactions that can take place in a well mixed patch of volume V. With parameters to specify the relative rates of the reactions, a model of individual level interactions that incorporates intrinsic noise is fully specified.

For an ILM version of the Levin-Segel model we consider the following reactions

$$\begin{aligned} P &\xrightarrow{b} PP \\ PP &\xrightarrow{e/V} PPP \\ PH &\xrightarrow{p_1/V} H \\ PH &\xrightarrow{p_2/V} HH \\ HH &\xrightarrow{d/V} H \end{aligned} \tag{5.8}$$

where P denotes plankton and H denotes herbivores, with the parameters as described above. Stochastic trajectories of H and P, enumerated by m and n respectively, are described by the master equation

$$\begin{aligned} \partial_t P(m,n) &= b(-nP(m,n) + (n-1)P(m,n-1)) \\ &+ \frac{e}{V}((n-1)(n-2)P(m,n-1) - n(n-1)P(m,n)) \\ &+ \frac{p_1}{V}(-mnP(m,n) + (m)(n+1)P(m,n+1)) \\ &+ \frac{p_2}{V}(-mnP(m,n) + (m-1)(n+1)P(m-1,n+1)) \\ &+ \frac{d}{V}\left[(m+1)mP(m+1,n) - m(m-1)P(m,n)\right] \end{aligned} \tag{5.9}$$

The master equation, which is exactly equivalent to the specification of the model as a collection of reactions in Eq. 5.8, can then be used to analyze the ILM version of the Levin-Segel model by applying techniques from non-equilibrium statistical mechanics.

5.3.1 Field theory representation of the model

While several options exist for analysis of the master equation, such as direct expansion of the master equation [73], we analyze the master equation by a mapping to field theory, because it is convenient for handling spatially extended systems. To analyze the master

equation using the techniques of field theory, we introduce the operators

$$\begin{aligned} a|m,n\rangle &= m|m-1,n\rangle \\ \hat{a}|m,n\rangle &= |m+1,n\rangle \\ [a,\hat{a}] &= 1 \\ b|m,n\rangle &= n|m,n-1\rangle \\ \hat{b}|m,n\rangle &= |m,n+1\rangle \\ \left[b,\hat{b}\right] &= 1 \end{aligned} \tag{5.10}$$

and the state $|\psi\rangle = \sum P(n)|n\rangle$. These definitions allow the master equation to be mapped to a bosonic field theory [80, 83, 81, 82, 87]. As an explicit example of how to convert the master equation to a field theory, consider the master equation corresponding to the second reaction in Eq. 5.8 alone.

$$\partial_t P(n) = \frac{e}{V}[(n-1)(n-2)P(n-1) - n(n-1)P(n)] \tag{5.11}$$

Ignoring V for now, we multiply both sides by $|n\rangle$ and sum over n

$$\sum_n \partial_t P(n)|n\rangle = e\sum_n \left[(n-1)(n-2)P(n-1) - n(n-1)P(n)\right]|n\rangle \tag{5.12}$$

We next shift the sums, and manipulate the first term in the sum

Let $n' = n - 1 \rightarrow n = n' + 1$.

$$
\begin{aligned}
& e \sum_{n'} n'(n'-1) P(n') |n'+1\rangle \ , n' \rightarrow n \\
&= e \sum n(n-1) P(n) \hat{b} |n\rangle \\
&= e \hat{b} \sum (n-1) \hat{b} b P(n) |n\rangle \\
&= e \hat{b} \hat{b} \sum (n-1) P(n) b |n\rangle \\
&= e \hat{b} \hat{b} \sum \hat{b} b b P(n) |n\rangle \\
&= e \hat{b}^3 b^2 \sum P(n) |n\rangle \\
&= e \hat{b}^3 b^2 |\psi\rangle
\end{aligned} \tag{5.13}
$$

We now work out the second term in the sum

$$
\begin{aligned}
& e \sum_{n} n(n-1) P(n) |n\rangle \\
& \qquad = e \hat{b} \hat{b} b b |\psi\rangle
\end{aligned} \tag{5.14}
$$

This yields

$$
\partial_t |\psi\rangle = e \left[\hat{b}^3 - \hat{b}^2 \right] b^2 |\psi\rangle \tag{5.15}
$$

Similar analyses lead to second quantized forms for the rest of the master equation. We can now assemble the entire Hamiltonian. We start by writing the master equation in second quantized form

$$
\partial_t |\psi\rangle = \left[b(\hat{b}^2 - \hat{b}) b + \frac{e}{V} (\hat{b}^3 - \hat{b}^2) b^2 + \frac{p_1}{V} (\hat{a} a b - \hat{a} a \hat{b} b) + \frac{p_2}{V} (\hat{a}^2 a b - \hat{a} a \hat{b} b) + \frac{d}{V} (1 - \hat{a}) \hat{a} a^2 \right] |\psi\rangle \tag{5.16}
$$

Since the standard definition of the Hamiltonian is

$$\partial_t|\psi\rangle = -\hat{H}|\psi\rangle \tag{5.17}$$

we have

$$\hat{H} = -\left[b(\hat{b}^2 - \hat{b})b + \frac{e}{V}(\hat{b}^3 - \hat{b}^2)b^2 + \frac{p_1}{V}(\hat{a}ab - \hat{a}a\hat{b}b) + \frac{p_2}{V}(\hat{a}^2ab - \hat{a}a\hat{b}b) + \frac{d}{V}(1-\hat{a})\hat{a}a^2\right] \tag{5.18}$$

According to the standard mapping between Hamiltonians represented by bosonic operators we can easily derive the Lagrangian, generalized to space. The boundary conditions are ignored, because the focus of this paper is the long time limit and there is only one attractor in the system. This gives a Lagrangian in the form

$$\begin{aligned}\mathcal{L} = a^*\partial_t a + b^*\partial_t b - \nu a^*\nabla^2 a - \mu b^*\nabla^2 b \\ + H(\hat{b}, \hat{a}, b, a)\end{aligned} \tag{5.19}$$

To transform to more physical variables, the standard Cole-Hopf transformation can be applied to transform the field variables to direct number and noise representations. This transformation is given by

$$\begin{aligned} a &= ze^{-\hat{z}} \\ \hat{a} &= e^{\hat{z}} \end{aligned} \tag{5.20}$$

$$\begin{aligned} b &= \rho e^{-\hat{\rho}} \\ \hat{b} &= e^{\hat{\rho}} \end{aligned} \tag{5.21}$$

the new field variables z and ρ can be heuristically interpreted as the number of predator and prey respectively (the precise interpretation is that their expectation values correspond, i.e. $\langle f(\rho, z)\rangle = \langle f(N_P, N_H)\rangle$) and the auxiliary fields denoted by carets generate the intrinsic

noise, as will be seen below by showing that the minimum of the action, which corresponds to mean field theory is at $\hat{\rho} = \hat{z} = 0$. The Lagrangian in the new variables is

$$\begin{aligned}\mathcal{L} = \hat{x}\partial_t z + \hat{\rho}\partial_t \rho - \nu\hat{z}\nabla^2 z - \mu\hat{\rho}\nabla^2\rho - \nu z(\nabla\hat{z})^2 - \mu\rho(\nabla\hat{\rho})^2 + b\rho(1-e^{\hat{\rho}}) \\ + \frac{e}{V}\rho^2(1-e^{\hat{\rho}}) + \frac{p_1}{V}z\rho(1-e^{-\hat{\rho}}) + \frac{c}{V}z\rho(1-e^{\hat{z}-\hat{\rho}}) + \frac{d}{V}z^2(1-e^{-\hat{z}})\end{aligned} \tag{5.22}$$

5.3.2 System size expansion

We now can carry out the system size expansion. We expand the fields as

$$\begin{aligned}\hat{z} \to \frac{\hat{z}}{\sqrt{V}}, \quad \hat{\rho} \to \frac{\hat{\rho}}{\sqrt{V}} \\ z = V\varphi + \sqrt{V}\eta, \quad \rho = V\psi + \sqrt{V}\xi\end{aligned} \tag{5.23}$$

To perform this expansion to consistent order, it is necessary to expand the exponentials out to second order. This is because the expansion will promote second order terms to first order. The result will be an expansion of the Lagrangian in the form

$$\mathcal{L} = \sqrt{V}\mathcal{L}_1 + \mathcal{L}_2 + O(1/\sqrt{V}) \tag{5.24}$$

We once again carry out the expansion explicitly for the term coupled by e/V. The other terms are worked out in appendix three, except those worked out in the previous chapter, or in [49]

$$\begin{aligned}&\frac{e}{V}\rho^2(1-e^{\hat{\rho}}) \\ &= \frac{e}{V}\left(V\psi + \sqrt{V}\xi\right)\left(V\psi + \sqrt{V}\xi\right)\left(1 - \left(1 + \frac{\hat{\rho}}{\sqrt{V}} + \frac{\hat{\rho}^2}{2V}\right)\right) \\ &= -e\left(\psi + \frac{\xi}{\sqrt{V}}\right)\left(\sqrt{V}\psi\hat{\rho} + \frac{\psi\hat{\rho}^2}{2} + \xi\hat{\rho}\right) \\ &= -e\left(\sqrt{V}\psi^2\hat{\rho} + \frac{\psi^2\hat{\rho}^2}{2} + 2\xi\hat{\rho}\right)\end{aligned} \tag{5.25}$$

Our next task is to gather terms of the same order in V to assemble the expansion and the resulting MFT and SDE. At leading order, $\sqrt{V}$, we have

$$\mathcal{L}_1 = \hat{\rho}\partial_t\psi + \hat{z}\partial_t\varphi - \nu\hat{z}\nabla^2\varphi - \mu\hat{\rho}\nabla^2\psi - b\psi\hat{\rho} - e\psi^2\hat{\rho} + b\varphi\psi\hat{\rho} - c\varphi\psi(\hat{z}-\hat{\rho}) + d\varphi^2\hat{z} \quad (5.26)$$

It is trivial to extract the mean field PDE's by using the Euler-Lagrange equations. The equations that result are

$$\frac{\delta\mathcal{L}_1}{\delta\hat{z}} = \partial_t\psi - \mu\nabla^2\psi + b\psi + e\psi^2 - (p_1+p_2)\psi\varphi = 0 \quad (5.27)$$

which is the first of the equations for the Levin-Segel model. The second equation is

$$\frac{\delta\mathcal{L}_1}{\delta\hat{\rho}} = \partial_t\varphi - \nu\nabla^2\varphi + p_2\varphi\psi - d\varphi^2 = 0 \quad (5.28)$$

again reproducing the Levin-Segel model equation of motion. Note that the auxiliary fields have zero expectation value at mean field, which confirms the interpretation that they correspond to noise. Now we wish to assemble the terms in the next to leading order terms in the Lagrangian. The terms in $\mathcal{L}_2$ that are linear in η or ξ correspond to the stability matrix of the MFT. We'll then consider the noise terms, which are of the form $\hat{\rho}^2$, $\hat{z}^2$ and $\hat{\rho}\hat{z}$

Proceeding, we have

$$\begin{aligned}&\hat{z}\partial_t\eta + \hat{\rho}\partial_t\xi - \hat{z}\nu\nabla^2\eta - \hat{\rho}\mu\nabla^2\xi \\ &+ p_1\eta\psi\hat{\rho} - p_2\eta\psi(\hat{z}-\hat{\rho}) + 2d\eta\varphi\hat{z} + b\xi\hat{\rho} + 2e\xi\psi\hat{\rho} + b\xi\varphi\hat{\rho} - p_2\xi\varphi(\hat{z}-\hat{\rho}) \end{aligned} \quad (5.29)$$

We convert this into a Fourier transformed matrix form that includes time and space

$$\mathcal{L}_2 = \boldsymbol{y}^T\partial_t\boldsymbol{x} - \boldsymbol{y}^T\boldsymbol{A}\boldsymbol{x} - \frac{1}{2}\boldsymbol{y}^T\boldsymbol{B}\boldsymbol{y} \quad (5.30)$$

with vectors given by

$$\boldsymbol{x} = \begin{pmatrix} \eta \\ \xi \end{pmatrix}, \ \boldsymbol{y} = \begin{pmatrix} \hat{z} \\ \hat{\rho} \end{pmatrix} \tag{5.31}$$

so that the predator variables are on top. The matrix A is the Jacobian J of the MFT with space added and is given by

$$\boldsymbol{A} = \begin{pmatrix} -\nu k^2 + p_2\psi - 2d\varphi & p_2\varphi \\ -(p_1+p_2)\psi & -\mu k^2 + b + 2e\psi - (p_1+p_2)\varphi \end{pmatrix} \tag{5.32}$$

The matrix for the correlations of the noise is given by

$$\boldsymbol{B} = \begin{pmatrix} d\varphi^2 + p_2\varphi\psi + \nu\varphi k^2 & -p_2\varphi\psi \\ -p_2\varphi\psi & b\psi + e\psi^2 + b\varphi\psi + p_2\varphi\psi + \mu\psi k^2 \end{pmatrix} \tag{5.33}$$

We also now note that $\mathcal{L}_2$ is in the form of a Lagrangian in the Martin-Siggia-Rose (MSR) response function formalism (summarized in appendix 1) for Langevin equations [89, 90].

5.3.3 The power spectrum

We now extract the stochastic differential equations (SDE) that govern the dynamics of the fluctuations, and calculate the power spectrum of the fluctuations. The Langevin equations from the response function formalism are

$$\begin{aligned} i\omega\boldsymbol{x} &= \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\gamma}(\omega) \\ \langle \gamma_i(\omega)\gamma_j(-\omega)\rangle &= B_{ij} \end{aligned} \tag{5.34}$$

We solve formally to obtain

$$\boldsymbol{x} = (\boldsymbol{A} + i\omega)^{-1}\boldsymbol{\gamma}(\omega) \equiv \boldsymbol{D}(\omega)^{-1}\boldsymbol{\gamma}(\omega) \tag{5.35}$$

The power spectrum is

$$\begin{aligned} x_1 &= \frac{1}{det(D)}\left(D_{22}\gamma_1 - D_{12}\gamma_2\right) \\ \langle x_1 x_1^* \rangle &= \frac{1}{|det(D)|^2}\langle (D_{22}\gamma_1 - D_{12}\gamma_2)(D_{22}^*\gamma_1 - D_{12}^*\gamma_2)\rangle \\ &= \frac{1}{|det(D)|^2}\left(|D_{22}|^2 B_{11} - 2D_{12}Re(D_{22})B_{21} + |D_{12}|^2 B_{22}\right) \end{aligned} \tag{5.36}$$

To find the phase diagram, take $p_1 \to 0$, $p_2 = p$. This simplification does not substantially change the dynamics of the model. In terms of elements of the stability matrix, the denominator of the power spectrum is

$$\begin{aligned} \det(D) &= (J_{11} + i\omega - \nu k^2)(J_{22} + i\omega - \mu k^2) - J_{12}J_{21} \\ &= J_{11}J_{22} - J_{12}J_{21} + i\omega J_{11} - J_{11}\mu k^2 + i\omega J_{22} - \omega^2 - i\omega\mu k^2 - \nu k^2 J_{22} - i\omega\nu k^2 + \mu\nu k^4 \\ &= det(J) + i\omega(Tr(J) - (\mu+\nu)k^2) - (J_{11}\mu + J_{22}\nu)k^2 + \mu\nu k^4 - \omega^2 \end{aligned} \tag{5.37}$$

The full expression for the power spectrum is

$$P(k,\omega) = \frac{|D_{22}|^2 B_{11} - 2D_{12}Re(D_{22})B_{21} + |D_{12}|^2 B_{22}}{(det(J) + \mu\nu k^4 - \omega^2 - (J_{11}\mu k^2 + J_{22}\nu k^2))^2 + \omega^2(Tr(J) - (\mu+\nu)k^2)^2} \tag{5.38}$$

Note that $Det(J) > 0$ and $Tr(J) < 0$ because the ODE has stable spirals in the regime with stable coexistence [67]. For $k = 0$, if variation in the numerator is neglected, there is a peak in ω.

Recall the fixed point values at coexistence are

$$\varphi = \frac{pb}{p^2 - de}$$
$$\psi = \frac{db}{p^2 - de} \tag{5.39}$$

Using the fixed point values, the matrix $\boldsymbol{A}$ can be further simplified to

$$\boldsymbol{A} = \begin{pmatrix} -\nu k^2 - p\psi & p\varphi \\ -p\psi & -\mu k^2 + e\psi \end{pmatrix} \tag{5.40}$$

Now we evaluate the determinant of the ODE stability matrix (J above, and equal to $\boldsymbol{A}$ with space removed) and the trace

$$\det(J) = p\psi b \tag{5.41}$$

The trace is

$$Tr(J) = (e - p)\psi \tag{5.42}$$

Simplifying the denominator in Eq 5.38 yields

$$\begin{aligned} |det(D)|^2 &= (det(J) + \mu\nu k^4 - \omega^2 - (J_{11}\mu k^2 + J_{22}\nu k^2))^2 + \omega^2(Tr(J) - (\mu + \nu)k^2)^2 \\ &= (pb\psi + \mu\nu k^4 - \omega^2 - \psi(-p\mu k^2 + e\nu k^2))^2 + \omega^2((e - p)\psi - (\mu + \nu)k^2)^2 \\ &= \left(pb\psi + \mu\nu k^4 - \omega^2 - \psi k^2 e\nu \left(1 - \frac{p\mu}{e\nu}\right)\right)^2 + \omega^2((e - p)\psi - (\mu + \nu)k^2)^2 \end{aligned} \tag{5.43}$$

The form of the denominator is $(A - Bk^2 + Ck^4)^2$, which has a minimum at non zero k. This minimum corresponds to an emergent length scale, and is the first indication of

pattern formation. Systematic demonstration of the emergence of pattern formation requires accounting for the k dependence in the numerator. The noise matrix $\boldsymbol{B}$ can be simplified to

$$\boldsymbol{B} = \begin{pmatrix} 2p\varphi\psi + \nu\varphi k^2 & -p\varphi\psi \\ -p\varphi\psi & 2p\varphi\psi + \mu\psi k^2 \end{pmatrix} \tag{5.44}$$

Notice the symmetry in the noise correlations. We now can expand out the numerator of Eq. 5.38

$$\begin{aligned} &|D_{22}|^2 B_{11} - 2D_{12}Re(D_{22})B_{21} + |D_{12}|^2 B_{22} \\ &= |e\psi - \mu k^2 + i\omega|^2(2p\varphi\psi + \nu\varphi k^2) + 2p\varphi(e\psi - \mu k^2)(p\varphi\psi) + p^2\varphi^2(2p\varphi\psi + \mu\psi k^2) \\ &= (e\psi - \mu k^2)^2(2p\varphi\psi + \nu\varphi k^2) + \omega^2(2p\varphi\psi \\ &+ \nu\varphi k^2) + 2p^2\varphi^2\psi(e\psi - \mu k^2) + p^2\varphi^2(2p\varphi\psi + \mu\psi k^2) \end{aligned} \tag{5.45}$$

This gives the final form of the power spectrum

$$P(k,\omega) = \frac{(e\psi - \mu k^2)^2(2p\varphi\psi + \nu\varphi k^2) + \omega^2(2p\varphi\psi + \nu\varphi k^2) + 2p^2\varphi^2\psi(e\psi - \mu k^2) + p^2\varphi^2(2p\varphi\psi + \mu\psi k^2)}{\left(pb\psi + \mu\nu k^4 - \omega^2 - \psi k^2 e\nu\left(1 - \frac{p\mu}{e\nu}\right)\right)^2 + \omega^2((e-p)\psi - (\mu+\nu)k^2)^2} \tag{5.46}$$

5.4 Analysis of the power spectrum

5.4.1 Phase diagram for quasi-patterns

The expression for the power spectrum in Eq. 5.46 is not very illuminating, and does not simplify a great deal. To find quasi-patterns we note that the highest power of k in the denominator of Eq. 5.46 is larger than the highest power in the numerator. That means for sufficiently large k, the power spectrum is always decreasing. Thus, to show the existence of

a maximum, it is sufficient to show that for small k, the power spectrum is increasing. This can be shown by computing $\frac{dP}{dk^2}$ and evaluating at $k^2 = 0$. When this expression is greater than 0, there is pattern formation. This yields the analytical criterion

$$\frac{\nu}{\mu} > \frac{p^3(5p^2 + 7de)}{e(4p^4 + 5p^2de + 3d^2e^2)} \tag{5.47}$$

From this criterion, we can detect when there will be spatial pattern formation. It is much less stringent than the criterion for Turing instabilities. The conditions for a Turing instability are

$$\frac{\nu}{\mu} > \left(\frac{1}{\sqrt{p/d} - \sqrt{p/d - e/p}} \right)^2 \tag{5.48}$$

For the generic parameters $b = 1/2,\ p = 1,\ d = 1/2,\ e = 1/2$ the criterion 5.47 yields $\nu/\mu > 2.48$, while the Turing condition yields $\nu/\mu > 27.8$. This behavior is typical of generic parameters. The phase diagram of the system bears out this conclusion as shown in figure 5.2.

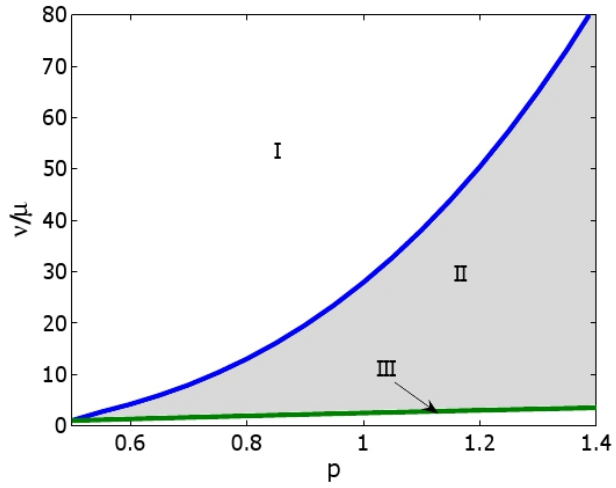

Figure 5.2: Phase diagram over stable parameter region in p/d. Region I is MFT level pattern formation, the shaded region II is the portion of the phase diagram where fluctuation driven quasi-patterns are present, region III is a spatially homogeneous phase.

An additional feature of the model is that oscillations and spatial pattern formation are essentially decoupled. This means that the model predicts global population oscillations and spatial pattern formation, but not traveling waves. The mathematical origin of this can be seen in Eq. 5.7. The k^2 term with a negative coefficient at $\omega = 0$ is quickly overwhelmed by the positive k^2 dependence of the ω^2 term as the frequency begins to grow. In the power spectrum (fig. 5.3) this can be seen as the deep valley between the peaks in k and ω. This interpretation is supported by preliminary simulations.

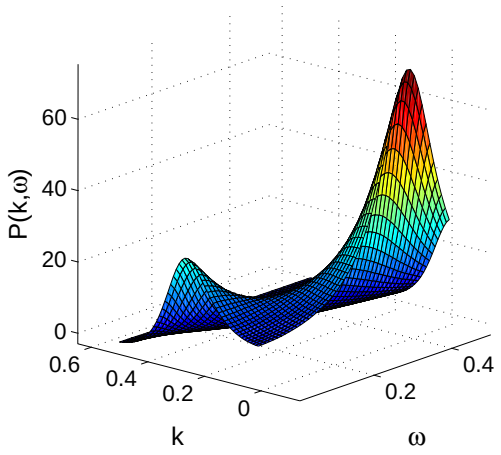

Figure 5.3: Power spectrum with p=1, ν/μ=15

5.4.2 Wavelength of fluctuation driven patterns

To a fairly good approximation, the wavelength of the Turing quasi-patterns can be calculated. The wavelength corresponds to the wave vector that maximizes the power spectrum. To calculate that value, consider the numerator of the power spectrum only at $\omega = 0$.

$$\left(pb\psi + \mu\nu k^4 - \psi k^2 e\nu \left(1 - \frac{p\mu}{e\nu}\right)\right)^2)^2 \tag{5.49}$$

The minimum of this expression will correspond with reasonable accuracy to the real

wavelength and can be obtained through straightforward calculation to be

$$\lambda_m = \frac{2\pi}{k_m} = \sqrt{\frac{2\mu}{\psi}\left(1 - \frac{c\mu}{e\nu}\right)} \tag{5.50}$$

This shows that for a fixed ratio of diffusivities, the wavelength increases as the square root of the diffusivity. In addition, while the phase diagram of the system (fig 5.2) and therefore the presence of Turing quasi-patterns depends on diffusivity only through the ratio ν/μ, the wavelength of the patterns depends on the values of the diffusivities.

This calculation also implies that the wavelength of the quasi-patterns is closely related to the wavelength of patterns in the region of the phase diagram where patterns are generated at mean field. In the standard theory of Turing patterns, patterns are formed when the homogeneous steady state is unstable to perturbations with a specific set of wave vectors k. The wavelength is then the wavelength corresponding to the mode that is most unstable. In the calculation above, we have picked out the mode that in mean field theory corresponds to the slowest decaying mode as the wavelength of the quasi-patterns. This is because the denominator of the power spectrum is equal to the product of the eigenvalues of the stability matrix squared. This product is smallest for the slowest decaying mode, which is also the mode that will go unstable in mean field theory first as parameters are varied. Therefore the wavelength of the quasi-patterns corresponds to the wavelength of the mean field patterns.

5.4.3 Period of quasi-cycles

A similar calculation to the calculation above for the wavelength of the quasi-patterns can be carried out for the period of the quasi-cycles. Consider the denominator of the power spectrum with $k = 0$

$$\left(pb\psi - \omega^2\right)^2 + \omega^2((e-p)\psi)^2 \tag{5.51}$$

Analogous to the wavelength calculation, we seek the minimum in ω. Simple calculation

yields a period of

$$T = \frac{2\pi}{\omega_m} = \frac{4\pi}{\sqrt{2bp\psi - (e-p)^2\psi^2}} \tag{5.52}$$

Similar arguments to those for the wavelength indicate that the period for the quasi-cycles is approximately the period for the stable spirals present in mean field theory [2].

5.5 Distinguishing quasi-patterns and quasi-cycles from other spatiotemporal patterns

To distinguish spatiotemporal patterns generated by intrinsic noise from those generated by feedbacks alone (i.e. mean field patterns) or by extrinsic noise, it is necessary to develop theoretical predictions that differ for each of these cases. Previous work has focused primarily on time series data, focusing on problems such as distinguishing quasi-cycles from limit cycles [110] as well as the simpler task of simply determining the amount of extrinsic versus intrinsic noise in ecosystems [65]. This work has confirmed that both extrinsic noise and intrinsic noise are important in real ecosystems for populations such as temperate songbirds in Norway, and the beetle species Tribolium [65, 118, 119] and that quasi-cycles are present in real ecological time series data[110]. The work also confirms that the importance of intrinsic noise decreases as population density increases, in line with the expectation that the scale of intrinsic noise depends on the scale of the population density [119].

While separate signatures of quasi-cycles and quasi-patterns will be discussed below, one common feature that distinguishes quasi-cycles and quasi-patterns from their counterparts in mean field theory is that they depend on the concentration of the population being studied. To leading order only the fluctuations have patterns, implying that the local populations can be written as mean value plus fluctuations scaled by the size of a locally well mixed region (see below). Thus the amplitude of the patterns relative to the mean population size of the fluctuation driven patterns will change as the size of a locally well mixed area changes,

while the relative amplitude of mean field patterns and limit cycles would not change. Such a variation of the size of a locally well mixed area could presumably be used to detect quasi-patterns and quasi-cycles.

5.5.1 Distinguishing quasi-cycles from limit cycles

Given a population that has oscillatory abundance in time, theory indicates that the oscillations can come from either quasi-cycles driven by noise or from population density dependent feedbacks alone, perturbed by noise (mean field cycles). The key difference mathematically is that the power spectrum of limit cycles has a pole at its frequency while the power spectrum of quasi-cycles does not. In the time domain, this means that the cycles driven by intrinsic noise have a short correlation time while limit cycles have an infinite correlation time. Since poles do not exist in real population data due to stochasticity, finite size populations, measurement error, etc. what this means for real data is that there is a separation of scales between the correlation time of limit cycles and quasi-cycles. This was first pointed out in detail by Pineda-Krch et al. [110]. These authors also showed that wolverine population cycles are likely quasi-cycles, while the celebrated lynx-hare cycles from the Hudson Bay company's trapping records are most likely limit cycles [110].

Other studies of the role of intrinsic noise have focused on intrinsic noise contributions compared to extrinsic noise contributions as a function of local population size [65, 119]. In frequency space, the best frequencies to analyze to distinguish the relative importance of noise are high frequencies, corresponding to the short timescale fluctuations of the system. To extract predictions for the case of intrinsic noise, we look at the large ω asymptotics of the power spectrum Eq. 5.46 at $k = 0$

$$P(k = 0, \omega) = \frac{2p\psi\varphi}{\omega^2}, \;\; \omega \gg \omega_m \tag{5.53}$$

where ω_m is the modal frequency of the quasi-cycles. For cycles driven by extrinsic additive

noise, we look at the same asymptotics for the power spectrum from the heuristic calculation, which, as we noted above, can be considered as a calculation for extrinsic noise. In this case, the asymptotic form is

$$P(k=0,\omega) = \frac{p^2\psi^2 + e^2\varphi^2}{\omega^4}\langle\xi\xi\rangle, \quad \omega \gg \omega_m \tag{5.54}$$

where the variance $\langle\xi\xi\rangle$ is independent of population density and ω_m is the frequency of the quasi-cycles. While in this case, both the expressions depend on the square of population density, the decay in ω has a power of two for intrinsic noise, and of four in the case of extrinsic noise. Thus the tails can be easily distinguished in real data.

5.5.2 Distinguishing quasi-patterns from mean field patterns

Similar considerations can be applied to quasi-patterns. While further study is needed, quasi-patterns should generically have a shorter correlation length than mean field patterns and techniques similar to those outlined above for distinguishing mean field patterns from quasi-patterns and applied to real populations in [110] should translate directly.

For distinguishing extrinsic and intrinsic noise contributions, the asymptotics for short wavelength fluctuations can again be derived for the intrinsic and extrinsic noise cases. For intrinsic noise, we have

$$P(k,\omega=0) = k^{-2}\frac{\varphi}{\nu}, \quad k \gg k_m \tag{5.55}$$

where k_m is the wave vector of the mode of the power spectrum. For extrinsic noise, we have

$$P(k,\omega=0) = \frac{k^{-4}}{\nu^2}\langle\xi\xi\rangle, \quad k \gg k_m \tag{5.56}$$

where the variance $\langle\xi\xi\rangle$ is independent of population density. Like the quasi-cycle case, the scaling in k differs by a power of two between the extrinsic and intrinsic noise cases. Contrary to quasi-cycles in the previous section, the extrinsic and intrinsic noise lead to different

powers of population density for large k. This provides a useful tool for distinguishing between the effects of extrinsic and intrinsic noise on the formation of patterns especially if the density of the populations can be varied through comparative study of field data in different ecosystems, or through experiments. These considerations are quite broad, and should qualitatively apply to other systems, such as chemical reaction systems in which quasi-patterns or cycles may be present, such as the Brusselator model of chemical pattern formation [112].

One major difficulty with the arguments above for distinguishing extrinsic and intrinsic noise effects is that in ecology, extrinsic noise often is not additive as assumed above, but is manifested through stochasticity in the kinetic parameters of the system [11]. How such noise affects the spectrum of quasi-patterns and quasi-cycles is not fully settled and requires further research. Another potential difficulty is that numerical study of quasi-patterns indicates that the form of the power spectrum used above is only strictly valid near the onset of the patterns due to higher order corrections to the mean population. Evidence for this claim is discussed below. The implications of this possibility for distinguishing different kinds of patterns is a subject for future research.

5.6 Thermodynamic limit

To compare to data, we also must be able to estimate the conditions under which the fluctuation driven effects described above are important. While the considerations that follow are mathematically elementary, they are important for the analysis of real data and have not always been clearly elucidated in the ecological literature, where it has sometimes been incorrectly assumed that intrinsic noise effects are only important if the total population of each species is small [11]. In fact, the scale of fluctuations is governed instead by the population size in a volume (indicated by the parameter V in the calculation above) sufficiently small that the crossing time for diffusion is smaller than the reaction times. The

confusion arises because when space is neglected, the organisms are all confined to such a small volume, so the scale of fluctuations is determined by the total population size [66, 2].

What the current calculation shows is that there are two separate limits in the construction of reaction-diffusion models. One of these limits yields a particular kind of mean field theory, and the other, corresponding to what would traditionally be called the thermodynamic limit in statistical physics, does not yield a mean field theory at all. Only in $d = 0$ do these limits coincide. Recall that the theory was constructed by creating a lattice of patches, each patch of volume V, and then taking the limit of an infinite number of patches, and looking at the continuum version of the theory. The thermodynamic limit corresponds to the limit as the number of patches goes to infinity, while the mean field limit corresponds to taking the volume of each patch, V, to infinity.

The parameter V is determined by the kinetics of the system and is finite whenever diffusion is significant. The finite V dynamics described in this chapter can only be neglected when the typical number of organisms in a well mixed patch of volume V is extremely large, since the finite V fluctuations are large. This is discussed in more detail in the next section.

5.7 Validity of the large V expansion and the scale of quasi-patterns

The expansion considered above is only strictly valid near the onset of quasi-patterns. While in the absence of space, the expansion is valid quite generally, leading to excellent agreement between theory and simulation for the power spectrum [2], the spatial structures do not seem to be as well captured by the expansion deep in the quasi-pattern regime. This is probably due to fluctuation corrections to the mean field not studied in the current chapter. This is suggested by simulated trajectories of the reaction-diffusion master equation using the exact algorithm of Gillespie [120]. The results of this calculation, along with the location on the phase diagram simulated are shown in fig 5.4

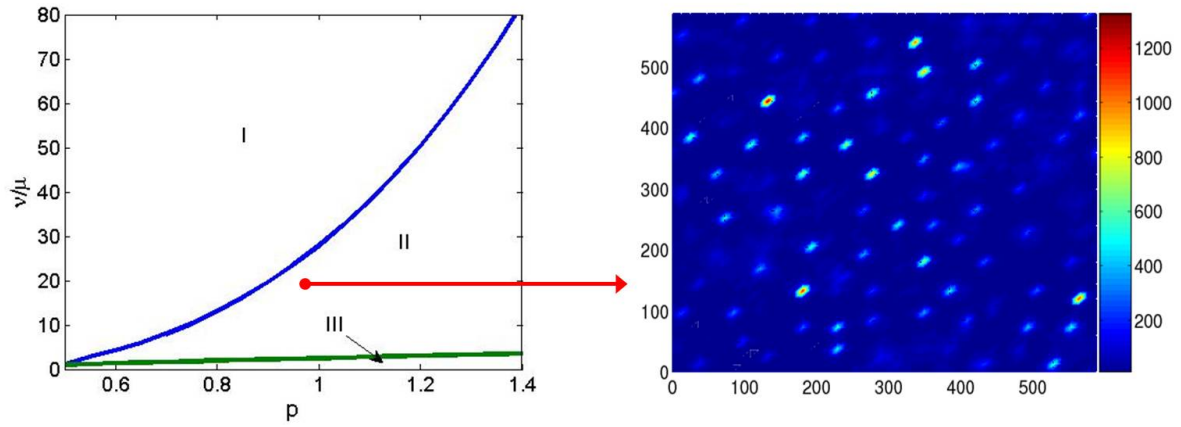

Figure 5.4: The phase diagram, with shading removed, is on the left. Region I is mean field pattern formation, region II is fluctuation driven pattern formation, and region III is homogeneous. The red arrow has its tail on the approximate location in parameter space simulated to produce the spatial patterns shown on the right. The image on the right hand is a heat map of population density in two dimensions. Note that the number of organisms is highly variable, even though mean field predicts a homogeneous state. The fluctuation effects are very large and variable, with patch populations ranging from 1200 to 0. The axes are the lattice index from simulation.

The calculation indicates that the patterns deep inside the quasi-pattern phase are non-perturbative, due to the large variance in populations. We expect that the non-perturbative corrections to the mean field solutions arise at higher order in the expansion. The analytical theory above does not predict the power spectrum of these patterns, but the calculation of wavelength and period above are still approximately valid, since they are obtained by finding the least stable modes, which are likely still dominant, even in the non-perturbative regime.

5.8 Explaining the failure of mean field theory

From the above calculation, as well as related calculations ranging from zero dimensional models of ecosystems to models of biochemical oscillations [13, 2, 75, 111] it is clear that in many applications where the fundamental physics contains intrinsic noise, mean field theory fails to describe the oscillatory dynamics in time and space of the system even for relatively

large systems with many degrees of freedom far from a critical point. Qualitatively, this failure can be understood quite generally by considering the nature of mean field theory.

While there are many ways to derive mean field theories [14], to understand the failure of mean field theory, the simplest approach for systems described by a master equation is to note that there are two essential steps to deriving a mean field theory: averaging and neglecting correlations.

Consider the first step, averaging. The average of the trajectories is given by

$$\varphi = \langle N(t,x) \rangle = \lim_{M_\zeta \to \infty} \frac{1}{M_\zeta} \sum_\zeta N_\zeta(t,x) \tag{5.57}$$

where ζ is the index for realizations of the discrete Markov process for the population dynamics, M_ζ is the number of realizations sampled, and $N_\zeta(t,x)$ is the realization of the discrete Markov process. Each individual realization may be oscillatory, but the oscillations will have a great deal of noise in their amplitude and phase. Summing over these oscillatory contributions will under many conditions lead to an average, φ, that is no longer oscillatory because the variance in amplitude and phase between different realizations ζ of the stochastic process will lead to cancellations of the oscillatory parts in the sum for the average above. That is, the sum of noisy oscillations is not always oscillatory. Since mean field theory considers the dynamics of averages, it will not capture the oscillations present in individual realizations of the dynamics unless the feedbacks that generate the oscillations are much more important than fluctuations (see fig. 5.5).

5.9 Application to field data and experiments

While the calculation above was intended primarily to shed light on the broad theoretical question of the fine tuning problem in Turing instabilities rather than the Levin-Segel model alone, it would still be satisfying to match the predictions above to plankton data. Such an

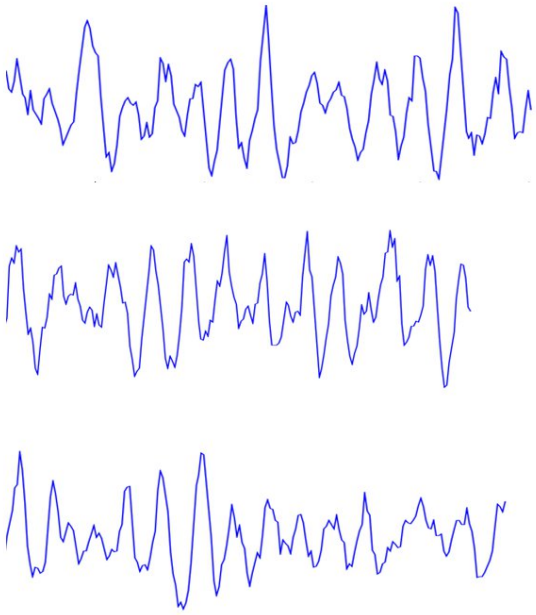

Figure 5.5: Sample trajectories of the Markov process for predator-prey dynamics. Note that while each is roughly oscillatory, a mean field theory derived from the average of many such trajectories would not contain oscillations.

application to current field data in planktonic systems is very difficult. In part this is because data on plankton patterns are primarily gathered for large scale spatial patterns that are driven by turbulent stirring, rather than biological interactions as in the theory presented here [102]. Convection accounts for most of the spatial heterogeneity of plankton at scales above tens of meters [102]. However, there do exist some limited data on plankton population heterogeneity at meter and shorter length scales [101]. Further data on the motility of plankton suggest that the ratio of diffusivities for predator-prey pairs is of order 10 [121]. We calculated above that with generic parameters, the criterion 5.47 yields $\nu/\mu > 2.48$, while the Turing condition yields $\nu/\mu > 27.8$. Under these conditions, it is likely that some populations have fluctuation driven patterns, if the Turing mechanism is responsible for the pattern formation. Current data are not, to our awareness, adequate to go much further.

There are several additional problems with applying the current theory to real planktonic systems, even if the data were to be much higher resolution. The first is that plankton

are enormously diverse, with many species interacting with many others, and body sizes and behaviors spanning several orders of magnitude [122, 123]. A second problem is that the current theory is so simplified that there is no clear connection between many of the parameters in the model and what is measured in real populations. The best way to carry out the identification of quasi-patterns is probably not to engage in detailed modeling of the population dynamics, but rather to use model independent predictions, such as the density dependence of the correlations described above, and the power of k and ω for large values of k and ω in the power spectrum. Data sets associated with plant systems are likely to be amenable to such analysis [15]. Additionally, laboratory experiments in engineered microbial [95], or even in chemical systems (see above comments on the thermodynamic limit) may provide more controlled ways to detect quasi-patterns.

5.10 Conclusions and prospects for future research

We conclude by noting that our analysis of the model in Eqs. 5.1 has demonstrated that Turing patterns are much more generic than is to be expected on the basis of mean field theory, partial differential equation analysis. We also have pointed out some possible ways in which the fluctuation driven spatiotemporal patterns discussed here can be identified in real data. While this chapter focused on a single model, we wish to emphasize again that the model was deliberately chosen to be highly generic with the goal of providing broad insight into the statistical mechanics of the Turing mechanism that can be widely applied. As noted in the introduction, the conjectured wide applicability of this result has received some support from calculations on the Brusselator model [112] and a model of Turing patterns in neuronal networks studied in the following chapter [104]. Further applications of this theory are potentially as wide as the applicability of the Turing mechanism, which, as was pointed out in the introduction, has been used to explain patterns in an enormous variety of systems. In fact, we conjecture that perhaps many or most observed Turing patterns

are the quasi-patterns predicted in this chapter. To demonstrate this conjecture, the next step is to apply the concepts in this chapter to an experimentally well-characterized system, such as an engineered bacterial system with Turing feedbacks. Another important avenue of investigation is to further explore ways to distinguish between quasi-patterns and mean field patterns. Further theoretical progress may also be made by addressing with a similarly detailed theory other noise driven spatiotemporal patterns such as intrinsic noise driven epidemic waves, which seem to be present in measles and dengue fever epidemics [124, 125].

Chapter 6

Emergent spatiotemporal pattern formation in the neocortex

In the previous chapters, we used the formalism of many body physics to study the emergence of fluctuation driven emergent states of matter, which we called quasi-patterns and quasi-cycles. The applications were drawn from predator-prey ecology, but we noted that the long wavelength and long time physics of the emergent quasi-patterns and quasi-cycles should be universal across systems with couplings qualitatively similar to those studied in ecology. In this chapter we study a nontrivial application of quasi-patterns and quasi-cycles to the visual cortex in the brain. The work has connections to a variety of interesting speculations on the relationship between the universal forms of the patterns of geometric visual hallucinations that correspond to the spatially patterned states of the visual cortex and the nature of universal geometric motifs in shamanic religion and art, which will be discussed below. We then will show through systematic calculation that the structure of the visual cortex supports both quasi-cycles and quasi-patterns, but is configured to minimize the risk that the normal function of vision will be overwhelmed by spontaneously generated internal brain activity.

Under typical conditions, the visual cortex processes data from the eye by representing an image with excitation patterns of neurons [8]. However, under the influence of hallucinogenic drugs, exhaustion, migraines, or other conditions, the visual cortex may produce spontaneous patterns of excitation that overwhelm visual function [126]. These patterns of excitation are experienced as visual hallucinations [126] and are generally in one of four classes of spatial patterns, called form constants [127]. There is now good evidence that the form constants correspond to the basic symmetries (euclidean plus a twist symmetry) of the visual cortex that are broken as pattern formation emerges due to the Turing mechanism [128]. These

hallucinations have forms like star shapes, spirals, and honeycombs (for a couple of examples, see fig. 6.1).

The rotational structure of the patterns in fig. 6.1 is due to the mapping between the visual field and the visual cortex, which away from singularities takes the form of a rescaled complex logarithm [126, 8]. With visual field coordinates subscripted with V and the coordinates of the visual cortex in rectangular form, the mapping is

$$\begin{aligned} x &= \alpha ln(\epsilon r_V) \\ y &= \beta \theta_V \end{aligned} \tag{6.1}$$

with α, ϵ and β parameters. A straight line in visual cortex coordinates is given by the equation $x = \eta y + m$. Inverting the Eqs. 6.1 to get the form of the straight line in visual field coordinates yields

$$r_V = \frac{1}{\epsilon} \exp\left[\frac{\beta\eta}{\alpha}\theta + \frac{m}{\alpha}\right] \tag{6.2}$$

The equation 6.2 is the equation for a logarithmic spiral, and the mapping of straight lines on the visual cortex to spirals in visual coordinates is the explanation for how regular pattern formation on the visual cortex gives rise to the exotic spiral shapes observed in the form constants.

The geometry of hallucinatory form constants is extremely common in art associated with shamanism in prehistoric religions [8, 9]. This is apparently because shamanism frequently involves hallucinogen induced journeys to what anthropologists have called "alternative states of consciousness" [9]. It is speculated by specialists in rock art and ancient religion that the spiral shapes seen in the early stages of the influence of hallucinogenic drugs were interpreted as gateways to a spirit world and that much of rock art describes these journeys (see fig. 6.2) [9].

Others, such as the doctor and neuroscientist Oliver Sacks who regularly experiences visual hallucinations himself through "visual migraines," speculate based on the frequency with which people experience geometric visual hallucinations without the input of drugs (about one in ten) that geometric hallucinations have led to "human universals" in art:

"I started to wonder whether what I had taken to be a personal experience and resonance might in fact be part of a larger whole, whether certain basic forms of geometric art, going back for tens of thousands of years, might also reflect the external expression of universal experiences. Migraine-like patterns, so to speak, are seen not only in Islamic art, but in classical and medieval motifs, in Zapotec architecture, in the bark paintings of Aboriginal artists in Australia, in Acoma pottery, in Swazi basketry — in virtually every culture."

While such hypotheses are still deep in the realm of speculation, images of the art Sacks mentions do present remarkable parallels, as can be seen by comparing a piece of Acoma pottery in fig.6.3 to the hallucinations in fig.6.1. But even without the connection to art and shamanism, the origins of geometric visual hallucinations as a failure mode of the visual cortex present a well formulated problem in the theory of the Turing mechanism and quasi-patterns.

In the previous chapter on ecological pattern formation, we observed that the widespread spatial pattern formation present field studies was mysterious given that typical Turing type

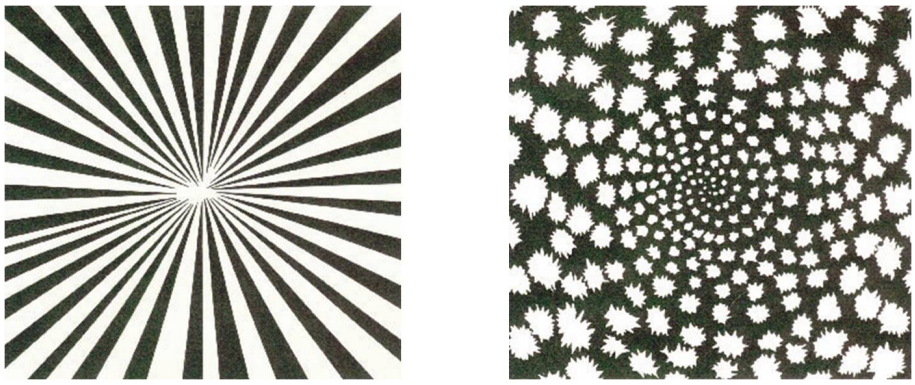

Figure 6.1: Two examples of form constants in visual hallucinations. Figure taken from [8].

Figure 6.2: Shamanic cave art from the Texas-Mexico border. The shaman, on the left, is diving into to a spirit world. Note the spiral images in the upper right, probably interpreted as gateways to the spirit world. The spiral image is also a reasonable representation of typical spiral hallucinations seen at the onset of the effects of hallucinogenic drugs. Figure taken from [9].

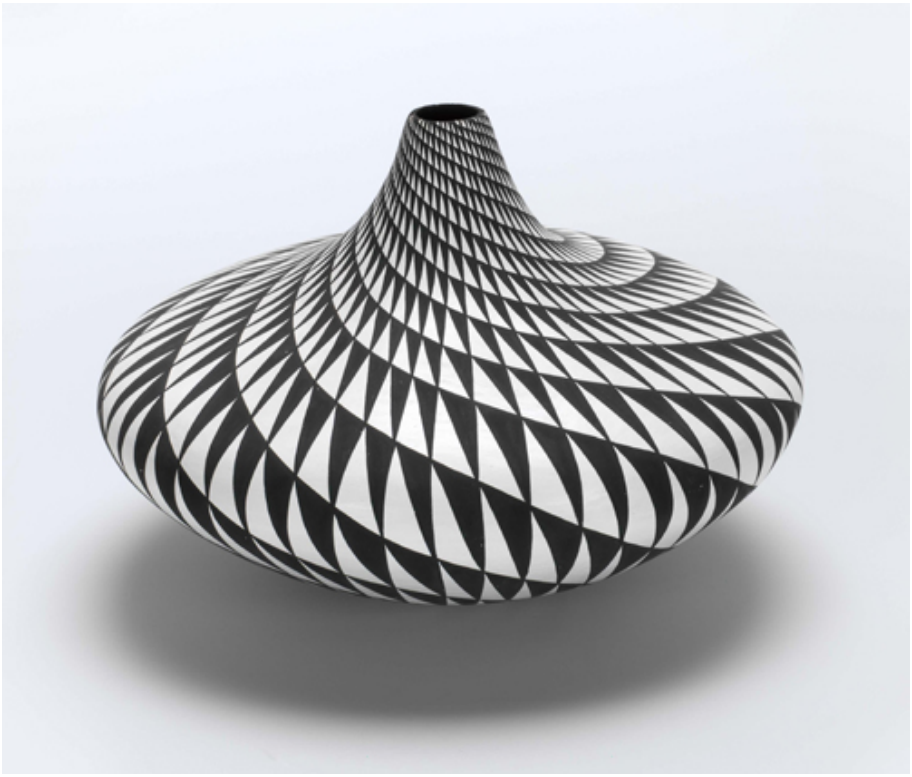

Figure 6.3: A contemporary piece of Acoma pottery from the American Southwest. Figure taken from an online display of southwestern native art at `http://faculty.vassar.edu/lucic/formsofexchange/pottery.html`.

models of pattern formation require fine tuning to generate patterns, which is equivalent to a prediction that pattern formation should be very rare in natural systems [13, 15]. We showed that this discrepancy between prediction and observation is resolved by including the intrinsic noise that is neglected in the typical mean field models of spatial pattern formation. The visual cortex has similar feedback structures, with long range inhibition and short range excitation, to pattern forming ecosystems [129, 130]. However, in the case of the visual cortex, spatial pattern formation corresponds to geometric visual hallucinations [126, 8]. So the visual cortex is either fine tuned away from pattern forming states, or has some mechanism for robust avoidance of spontaneously generated spatial structures. In this chapter we show that the network architecture of the visual cortex is configured such that spontaneous spatial pattern formation can only emerge under relatively extreme circumstances, even with intrinsic noise.

6.1 Modeling the visual and neocortex

Parallel to the development of a theory of the role of intrinsic noise in population biology outlined in the previous chapters, is a development of theories describing the role of intrinsic noise in neurobiology [131, 132, 133]. This last development grew out of the recognition of an analogy between predator-prey cycles in population biology and oscillations in networks of coupled excitatory and inhibitory neurons [134, 135]. Here we analyze quasi-cycles and quasi-patterns in a simple model of the neocortex, and study their implications for the robustness of the visual cortex to the formation of the hallucinatory state described above.

6.1.1 General anatomical features of the neocortex

The human neocortex consists of a three-dimensional slab of neurons and neuroglia connected together in a dense neuropil. There are about 3×10^{11} neurons in the human neocortex, which has a surface area of about $2m^2$ and a thickness of about $3mm$. Thus its total volume

is about $6.10^6 mm^3$, and ρ, the packing density of human neocortical neurons, is about $50,000$ cells/mm^3. There are about $150,000$ cells in a block of surface area $1mm^2$. It has now been established that the visual cortex is organized on a modular basis of such blocks [136]. It has also been established that each cell has about 4000 contacts [137], i.e. each cell is coupled to at most 2.5% of the others. So the connectivity is rather sparse.

6.1.2 A simplified canonical microcircuit for cortical modules

We now look at modular circuitry. Details of the functional anatomy of a simplified version of the Douglas-Martin canonical microcircuit [138] (our choice for the sub-blocks of modular circuitry) are shown in figure 6.4. This microcircuit summarizes the neuronal cell types and

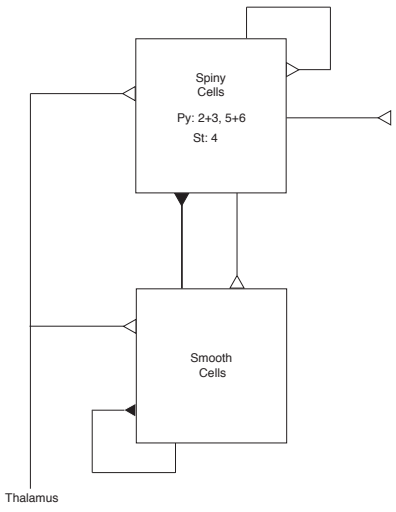

Figure 6.4: Block diagram of a reduced neocortical microcircuit. Py =pyramidal neurons, St = spiny stellate neurons. Unfilled triangles = excitatory synapses, Filled triangles = inhibitory synapses.

their interconnections found in a typical module. The spiny cells include Pyramidal (Py) in layers 2, 3, 5, and 6, and Stellate (St) cells in layer 4 of the neocortex. All these cells are

excitatory, meaning that when they are excited, they tend to excite the neurons they have downstream connections to. The smooth cells include Basket, Martinotti and Chandelier cells at all layers and are all inhibitory, meaning they tend to inhibit the excitation of neurons that they have downstream connections to when they are excited. In the next section we introduce a stochastic model of such a microcircuit.

6.1.3 The statistical dynamics of a single microcircuit

In order to analyze the action of such a circuit we first introduce master equations and actions for neural networks. We first introduce a simple Markov model of the action of a single neuron [131]. Each neuron can be in one of two states, quiescent (q) or activated (a). The rate for the transition $q \to a$ is $f[s]$ where f is a smooth saturating function of the input current s, above a current threshold $s_{\rm TH}$. The rate for the transition $a \to q$ is a constant α.

Following [133] the master equation for $P_{m,n}(t)$, the probability of finding m active excitatory neurons, and n active inhibitory neurons at time t can then be written as:

$$\begin{aligned}
\frac{dP_{m,n}(t)}{dt} &= \alpha_E[(m+1)P_{m+1,n} - mP_{m,n}] \\
&+ (M-m+1)f_E[s_E(m-1,n)]P_{m-1,n} \\
&- (M-m)f_E[s_E(m,n)]P_{m,n} \\
&+ \alpha_I[(n+1)P_{m,n+1} - nP_{m,n}] \\
&+ (N-n+1)f_I[s_I(m,n-1)]P_{m,n-1} \\
&- (N-n)f_I[s_I(m,n)]P_{m,n}
\end{aligned} \tag{6.3}$$

where

$$s_E(m,n) = w_{EE}m - w_{EI}n + h_E, \quad s_I(m,n) = w_{IE}m - w_{II}n + h_I \tag{6.4}$$

are the currents driving the neurons.

The resulting action can be written down as:

$$\begin{aligned} M\int dt\,[\hat{\varphi}_E\partial_t\varphi_E &+ \hat{\psi}_E\partial_t\psi_E + \alpha(\hat{\varphi}_E\varphi_E - \hat{\psi}_E\varphi_E) \\ &- (\hat{\varphi}_E\psi_E - \hat{\psi}_E\psi_E)f_E[s_E]\Big] \\ +N\int dt\,[\hat{\varphi}_I\partial_t\varphi_I &+ \hat{\psi}_I\partial_t\psi_I + \alpha(\hat{\varphi}_I\varphi_I - \hat{\psi}_I\varphi_I) \\ &- (\hat{\varphi}_I\psi_I - \hat{\psi}_I\psi_I)f_I[s_I]\Big] \end{aligned} \tag{6.5}$$

where

$$\begin{aligned} s_E &= w_{EE}(\hat{\varphi}_E\varphi_E + \varphi_E) - w_{EI}(\hat{\varphi}_I\varphi_I + \varphi_I) + h_E \\ s_I &= w_{IE}(\hat{\varphi}_E\varphi_E + \varphi_E) - w_{II}(\hat{\varphi}_I\varphi_I + \varphi_I) + h_I \end{aligned} \tag{6.6}$$

this can be written in the density representation, using

$$\begin{aligned} \varphi_E &= me^{-\hat{m}}, \quad \hat{\varphi_E} = e^{\hat{m}} - 1, \quad \hat{\varphi_E}\varphi_E + \varphi_E = m \\ \psi_E &= pe^{-\hat{p}}, \quad \hat{\psi_E} = e^{\hat{p}} - 1, \quad \hat{\psi_E}\psi_E + \psi_E = p \\ \varphi_I &= ne^{-\hat{n}}, \quad \hat{\varphi_I} = e^{\hat{n}} - 1, \quad \hat{\varphi_I}\varphi_I + \varphi_I = n \\ \psi_I &= qe^{-\hat{q}}, \quad \hat{\psi_I} = e^{\hat{q}} - 1, \quad \hat{\psi_I}\psi_I + \psi_I = q \end{aligned} \tag{6.7}$$

as

$$
\begin{aligned}
S = \int dt \quad M \quad & \big\{(\hat{m}\partial_t m + \hat{p}\partial_t p + \alpha_E m(1 - e^{-(\hat{m}-\hat{p})}) \\
- \quad & p(e^{(\hat{m}-\hat{p})} - 1) f_E[s_E]) \\
+ \quad & \gamma\,(\hat{n}\partial_t n + \hat{q}\partial_t q + \alpha_I n(1 - e^{-(\hat{n}-\hat{q})}) \\
- \quad & q(e^{(\hat{n}-\hat{q})} - 1) f_I[s_I])\big\}
\end{aligned}
\tag{6.8}
$$

where $\gamma = N/M = 0.25$, m and n are densities of active neurons, and $\hat{m}$ and $\hat{n}$ represent the effects of intrinsic fluctuations. Note also that the action $S[m, \hat{m}; n, \hat{n}]$ is that of a single cortical microcircuit, and that $\delta S = \mathcal{L}dt$ where $\mathcal{L}$ is the neural network Lagrangian.

6.1.4 Neocortex as a two dimensional lattice

We now introduce a model of an extended slice of neocortex as a two-dimensional lattice of canonical microcircuits. However we first need to distinguish two length-scales. The first we call *local*. This is the length-scale on which canonical microcircuits interact with their neighbors. We assume that this local scale comprises all the microcircuits in one module. In the simplified model described in this article we assume that there are 9 such microcircuits per module. Thus there are about 1.67×10^4 cells per microcircuit, of which 1.33×10^4 cells are excitatory and 0.33×10^4 cells are inhibitory. Based on neuroanatomical data we assume that intramodular excitatory and inhibitory connections exist between all microcircuits within a module [See [129]]. However, it is also possible, depending on which inhibitory interneurons we choose to model, that the local inhibition is longer ranged than the local excitation.

The second length-scale we call *intermediate*. On this length-scale modules are coupled together by patchy excitatory connections. These connections have a range of about $\pm 4mm$ with axonal arbors every 1 mm or so. [130]. Thus they are intermodular. If we take the

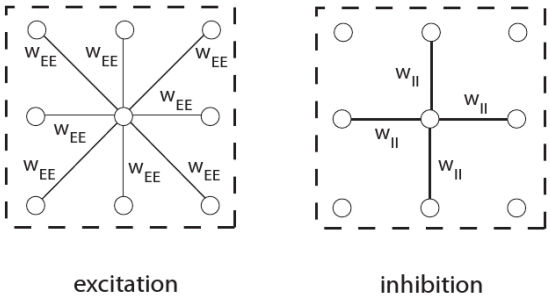

Figure 6.5: Block diagram of intra-modular coupling.

lattice spacing between individual microcircuits to be Lmm then the intermodular spacing is $\sqrt{a}Lmm$ where a is the number of microcircuits per module. There is also some data on the connections made by a special class of Basket cells called Large Basket cells (LBC), which have long axons and so can provide intermodular inhibition [139]. Thus the connections between differing lattice sites support a mixture of both local and intermediate excitation and inhibition, and the currents driving neocortical neurons are now functions of position, i.e., $s_E \rightarrow s_i^E(m,n), \quad s_I \rightarrow s_i^I(m,n)$, and the intra- and inter-modular couplings are as shown in figures [6.5] and [6.6].

As to long-range connections, which would provide a third length-scale, the best evidence to date [140] suggests that they are small-world [141]. However in this chapter we do not consider their effects.

It follows that we can extend both the master equations and the actions derived for a single canonical microcircuit to that of a lattice of such microcircuits simply by indexing them for position in the lattice. Let Ω be the number of microcircuits in the lattice, and let $j = 1, \cdots, \Omega$ denote the lattice coordinates of any microcircuit. Then, for example, the

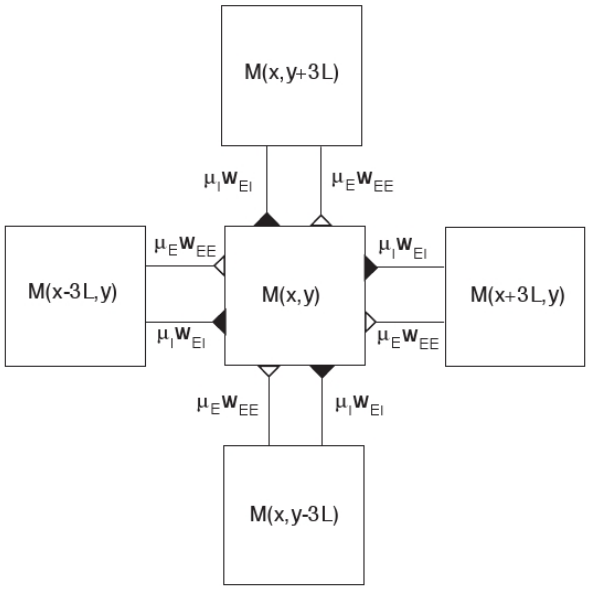

Figure 6.6: Block diagram of inter-modular coupling.

generalization of the action given in equation [6.8] is

$$\begin{aligned}\int dt \quad M \quad & \sum_{j=1}^{\Omega} \big\{ (\hat{m_j}\partial_t m_j + \hat{p_j}\partial_t p_j + \alpha_E m_j (1 - e^{-(\hat{m_j} - \hat{p_j})}) \\ - \quad & p_j (e^{(\hat{m_j} - \hat{p_j})} - 1) f_E[s_j^E]) \\ + \quad & \gamma \, (\hat{n_j}\partial_t n_j + \hat{q_j}\partial_t q_j + \alpha_I n_j (1 - e^{-(\hat{n_j} - \hat{q_j})}) \\ - \quad & q_j (e^{(\hat{n_j} - \hat{q_j})} - 1) f_I[s_j^I]) \big\}\end{aligned} \tag{6.9}$$

6.1.5 A Lagrangian for neocortical networks

It follows directly from equation [6.9] that a neocortical Lagrangian can be written down in the form:

$$\begin{aligned}
\mathcal{L} &= M\sum_{j=1}^{\Omega}\big\{(\hat{m_j}\partial_t m_j + \hat{p_j}\partial_t p_j + \alpha_E m_j(1-e^{-(\hat{m_j}-\hat{p_j})}) \\
&- p_j(e^{(\hat{m_j}-\hat{p_j})}-1)f_E[s_j^E]) \\
&+ \gamma\,(\hat{n_j}\partial_t n_j + \hat{q_j}\partial_t q_j + \alpha_I n_j(1-e^{-(\hat{n_j}-\hat{q_j})}) \\
&- q_j(e^{(\hat{n_j}-\hat{q_j})}-1)f_I[s_j^I])\big\}
\end{aligned} \tag{6.10}$$

The currents given in equation [6.6] can be written in the form:

$$\begin{aligned}
s_i^E(m,n) &= g_E(\Delta)w_{EE}m_i - g_I(\Delta)w_{EI}n_i + h_i^E \\
s_i^I(m,n) &= g_E(\Delta)w_{IE}m_i - g_I(\Delta)w_{II}n_i + h_i^I
\end{aligned} \tag{6.11}$$

where $g_E(\Delta)$ and $g_I(\Delta)$ are given by

$$\begin{aligned}
g_E(\Delta) &= (1+\mu_1^E+\mu_2^E) + \mu_1^E\Delta_1^E + \mu_2^E\Delta_2^E \\
g_I(\Delta) &= (1+\mu_1^I+\mu_2^I) + \mu_1^I\Delta_1^I + \mu_2^I\Delta_2^I
\end{aligned} \tag{6.12}$$

and are given by

$$\Delta_1^E = \frac{1}{4d}\sum_{j\neq i} - \sum_j \delta_{ij}, \quad \Delta_2^E = \frac{1}{2d}\sum_{j\neq i} - \sum_j \delta_{ij} \tag{6.13}$$

Δ_1^E and Δ_2^E are, respectively, the discrete 9-point and 5-point Laplacian operators in $d-$dimensions, and Δ_1^I, Δ_2^I are defined as the five-point operators:

$$\Delta_1^I = \frac{1}{2d}\sum_{j\neq i} - \sum_j \delta_{ij}, \quad \Delta_2^I = \frac{1}{2d}\sum_{j\neq i} - \sum_j \delta_{ij} \tag{6.14}$$

But since the weights refer to interactions between homogeneous microcircuits, they are effectively constants, so we can rewrite the current equations in the slightly modified form

$$\begin{aligned} s_i^E(m,n) &= g_E(\Delta)w_{EE}m_i - g_I(\Delta)w_{EI}n_i + h_i^E \\ s_i^I(m,n) &= g_E(\Delta)w_{IE}m_i - g_I(\Delta)w_{II}n_i + h_i^I \end{aligned} \tag{6.15}$$

6.1.6 System-size expansion of the Lagrangian

We now carry out the van Kampen system size expansion of the Master equation. To do this we assume that the state of the system characterized by (m,n) is at a stable fixed point (φ_0, ϕ_0). We then assume that there are Gaussian fluctuations (η, ξ) about this state so that in general:

$$\begin{aligned} m &= \varphi + \frac{\eta}{\sqrt{M}}, \quad n = \phi + \frac{\xi}{\sqrt{N}} \\ p &= \vartheta + \frac{\zeta}{\sqrt{M}}, \quad q = \theta + \frac{\xi}{\sqrt{N}} \end{aligned} \tag{6.16}$$

Using equation [6.16] and expanding the exponentials, we obtain the expanded Lagrangian in the form:

$$\mathcal{L} = \sqrt{M}\mathcal{L}_1 + \mathcal{L}_2 + O(\frac{1}{\sqrt{M}}) \tag{6.17}$$

where $M/N = \gamma$ remains constant as we increase both M and N, and

$$\begin{aligned}
\mathcal{L}_1 &= \sum_{j=1}^{\Omega} \{[\hat{m}_j \partial_t \varphi_j + \hat{p}_j \partial_t \vartheta_j + \ \alpha_E \varphi_j (\hat{m}_j - \hat{p}_j) \\
&- \ \vartheta_j (\hat{m}_j - \hat{p}_J) f_E[s_j^E]] \\
&+ \ \sqrt{\gamma}\, [\hat{n}_j \partial_t \phi_j + \hat{q}_j \partial_t \theta_J + \alpha_I \phi_j (\hat{n}_j - \hat{q}_j) \\
&- \ \theta_j (\hat{n}_j - \hat{q}_j) f_I[s_j^I]]\}
\end{aligned} \tag{6.18}$$

where:

$$\begin{aligned}
s_i^E &= g_E(\Delta) w_{EE} \varphi_i - g_I(\Delta) w_{EI} \phi_i + h_i^E \\
s_i^I &= g_E(\Delta) w_{IE} \varphi_i - g_I(\Delta) w_{II} \phi_i + h_i^I
\end{aligned} \tag{6.19}$$

and

$$\begin{aligned}
\mathcal{L}_2 &= \sum_{j=1}^{\Omega} \{[\hat{m}_j \partial_t \eta_j + \hat{p}_j \partial_t \zeta_j + \alpha_E \eta_j (\hat{m}_j - \hat{p}_j) \\
&- \ \frac{1}{2} \alpha_E \varphi_j (\hat{m}_j - \hat{p}_j)^2 - \zeta_j (\hat{m}_j - p_j) f_E[s_j^E] \\
&- \ \frac{1}{2} \vartheta_j (\hat{m}_j - \hat{p}_j)^2 f_E[s_j^E] - \vartheta_j (\hat{m}_j - \hat{p}_j) f_E' \cdot (\delta s_j^E)] \\
&+ \ \gamma\, [\hat{n}_j \partial_t \xi_j + \hat{q}_j \partial_t \chi_j + \alpha_I \xi_j (\hat{n}_j - \hat{q}_j) \\
&- \ \frac{1}{2} \alpha_I \phi_j (\hat{n}_j - \hat{q}_j)^2 - \chi_j (\hat{n}_j - q_j) f_E I[s_j^I] \\
&- \ \frac{1}{2} \theta_j (\hat{n}_j - \hat{q}_j)^2 f_I[s_j^I] - \theta_j (\hat{n}_j - \hat{q}_j) f_I' \cdot (\delta s_j^I)] \Big\}
\end{aligned} \tag{6.20}$$

where:

$$\begin{aligned} \delta s_i^E &= g_E(\Delta) w_{EE} \eta_i - g_I(\Delta) w_{EI} \xi_i / \sqrt{\gamma} \\ \delta s_i^I &= g_E(\Delta) w_{IE} \eta_i - g_I(\Delta) w_{II} \xi_i / \sqrt{\gamma} \end{aligned} \tag{6.21}$$

6.2 Variational Derivatives of the Lagrangians

If we now form variational derivatives of the Lagrangian we obtain at $O(\sqrt{M})$:

$$\begin{aligned} \frac{\delta \mathcal{L}_1}{\delta \hat{m}_i} &= \partial_t \varphi_i + \alpha_E \varphi_i - \vartheta_i f_E[s_i^E] \\ \frac{\delta \mathcal{L}_1}{\delta \hat{p}_i} &= \partial_t \vartheta_i - \alpha_E \varphi_i + \vartheta_i f_E[s_i^E] \\ \frac{\delta \mathcal{L}_1}{\delta \hat{n}_i} &= \sqrt{\gamma}(\partial_t \phi_i + \alpha_I \phi_i - \theta_i f_I[s_i^I]) \\ \frac{\delta \mathcal{L}_1}{\delta \hat{q}_i} &= \sqrt{\gamma}(\partial_t \theta_i - \alpha_I \phi_i + \theta_i f_I[s_i^I]) \end{aligned} \tag{6.22}$$

Similarly at $O(1)$ we obtain:

$$\begin{aligned}
\frac{\delta \mathcal{L}_2}{\delta \hat{m}_i} &= \partial_t \eta_i + \alpha_E \eta_i - \alpha_E \varphi_i (\hat{m}_i - \hat{p}_i) - \zeta_i f_E[s_i^E] \\
&- \vartheta_i (\hat{m}_i - \hat{p}_i) f_E - \vartheta_i f_E' \cdot (\delta s_i^E) \\
\frac{\delta \mathcal{L}_2}{\delta \hat{p}_i} &= \partial_t \zeta_i - \alpha_E \eta_i + \alpha_E \varphi_i (\hat{m}_i - \hat{p}_i) + \zeta_i f_E[s_i^E] \\
&+ \vartheta_i (\hat{m}_i - \hat{p}_i) f_E + \vartheta_i f_E' \cdot (\delta s_i^E) \\
\frac{\delta \mathcal{L}_2}{\delta \hat{n}_i} &= \gamma \left[\partial_t \xi_i + \alpha_I \xi_i - \alpha_I \phi_i (\hat{n}_i - \hat{q}_i) - \chi_i f_I[s_i^I] \right. \\
&- \left. \theta_i (\hat{n}_i - \hat{q}_i) f_I - \theta_i (f_I' \cdot (\delta s_i^I)) \right] \\
\frac{\delta \mathcal{L}_2}{\delta \hat{q}_i} &= \gamma \left[\partial_t \chi_i - \alpha_I \xi_i + \alpha_I \phi_i (\hat{n}_i - \hat{q}_i) + \chi_i f_I[s_i^I] \right. \\
&+ \left. \theta_i (\hat{n}_i - \hat{q}_i) f_I + \theta_i f_I' \cdot (\delta s_i^I) \right]
\end{aligned} \tag{6.23}$$

6.3 Mean-field Wilson-Cowan equations

At a minimum all these variations are zero, whence we obtain at $O(\sqrt{M})$ the equations:

$$\begin{aligned}
\partial_t \varphi_i &= -\alpha_E \varphi_i + \vartheta_i f_E[s_i^E] \\
\partial_t \vartheta_i &= \alpha_E \varphi_i - \vartheta_i f_E[s_i^E] \\
\partial_t \phi_i &= -\alpha_I \phi_i + \theta f_I[s_i^I] \\
\partial_t \theta_i &= +\alpha_I \phi_i - \theta f_I[s_i^I]
\end{aligned} \tag{6.24}$$

Evidently we have:

$$\partial_t (\varphi_i + \vartheta_i) = 0, \quad \partial_t (\phi_i + \theta_i) = 0 \tag{6.25}$$

It follows that we can set

$$\varphi_i + \vartheta_i = \text{constant} = 1, \quad \phi_i + \theta_i = \text{constant} = 1 \tag{6.26}$$

and therefore we can eliminate the equations for the quiescent states p and q and we are left with the mean field equations:

$$\begin{aligned} \partial_t \varphi_i &= -\alpha_E \varphi_i + (1-\varphi_i) f_E[s_I^E] \\ \partial_t \phi_i &= -\alpha_I \varphi_i + (1-\phi_i) f_I[s_I^I] \end{aligned} \tag{6.27}$$

With appropriate rescaling these are the Wilson-Cowan mean-field equations [142], extended to include local spatial interactions between patches.

6.4 Langevin equations

In similar fashion functional derivatives of the $O(1)$ equations yield:

$$\begin{aligned} \partial_t \eta_i &= -\alpha_E \eta_i + \alpha_E \varphi_i (\hat{m}_i - \hat{p}_i) + \zeta_i f_E[s_i^E] \\ &+ \vartheta_i (\hat{m}_i - \hat{p}_i) f_E + \vartheta_i f_E' \cdot (\delta s_i^E) \\ \partial_t \zeta_i &= +\alpha_E \eta_i - \alpha_E \varphi_i (\hat{m}_i - \hat{p}_i) - \zeta_i f_E[s_i^E] \\ &- \vartheta_i (\hat{m}_i - \hat{p}_i) f_E - \vartheta_i f_E' \cdot (\delta s_i^E) \\ \partial_t \xi_i &= -\alpha_I \xi_i + \alpha_I \phi_i (\hat{n}_i - \hat{q}_i) + \chi_i f_I[s_i^I] \\ &+ \theta_i (\hat{n}_i - \hat{q}_i) f_I + \theta_i f_I' \cdot (\delta s_i^I) \\ \partial_t \chi_i &= \alpha_I \xi_i - \alpha_I \phi_i (\hat{n}_i - \hat{q}_i) - \chi_i f_I[s_i^I] \\ &- \theta_i (\hat{n}_i - \hat{q}_i) f_I - \theta_i f_I' \cdot (\delta s_i^I) \end{aligned} \tag{6.28}$$

Evidently

$$\partial_t(\eta_i + \zeta_i) = 0, \quad \partial_t(\xi_i + \chi_i) = 0 \tag{6.29}$$

so that we can set

$$\eta_i + \zeta_i = \text{constant} = 0, \quad \xi_i + \chi_i = \text{constant} = 0 \tag{6.30}$$

since all the fluctuations have zero mean.

We can use all these constraints to eliminate the fluctuations ζ and χ, and the noise densities $\hat{p}$ and $\hat{q}$. Let $\hat{u} = \hat{m} - \hat{p}, \hat{v} = \hat{n} - \hat{q}$, and let $\zeta = -\eta, \chi = -\xi$. Equations [6.28] now reduce to the pair:

$$\begin{aligned}
\partial_t \eta_i &= -\alpha_E \eta_i + \alpha_E \varphi_i \hat{u}_i - \eta_i f_E[s_i^E] \\
&+ (1 - \varphi_i)\hat{u}_i f_E + (1 - \varphi_i) f_E' \cdot (\delta s_i^E) \\
\partial_t \xi_i &= -\alpha_I \xi_i + \alpha_I \phi_i \hat{v}_i - \xi_i f_I[s_i^I] \\
&+ (1 - \phi_i)\hat{v}_i f_I + (1 - \phi_i) f_I' \cdot (\delta s_i^I)
\end{aligned} \tag{6.31}$$

To obtain fluctuation or Langevin equations we first Laplace transform equation [6.31] with respect to time, and then rewrite the equations of motion for the fluctuations in vector form as:

$$s\mathbf{x} = A\mathbf{x} + B\mathbf{y} \tag{6.32}$$

assuming $\mathbf{x}_j(0) = 0$, and where

$$\mathbf{x}_j = \begin{pmatrix} \eta_j \\ \xi_j \end{pmatrix} \qquad \mathbf{y}_j = \begin{pmatrix} \hat{u}_j \\ \hat{v}_j \end{pmatrix} \tag{6.33}$$

and the matrices A and B are given as:

$$A_j = \begin{pmatrix} \substack{-\alpha_E - f_E + \\ (1-\varphi_j) f'_E g_E(\Delta) w_{EE}} & -(1-\varphi_j) f'_E g_I(\Delta) w_{EI}/\sqrt{\gamma} \\ \\ (1-\phi_j) f'_I g_E(\Delta) w_{IE} & \substack{-\alpha_I - f_I \\ -(1-\phi_j) f'_I g_I(\Delta) w_{II}/\sqrt{\gamma}} \end{pmatrix} \tag{6.34}$$

and

$$B_j^2 = \begin{pmatrix} \alpha_E \varphi_j + (1-\varphi_j) f_E & \cdot \\ \cdot & \alpha_I \phi_j + (1-\phi_j) f_I \end{pmatrix}, \tag{6.35}$$

These equations describe the dynamics of the fluctuations of the activity η and ξ in terms of the mean field variables φ and ϕ, and the variables $\hat{u}$ and $\hat{v}$. As noted in [49] $\hat{u}$ and $\hat{v}$ are the components of the response field in the Martin-Siggia-Rose response function formalism (see appendix one).

We now Fourier transform equation [6.32] using the discrete Fourier transform

$$f_{\mathbf{k}} = L^d \sum_{\mathbf{j}} e^{-i\mathbf{k}\cdot L\mathbf{j}} f_{\mathbf{j}}, \qquad f_{\mathbf{j}} = L^{-d}\Omega^{-1} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot L\mathbf{j}} F_{\mathbf{k}} \tag{6.36}$$

subject to the condition:

$$\sum_{\mathbf{j}} e^{-i\mathbf{k}\cdot L\mathbf{j}} = \Omega \delta_{\mathbf{j}.0} \tag{6.37}$$

The resulting Langevin equation can now be written down in the form

$$\mathbf{x}_{\mathbf{k}} = (s - A(\mathbf{k}))^{-1} B\mathbf{y} = D(\mathbf{k}, s) B\mathbf{y} \tag{6.38}$$

whence

$$\langle \mathbf{x}_{\mathbf{k}} \mathbf{x}_{\mathbf{k}}^{\dagger} \rangle = D(\mathbf{k}, s) B \langle \mathbf{y}\mathbf{y}^{\dagger} \rangle B^T D(\mathbf{k}, s)^{\dagger} \tag{6.39}$$

where $\dagger$ denotes the adjoint. But $\mathbf{y}$ is delta-correlated Gaussian noise with unit variances,

so:

$$\langle \mathbf{y_k} \mathbf{y_k}^\dagger \rangle = \begin{pmatrix} 1 & \cdot \\ \cdot & 1 \end{pmatrix} = I \tag{6.40}$$

so

$$\langle \mathbf{x_k} \mathbf{x_k}^\dagger \rangle = D(\mathbf{k}, s) B B^T D(\mathbf{k}, s)^\dagger \tag{6.41}$$

We note that in our Fokker-Planck-Langevin approach we assume that

$$BB^T = \begin{pmatrix} 2\alpha_E \varphi_0 & \cdot \\ & \cdot \\ \cdot & 2\alpha_I \phi_0 \end{pmatrix} \tag{6.42}$$

where φ_0 and ϕ_0 are the fixed point solutions of the mean-field Wilson-Cowan equations [6.27]. We use this assumption in what follows.

We also note that in taking the discrete Fourier transform of A_j we require the discrete Fourier transform of the discrete Laplacian operators Δ^E_{1j}, Δ^E_{2j}, Δ^I_{1j} and Δ^I_{2j} i.e.

$$\Delta^E_1(\mathbf{k}) = \frac{1}{2d} \sum_{\gamma=1}^{2d} [\cos(k_\gamma L) - 1], \quad \Delta^E_2(\mathbf{k}) = \frac{1}{d} \sum_{\gamma=1}^{d} [\cos(k_\gamma L) - 1] \tag{6.43}$$

and

$$\Delta^I_1(\mathbf{k}) = \frac{1}{d} \sum_{\gamma=1}^{d} [\cos(k_\gamma L) - 1], \quad \Delta^I_2(\mathbf{k}) = \frac{1}{d} \sum_{\gamma=1}^{d} [\cos(k_\gamma L) - 1] \tag{6.44}$$

where $\mathbf{k}$ is restricted to the first Brillouin zone $-\pi \leq k_\gamma \leq \pi$

We can approximate this transform using the expansion:

$$\cos(k_\gamma) \approx 1 - \frac{k_\gamma^2}{2} + O[k^4] \tag{6.45}$$

whence:

$$\begin{aligned}\Delta_1^E(\mathbf{k}) &\approx -\frac{3}{8}k^2, \quad \Delta_2^E(\mathbf{k}) \approx -\frac{9}{4}k^2 \\ \Delta_1^I(\mathbf{k}) &\approx -\frac{1}{4}k^2, \quad \Delta_2^I(\mathbf{k}) \approx -\frac{9}{4}k^2\end{aligned} \tag{6.46}$$

Thus the matrices A_j and B_j transform to:

$$A_{\mathbf{k}} = \begin{pmatrix} \overset{-\alpha_E - f_E +}{(1-\varphi) f_E' g_E(\Delta_{\mathbf{k}}) w_{EE}} & -(1-\varphi) f_E' g_I(\Delta_{\mathbf{k}}) w_{EI}/\sqrt{\gamma} \\ \\ (1-\phi) f_I' g_E(\Delta_{\mathbf{k}}) w_{IE} & \overset{-\alpha_I - f_I}{-(1-\phi) f_I' g_I(\Delta_{\mathbf{k}}) w_{II}/\sqrt{\gamma}} \end{pmatrix} \tag{6.47}$$

where

$$\begin{aligned} g_E(\Delta_{\mathbf{k}}) &= (1 + \mu_1^E + \mu_2^E) - (\frac{3}{8}\mu_1^E + \frac{9}{4}\mu_2^E)k^2L^2 \\ g_I(\Delta_{\mathbf{k}}) &= (1 + \mu_1^I + \mu_2^I) - (\frac{1}{4}\mu_1^I + \frac{9}{4}\mu_2^I)k^2L^2 \end{aligned} \tag{6.48}$$

and

$$B_{\mathbf{k}}^2 = \begin{pmatrix} \alpha_E\varphi + (1-\varphi)f_E & \cdot \\ \cdot & \alpha_I\phi + (1-\phi)f_I \end{pmatrix} \tag{6.49}$$

under the assumption that k is small, or that the continuum limit is taken, in which the lattice spacing goes to zero, while, Ω the number of patches, increases so as to keep the cortical area (or volume) constant.

6.5 The emergence of quasi-cycles

To calculate the conditions that lead to quasi-cycles, we compute the autocorrelation function $\langle \mathbf{x_k} \mathbf{x_k^\dagger} \rangle$ where $\mathbf{x_k}$ is the discrete Fourier transform of $\mathbf{x}$, and $\mathbf{x_k^\dagger}$ is the *adjoint* of $\mathbf{x_k}$. The

result is that the autocorrelation of the fluctuations about m_j can be written as

$$\langle \eta_k \eta_k^\star \rangle = \frac{\alpha_k + \beta_k \omega^2}{(\omega^2 - \Omega_k^2)^2 + \Gamma_k^2 \omega^2} \tag{6.50}$$

where

$$\begin{aligned} \alpha_k &= A_{22}(k)^2 \cdot 2\alpha_E \varphi_0 + A_{12}(k)^2 \cdot 2\alpha_I \phi_0 \\ \beta_k &= 2\alpha_E \varphi_0 \\ \Omega_k^2 &= A_{11}(k) A_{22}(k) - A_{21}(k) A_{12}(k) \\ \Gamma_k &= A_{11}(k) + A_{22}(k) \end{aligned} \tag{6.51}$$

where $\beta_{\mathbf{k}}$ is obtained from equation [6.35] evaluated at the fixed point (φ_0, ϕ_0), and the $A_{ij}(k)$ and B_i coefficients are obtained from equations [6.34] and [6.35] after a Fourier transformation.

The power spectrum of the fluctuations about m_j, (for real $\langle \eta_k \eta_k^\star \rangle$) is obtained simply as

$$P_E(k, \omega) = 2\langle \frac{\eta_k}{\sqrt{M}} \frac{\eta_k^\star}{\sqrt{M}} \rangle = \frac{2}{M} \frac{\alpha_k + \beta_k \omega^2}{(\omega^2 - \Omega_k^2)^2 + \Gamma_k^2 \omega^2} \tag{6.52}$$

This expression is written in the same format as those in [3] and [49]. There is a peak in this spectrum at approximately Ω_k. There is no corresponding peak in the mean-field power spectrum, since there are no fluctuations or oscillations about the mean-field fixed point, which is in this case, a stable focus. This is the basis for quasi-cycles.

6.6 Mean field conditions for pattern formation

The conditions for mean-field spatial pattern formation were worked out initially in [143], and in more detail in [144]. The main result is that the lateral inhibition of excitatory cells must be of longer range that that of self-inhibition, and the lateral excitation of inhibitory

cells must be of longer range than that of self-excitation. We repeat the analysis within the current formulation, by introducing an additional parameter in the model, the length scale of the intermodular excitation of inhibitory neurons. This means that the model must be modified so that instead of a single function $g_E(\Delta)$, we now have two functions, $g_{IE}(\Delta)$ and $g_{EE}(\Delta)$. The form of the currents is then

$$\begin{aligned} s_i^E &= g_{EE}(\Delta) w_{EE} \varphi_i - g_I(\Delta) w_{EI} \phi_i + h_i^E \\ s_i^I &= g_{IE}(\Delta) w_{IE} \varphi_i - g_I(\Delta) w_{II} \phi_i + h_i^I \end{aligned} \tag{6.53}$$

The fields φ and ϕ are dimensionless. We re-scale the couplings w so that

$$g(\Delta) w = w + g^1 w \Delta \tag{6.54}$$

This defines the coupling constants g^1. We assume that when the values of g^1 are varied, the coupling constants w are kept fixed through a rescaling of the original w coupling constants. In the continuum limit, the dimensions of the coupling constants g^1 are L^2. To go to a dimensionless description, we rescale length by $x \to \sqrt{g^1_{EE}} x$ where the new x is a dimensionless length. We also define the ratios $\frac{g^1}{g^1_{EE}} \to g^1$ This puts the currents in the form

$$\begin{aligned} s_E &= w_{EE} \left(1 + \Delta\right) \varphi - w_{EI} \left(1 + g^1_I \Delta\right) \phi + h_E \\ s_I &= w_{IE} \left(1 + g^1_{IE} \Delta\right) \varphi - w_{II} \left(1 + g^1_I \Delta\right) \phi + h_I \end{aligned} \tag{6.55}$$

In Fourier space the Laplacian operator $\Delta \to -k^2$, so we expand the coupling matrix

$A(\Delta_k) = A(k^2)$ as

$$\begin{aligned} A(k^2) &= A(0) + \frac{dA}{dk^2}k^2 \\ &= A_0 + k^2\delta A \end{aligned} \tag{6.56}$$

Explicitly, these matrices are

$$A_0 = \begin{pmatrix} -\alpha'_E + (1-\varphi)f'_E w_{EE} & -(1-\varphi)f'_E w_{EI} \\ (1-\phi)f'_I w_{IE} & -\alpha'_I - (1-\phi)f'_I w_{II} \end{pmatrix} \tag{6.57}$$

where we have also rescaled $w_{II}/\sqrt{\gamma} \to w_{II}$ and $w_{EI}/\sqrt{\gamma} \to w_{EI}$ and defined $\alpha' \equiv \alpha + f$. The matrix δA is

$$\delta A = \begin{pmatrix} -(1-\varphi)f'_E w_{EE} & (1-\varphi)f'_E g^1_I w_{EI} \\ -(1-\phi)f'_I g^1_{IE} w_{IE} & (1-\phi)f'_I g^1_I w_{II} \end{pmatrix} \tag{6.58}$$

Note that the matrix B is unaffected by these rescalings.

To obtain the conditions for mean field pattern formation, we must obtain the eigenvalues of $A(k)$. Mean-field pattern formation occurs when the eigenvalues of $A(0)$ are negative, but for some finite range of $k \neq 0$, the eigenvalues become non-negative [16]. The eigenvalues are given by the equation

$$\lambda_\pm = \frac{1}{2}\left[TrA \pm \sqrt{(TrA)^2 - 4\det A}\right] \tag{6.59}$$

But $\lambda_+ \geq \lambda_-$, so we need only investigate λ_+ to find the transition to pattern formation.

It is clear from the eigenvalue formula that a necessary condition for spatial pattern formation is $\det A \leq 0$, for then $\lambda_+ \geq 0$. From equation [6.54] we can expand $\det A$ as:

$$\det A = \det A_0 + \Delta A\, k^2 + \det \delta A\, k^4 \tag{6.60}$$

where $\Delta A = \delta A_{11} A_{22} + A_{11} \delta A_{22} - \delta A_{12} A_{21} - A_{12} \delta A_{21}$.

It follows that we can obtain the necessary condition for spatial pattern formation, $\det A \leq 0$ over a finite range of wave numbers k if and only if

$$\begin{aligned} \Delta A &< 0, \\ \det \delta A &> 0 \\ 4 \det A_0 \det \delta A &< (\Delta A)^2 \end{aligned} \tag{6.61}$$

These conditions are obtained from seeking zeros of the biquadratic equation in k 6.60.

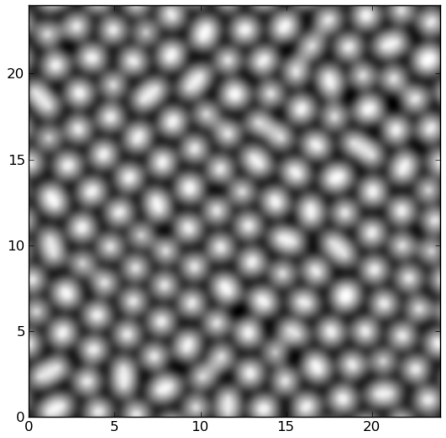

Figure 6.7: Spatial heterogeneity of neuronal excitation from the Wilson-Cowan equations.

To explore the qualitative nature of the patterns generated at mean field and beyond mean field, we solved the mean field equations numerically in the pattern forming regime using a simple forward Euler scheme. The resulting patterns are shown in fig. 6.7. To compare directly to reported visual geometric hallucinations as in fig.6.1 we transformed fig. 6.7 into visual field coordinates using the mapping between visual field and visual cortex coordinates in Eq. 6.1. The resulting image from numerical solution in fig. 6.8 is strikingly similar to a reported form constant as in fig. 6.1. Prediction of all four fundamental form

constants can be achieved by by including the full symmetry of the visual cortex into the model [8].

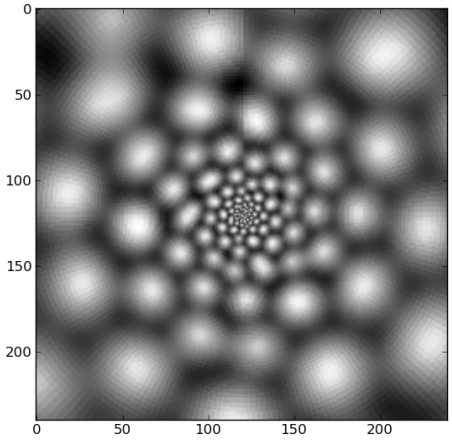

Figure 6.8: Mean field patterns of neural activity in visual field coordinates. Note the close resemblance to reported geometric visual hallucinations.

6.7 Pattern formation beyond mean field

To calculate the conditions under which patterns form beyond mean field, we follow [13]. We assume that mean field solutions are stable and constant. We set $\omega = 0$ and examine the power spectrum in the form:

$$P_E(k,0) = \frac{2}{M} \frac{A_{22}(k)^2(2\alpha_E\varphi_0) + A_{12}(k)^2(2\alpha_I\phi_0)}{|\det(A(k))|^2} \tag{6.62}$$

where $A(k,0)$ is strictly real.

To obtain the conditions for pattern formation, we note that the denominator of this equation will grow as the eighth power of k in the large k limit, so the power spectrum will be a decreasing function of k for large k. Thus a sufficient condition for pattern formation (corresponding to a non-zero peak in the power spectrum as in fig. 6.9) is that $dP_E/dk^2 > 0$.

To evaluate the conditions under which this inequality is satisfied, we examine the derivative

$$\frac{d}{dx}\frac{f}{g^2} = \frac{f'g - 2fg'}{g^3} \tag{6.63}$$

Dropping constant factors and taking $f = A_{22}(k)^2 B_1^2 + A_{12}(k)^2 B_2^2$ and $g = \det(A(k))$ (to satisfy the sufficient condition for pattern formation), we find the condition for fluctuation driven pattern formation in excitatory neurons to be

$$\begin{aligned}\left(A_{22}^0 \delta A_{22} B_1^2 + A_{12}^0 \delta A_{12} B_2^2\right) \det(A^0) \\ > \left({A_{22}^0}^2 B_1^2 + {A_{12}^0}^2 B_2^2\right) \Delta A\end{aligned} \tag{6.64}$$

We also want the criteria for fluctuation driven pattern formation in the inhibitory neuron population. To obtain this, we note that the power spectrum for inhibitory fluctuations is

$$P_I(k,0) = \frac{2}{M}\frac{A_{21}(k)^2(2\alpha_E\varphi_0) + A_{11}(k)^2(2\alpha_I\phi_0)}{\left|\det(A(k))\right|^2} \tag{6.65}$$

Through simple substitution of the explicit subscripts $22 \to 21$ and $12 \to 11$ we obtain the conditions for pattern formation in the inhibitory neurons as

$$\begin{aligned}\left(A_{21}^0 \delta A_{21} B_1^2 + A_{11}^0 \delta A_{11} B_2^2\right) \det A_0 \\ > \left({A_{21}^0}^2 B_1^2 + {A_{11}^0}^2 B_2^2\right) \Delta A\end{aligned} \tag{6.66}$$

This condition differs slightly from that for excitatory neurons. For completeness, we next obtain the conditions for quasi-cycles by setting $k = 0$ and retaining ω^2. Through identical calculations to those above with the independent variable k^2 replaced by ω^2, the condition for quasi-cycles is obtained

$$\beta_k \det A_0 > \alpha_k(Tr^2 A_0 - 2\det A_0) \tag{6.67}$$

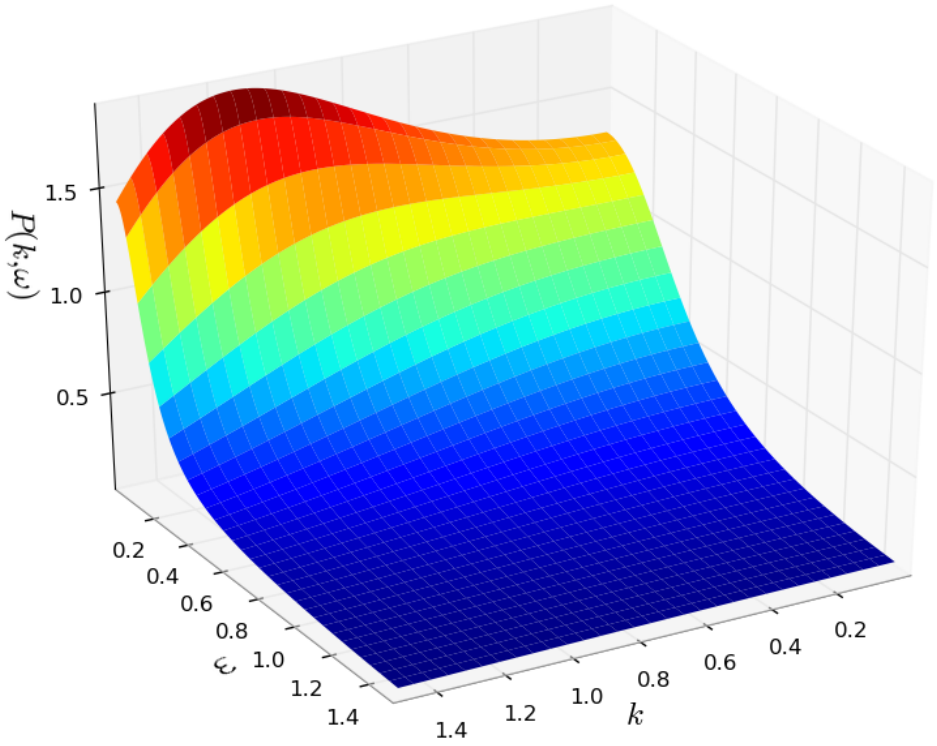

Figure 6.9: The power spectrum as a function of k and ω in the quasi-pattern regime, without quasi-cycles. The scaled parameters are $g_{IE}^1 = 3.0$ and $g_I^1 = 0.2$.

with $\beta_k > 0$ and $\alpha_k > 0$ as defined in equation [6.51] (with $k^2 = 0$).

6.8 Pattern formation with inhibition forbidden in the intermediate length scale

The most important case anatomically is the case where there are no inhibitory synapses at the intermediate length scale. In this case, the parameters satisfy $g_{EE}^1 \gg g_I^1$ and $g_{IE}^1 \gg g_I^1$. Under these conditions the left hand side of equation [6.64] is order g_I^1, and so its magnitude is much less than the right hand side, which is of higher order. So, to achieve fluctuation driven pattern formation, to order g_I^1 the requirement is $\Delta A < 0$.

From the explicit form of the explicit determinant of δA

$$\det \delta A = (1-\varphi)(1-\phi) f_I' f_E' g_I^1 [g_{IE}^1 w_{EI} w_{IE} - g_{EE}^1 w_{EE} w_{II}] \tag{6.68}$$

the lhs in the last equation of the mean field pattern formation conditions Eq. 6.61 is $(O(g_I^1))$. ΔA is not generally small. This implies that in the case of no intermediate inhibitory connections, the requirement for mean-field pattern formation is $O(g_I^1) < \Delta A^2$, Thus, in this case, the conditions for fluctuation driven and mean field pattern formation are identical to order g_I^1, provided $\det \delta A > 0$.

To derive the condition explicitly, terms in Eq. 6.64 of order g_I^1 can be neglected, whence

$$\Delta A = g_{IE}^1 (1-\phi) f_I' w_{IE} A_{12} - g_{EE}^1 (1-\varphi) f_E' w_{EE} A_{22} < 0 \tag{6.69}$$

Rearranging yields

$$\frac{g_{IE}^1}{g_{EE}^1} > \frac{w_{EE}\left[\alpha_I' + (1-\phi) f_I' w_{II}\right]}{w_{EI}(1-\phi) f_I' w_{IE}} + O\left(\frac{g_I^1}{g_{EE}^1}\right) \tag{6.70}$$

This equation describes the conditions under which patterns can form, provided there are no intermediate length scale inhibitory-inhibitory connections.

6.9 Scale-free spatial effects

In addition to spatial pattern formation, the stability matrix A yields eigenvalues that increase without bound as a function of k. This corresponds not to pattern formation, but most likely to scale-free spatial effects,. This possibility arises because there are several ways that the eigenvalue can become positive for non zero k. The first way is for TrA to become positive. In the standard formulation of Turing instabilities, this is forbidden, but in the model currently being investigated, this is a possibility. The key physical difference is that in the standard Turing models, diffusion of inhibitor into a neighboring cell increases the

local concentration of inhibitor. This is not the case for neural inhibition, which reduces the amount of activity in neighboring inhibitors.

The formula for $TrA(k)$ is

$$TrA = TrA_0 + \left((1-\phi) f_I' g_I^1 w_{II} - (1-\varphi) f_E' w_{EE}\right) k^2 \tag{6.71}$$

Thus, if

$$\frac{g_I^1}{g_{EE}^1} > \frac{(1-\varphi) f_E' w_{EE}}{(1-\phi) f_I' w_{II}} \tag{6.72}$$

λ_+ increases without bound as a function of increasing k (note we have restored g_{EE}^1 for clarity). An additional condition for unbounded growth of λ_+ is can be obtained by noting that if

$$\det \delta A < 0 \tag{6.73}$$

then the eigenvalue will increase as k^2 as well.

We conjecture that these conditions might provide a novel mechanism for the generation of neuronal activity avalanches [133]. A plot of eigenvalues of A in both the mean field pattern forming phase and the scale invariant phase is shown in fig. 6.10

Further research is required to fully characterize the spatio-temporal dynamics of the "scale invariant" phase.

6.10 Phase diagram when inhibition is introduced at the intermediate length scale

As noted above, in the visual cortex (and in the neocortex more generally) inhibition over the intermediate length scale is forbidden. Anatomical studies cannot explain why the neural architecture is configured in this way. However, the current model can be modified to see how its qualitative behavior would differ if intermediate length scale inhibition were

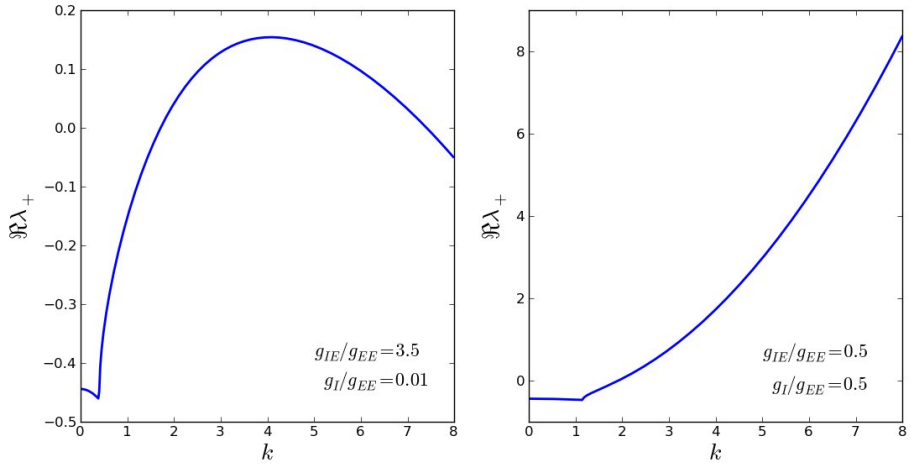

Figure 6.10: The real part of the largest eigenvalue of the stability matrix as a function of k in phases V and I respectively.

introduced. While the model is a highly simplified caricature of the real neocortex, the qualitative features that emerge when intermediate inhibition is introduced are likely to illuminate the real changes in function that would occur in the brain if such couplings existed. In this section, we report the behavior of the system as a function of the longest length scale of inhibitory and excitatory connections.

When intermediate length scale inhibition is introduced, the neocortex is overwhelmed by spontaneous, apparently scale free spatiotemporal dynamics. Analytically, this can be seen in Eq. 6.72, which shows that if g_I^1 is sufficiently large, the system spontaneously generates spatial structure. If intermediate length scale inhibition is forbidden, the model has rich phase behavior, including scale free, homogeneous steady states, and pattern formation. In the visual cortex in particular, there must be a stable homogeneous phase so that sensory data can determine the excitation patterns of the neurons rather than internally generated activity.

To study these behaviors systematically, we explored the above conditions numerically. The result is that increasing the length scale of inhibition increases substantially the likelihood that intrinsic fluctuations will trigger the formation of spatial structures. Figure [6.11]

shows the phase diagram.

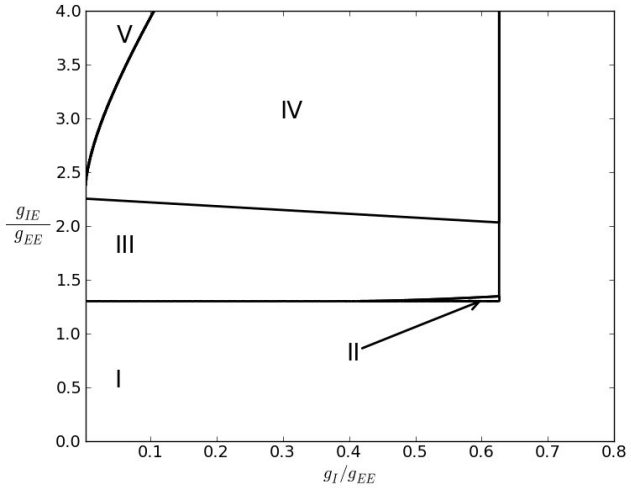

Figure 6.11: Phase diagram for pattern formation. This figure shows the conditions for the emergence of spatial structure. The x axis is g_I^1/g_{EE}^1 and the y axis is g_{IE}^1/g_{EE}^1. Region I is scale invariant structure at the mean field level, region II is a mean field pattern formation phase, region III is a homogeneous steady state phase, region IV is the quasi-pattern phase and region V is a mean field pattern forming phase.

Based on the phase diagram in fig. 6.11 we conjecture that the lack of long range inhibition in the visual cortex is connected with the normal spatial stability of the visual cortex to intrinsic fluctuations. As noted above, the spatial stability of the visual cortex required for normal visual function cannot be maintained with intermediate range inhibition.

6.11 Conclusion

In this chapter, we showed that quasi-patterns and quasi-cycles are both present in a minimal model of the visual cortex, corresponding to temporal oscillations of brain activity and to visual hallucinations. We also showed that there is novel scale invariant phase. The most interesting finding, however, is that the previously unexplained network structure of the visual cortex, with intermediate length scale inhibition forbidden, can be understood as resulting from the requirement that the visual cortex not be dominated by spontaneous

activity so that it can represent sensory input from the eyes. Forbidding intermediate scale inhibition makes the network much more flexible in its phase behavior and much more robust to unwanted spontaneously generated spatiotemporal behavior while still allowing the full range of dynamical behavior if desired.

While the goals of this research were primarily fundamental understanding of the nature of pattern formation in the brain and the robustness of the visual cortex, practical applications are also possible. It is easy to imagine that in silico neural networks used for optical identification software, etc. might be engineered to be more failure resistant by ensuring that inhibitory connections are sufficiently local, as in the real anatomy. We also conjecture that in more general complex networks, locality (defined appropriately) of outgoing links with roles analogous to inhibition, (such as repressor proteins in gene regulatory networks) [145], may be a more general network motif useful for controlling the emergence of unwanted internally generated network dynamics in order to preserve a wider variety of network behaviors.

Chapter 7

Sloppiness and robustness in a model circadian rhythm circuit

7.1 The problem of biological robustness

The interior of a cell contains thousands of different species of proteins, each with highly varying numbers of copies dependent on expression level, transcriptional noise, or other factors [145]. Such an environment is extremely noisy due to both intrinsic noise from fluctuations in copy number of proteins and other molecules as well as extrinsic noise from the environment [146, 147]. Additionally, differences in temperature, point mutations, and many other factors lead to differences in the kinetic parameters of the protein and gene interaction networks between organisms of the same species and in the same organism in different environmental conditions [148]. In spite of the enormous amount of variability in cell conditions suggested by the above considerations, the cell is able to carry out many crucial functions with precision [147, 149].

The ability of the cell to maintain function in the face of noise and variations in parameters is known as robustness (for a review, including several definitions, see [150]). Robustness has been investigated in many different contexts, and can arise from a variety of mechanisms, such as active control from feedback loops, redundancy, modularity, etc. and is widely believed to be an evolved feature of biological networks [150, 147]. An alternative line of investigation has focused on robustness to changes of system parameters as a universal emergent property of systems with many different parameters [148]. From this perspective, not all robustness is necessarily evolved. Rather, some systems are robust simply as a consequence of the fact that biological networks are high dimensional dynamical systems

with many governing parameters and the quantitative behavior of such systems often only depends weakly on the values of the underlying parameters. This property of high dimensional dynamical systems is called "sloppiness" and models that have it are called "sloppy models"[148, 151, 152, 153]. Evidence that sloppiness is not an evolved property is provided by examples from non biological natural systems from physics [152]. Both evolution and the high dimensionality of biological dynamics likely contribute to the robustness observed in real biological systems.

In this chapter we consider a model of circadian rhythms that is known to be sloppy [6, 151] and probe the sensitivity of the essential function of the network (maintaining a fixed period oscillation) to large changes of a simple, easily interpreted biological parameter. The goal of this calculation is to shed further light on the origins of robustness and sloppiness by examining their relationship in the context of a simple model. This is useful for bridging a currently existing gap between the study of sloppy models, which has tended to analyze network properties alone as a function of parameter variation without attention to the biological interpretation of the parameters varied [151] and a large literature that investigates the origin of robustness in simple, well characterized circuits [150]. We begin by summarizing some key results in the study of sloppy models. Next we will introduce a simple model of circadian rhythms [6] which we will analyze and demonstrate that the robust function of the circadian oscillator depends crucially on unpromoted mRNA production, but that the crucial dependence is, contrary to intuition, also robust to changes in the rate of unpromoted mRNA production over more than an order of magnitude. We will conclude with a discussion of the possible relationship between sloppiness and feedback generated robustness illustrated by this calculation.

7.2 Sloppy models

The theory of sloppy models is built around the observation that in dynamical systems, such as gene regulatory networks, with large numbers of governing parameters, the quantitative predictions of the models are insensitive to all but a few combinations of parameters [153]. This behavior is essentially universal across many parameter systems biology models [151] and also occurs in physics models with many parameters [152]. Sloppiness is observed empirically in models starting at around 8 parameters [152]. Specific calculations indicate that the sloppiness property can explain otherwise perplexing robustness phenomena, such as the fact that circadian rhythm of cyanobacteria is fairly constant over temperatures ranging from 25-35 degrees C while the Arrhenius law indicates that the period of circadian rhythms might be naïvely expected to change by a factor of two [148].

Sloppiness has been conjectured to have its origin in subspaces of degenerate parameters [152]. Specifically, if the operator P is a permutation operator that takes a given configuration of parameters $\boldsymbol{p}$ and permutes the parameters of the system to a new configuration $\boldsymbol{p}'$ it can be shown under specific circumstances that if the deviation from the desired behavior of the model, $C(\boldsymbol{p})$, has the property that

$$C(P\boldsymbol{p}) = C(\boldsymbol{p}) \tag{7.1}$$

then the model will with high likelihood be sloppy, and therefore robust. Since invariance in the deviation under permutation of parameters is an extremely restrictive requirement certainly not met for most complex systems, it is believed that more general complex systems have subspaces in parameter space that satisfy the requirement in Eq. 7.1. The stiff directions correspond to parameter combinations that mix these subspaces [152, 151]. This has led to some proposals to use sloppiness as a systematic tool for model reduction in systems biology, though such attempts have not yet been developed in detail [152]. Further progress requires identification of the origins of sloppiness for specific parameters of a known sloppy

model, where the findings can be interpreted biologically. To carry out such an analysis, we introduce a simple model of circadian rhythms that is known to be sloppy [6, 151].

7.3 A simple model of circadian rhythms

Circadian rhythms in many organisms share essential features of feedback and control [10]. These shared features include negative feedback, which generates oscillations, coupled to positive feedback that acts as a delay on the negative feedback, resulting in robust limit cycles [6, 10]. These features are also common in other biological cycles [154]. Vilar and coworkers formalized these features into a generic model of circadian rhythms. The model

contains the following 15 reactions

$$
\begin{aligned}
D_A + A &\xrightarrow{\gamma_A} D'_A \\
D_R + A &\xrightarrow{\gamma_R} D'_R \\
D'_A &\xrightarrow{\theta_A} A + D_A \\
D'_R &\xrightarrow{\theta_R} A + D_R \\
D_A &\xrightarrow{\alpha_A} M_A + D_A \\
D'_A &\xrightarrow{\alpha'_A} M_A + D'_A \\
M_A &\xrightarrow{\beta_A} A + M_A \\
A + R &\xrightarrow{\gamma_C} C \\
A &\xrightarrow{\delta_A} \emptyset \\
D'_R &\xrightarrow{\alpha'_R} M_R + D'_R \\
D_A + A &\xrightarrow{\gamma_A} D'_A \\
D_R &\xrightarrow{\alpha_R} M_R + D_R \\
M_R &\xrightarrow{\delta_M} \emptyset \\
M_R &\xrightarrow{\beta_R} R + M_R \\
C &\xrightarrow{\delta_A} R \\
R &\xrightarrow{\delta_R} \emptyset
\end{aligned}
\tag{7.2}
$$

with reactants labeled in the table 7.1

The key element of the circuit is the activator protein, denoted by A, which binds to the promoter region for its own corresponding gene (D_A) as well as the promoter region for a gene D_R corresponding to a repressor protein R. Thus when A is produced, its own production increases, as does the production of R. R binds to A to form a complex C, which

D_A	activator promoter
D'_A	activator promoter with activator bound to it
D_R	repressor promoter
D'_R	repressor promoter with activator bound to it
M_A	mRNA for activator
M_R	mRNA for repressor
C	complex formed from binding of activator and repressor

Table 7.1: Reactants for circadian rhythm model from [6]

inhibits A from binding to the promoters. This reduces the amount of A in the system through degradation and decreased production of A. Cycles emerge as follows:

1. The initial A proteins bind to their promoter leading to an increase in A copy number

2. The activator A also binds to the promoter for the repressor R leading to an increase in R copy number

3. The repressor R binds A into the complex C, so that it can no longer promote itself. This step involves delay because R must first be produced.

4. Copy numbers of R and A both drop through degradation

5. The decline in R allows the auto-promotion of A to resume, returning the system to step 1.

This picture is confirmed in both stochastic and deterministic simulations (see fig 7.1) that generate limit cycles with a period of around 24 hours with biologically realistic parameters [6].

As noted above, this is a 15 parameter model known to be sloppy [151]. The next step in the analysis is to specifically explore the nature of the sloppiness in the model to see if insight can be gained into the mechanisms from which it arises.

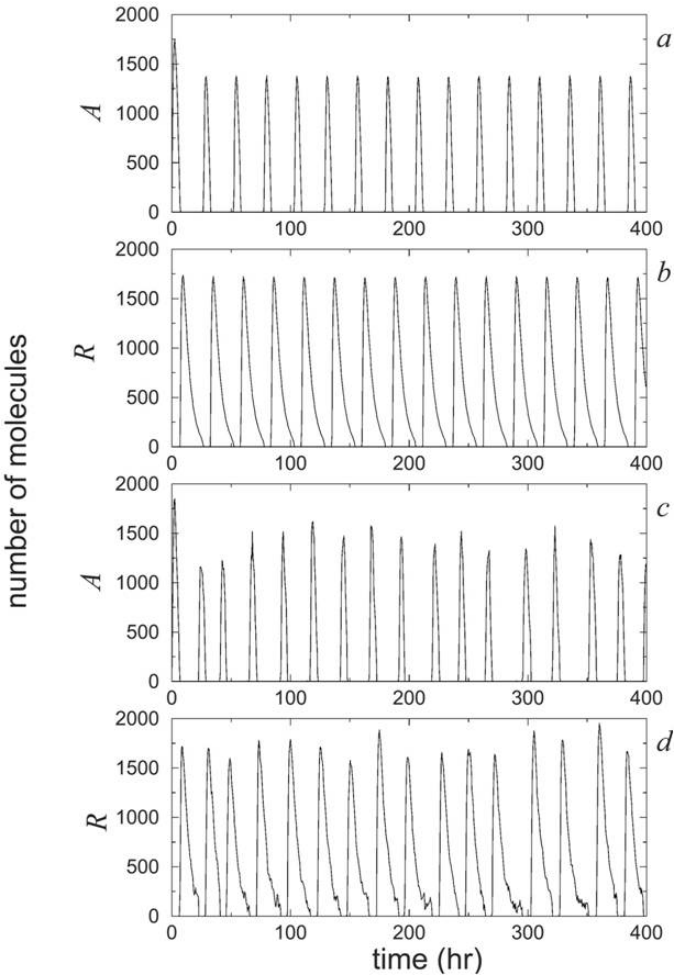

Figure 7.1: 24 hour circadian cycles in A and R from the reactions in Eq. 7.2. Figure taken from [6].

7.4 Robustness in the model circadian rhythm

For the circadian rhythm model presented above, the definition of robustness can be given a precise mathematical formulation: that the period predicted does not change substantially when the parameters are varied, or under the influence of noise. Parameters that can be varied without affecting the period are sloppy parameters. Systematically varying the parameters of the system will give insight into whether or not the sloppiness corresponds in some meaningful way to real biological robustness, as well as into the mechanisms leading both to sloppiness and to robustness. We will show that in the circadian rhythm model,

there is a parameter that when varied over orders of magnitude does not affect the dynamics of the system. However, while one might believe that this means that the model can be reduced by removing the parameter, in fact the parameter is essential, and if it is removed, the system fails. This result clarifies a few key points about sloppy models and robustness: First, it demonstrates that sloppiness can emerge from things other than the degeneracy of parameters as theorized above. Second, it shows that attempts to use sloppiness for model reduction must be carried out with care, as extremely sloppy parameters can still be essential to the performance of the system. Third, as will be shown below, it demonstrates one way the results of a model calculation with differential equations can fail at low copy number. Finally, as will also be shown below, it shows how natural biological oscillators exploit intrinsic limitations in gene regulation as a passive safeguard against failure from excessive negative feedback.

To arrive at these results, we vary a parameter corresponding to unpromoted production of mRNA for the production of A. In the reactions above, this corresponds to the rate α_A. To compute the effect of varying α_A, we compute the trajectory of the system for each value and obtain the power spectrum for A. We compute the trajectory in two different ways: First, we compute it deterministically by straightforward integration of deterministic differential equations corresponding to the reactions in Eq. 7.2 and written out explicitly in [6]. Second we compute the trajectory of the stochastic process corresponding to the reactions in Eq. 7.2. The stochastic description keeps an explicit count of the absolute number of each type of molecule and selects the next reaction to take place and the time the next reaction will take place from a rigorously derived exact distribution of probabilities. The algorithm for carrying out this simulation is called the Gillespie algorithm, and it corresponds to a master equation description of the dynamics [120]. From the power spectra corresponding to each trajectory, we obtain the period T from the peak frequency ω_P as $T = 2\pi/\omega_P$. Iterating the process for both the stochastic and deterministic descriptions we obtain the period of the cycle as a function of α_A for the deterministic and stochastic cases. The results for $\alpha_A > 0$

are shown in fig. 7.2

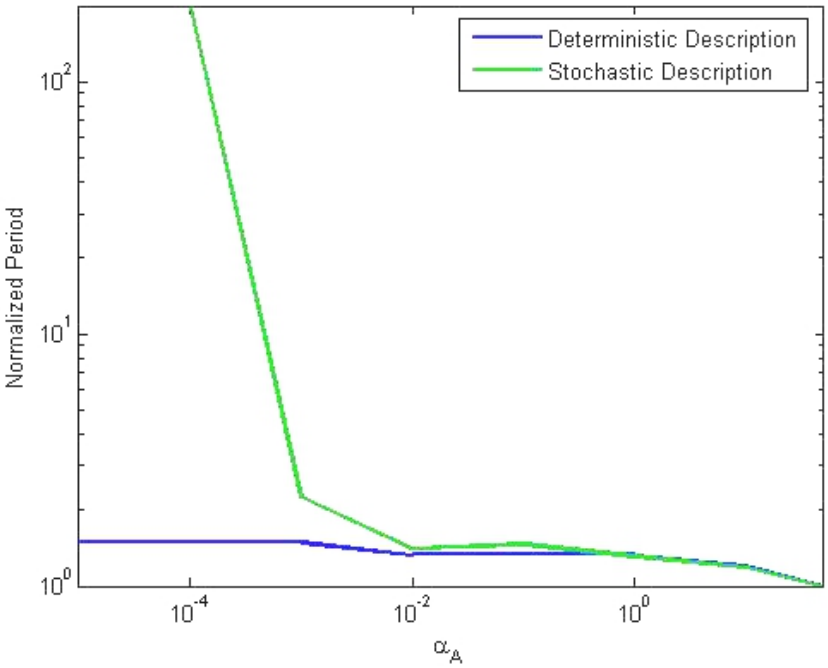

Figure 7.2: Change in the period of the circadian rhythm as a function of α_A, or unpromoted mRNA production for the gene for A. The y axis is the normalized period, meaning that all periods have been divided by 24 hours. Note that the period stays roughly constant over three orders of magnitude in the stochastic description, and over at least two more orders of magnitude in the deterministic description. Not shown is the failure of the deterministic description at $\alpha_A = 0$, where the period diverges to infinity.

The results have several features of interest. First note that in both the deterministic and stochastic description, the period is insensitive to more than an order of magnitude variation in α_A. Thus α_A is clearly a sloppy parameter. But it is also a necessary parameter. While the system continues to function well when the parameter is extremely small relative to other parameters (typical small parameters in the system are of order 0.2/hr), eliminating the parameter leads to total failure of the system. This implies that leaky mRNA production for the gene corresponding to A is essential for the functioning of the oscillator, but the rate of production is nearly irrelevant. This can be understood by noting that the oscillations in fig. 7.2 use negative feedback to drive the number of activator proteins in the system to

very low counts at the end of each cycle. This is dangerous, because a non zero quantity of activator proteins is required to start each cycle (see step one in the description of the cycles above). The system is protected from this failure mode by the leaky production of activator proteins, which therefore act as a passive fail safe mechanism for circadian rhythms. This is interesting, because leaky transcription might generally be considered an engineering limitation. But in the circadian system under consideration, it is exploited to increase the robustness of the circuit.

Also, it is clear that α_A is not a redundant parameter. Redundancy would imply that the parameter could be set to zero without the catastrophic failure of the system. This makes it unlikely that this parameter is part of a a subspace of parameters where the deviation from the desired behavior of the system, $C(\boldsymbol{p})$, is invariant under permutation of parameter values. Thus the origin of its sloppiness is unlikely to be explained by the theory outlined above [152].

7.5 Comparison of stochastic and deterministic approaches

One key feature of fig. 7.2 is that the deterministic and stochastic descriptions have dramatically different predictions for low values of α_A. Robust cycles are maintained over a much broader range of α_A in the deterministic description than in the stochastic description (see fig. 7.2). This can be explained by noting that due to the strength of the positive feedback on the activator protein, the cycle requires only a very small amount of activator protein to start. In the deterministic description, that small amount can be so small that in a real cellular environment, it would correspond to an unphysical fraction of an activator protein. In the stochastic version, such unphysical fractions of activator proteins are forbidden since the absolute number of activator proteins is modeled. This means that if α_A is sufficiently small, then the system must wait until an mRNA is produced and then transcribed into an

activator protein to start a new cycle. If the waiting time is sufficiently long, the cycles are no longer regular. In this case, it is clear that noise from the stochasticity of the system disrupts the tendency of the system to oscillate

Interestingly, in other regions of parameter space, the system responds to stochasticity in exactly the opposite way, as was noted in the original publication that introduced the model [6]. Choosing the bifurcation parameter δ_R, which corresponds to the degradation rate of the repressor protein R, we carry out the same calculation as above of the period as a function of the parameter δ_R for both the stochastic and deterministic representations of the model. The results, as seen in fig. 7.3, show that stochastic trajectories oscillate over a much larger range of parameters than in the deterministic version.

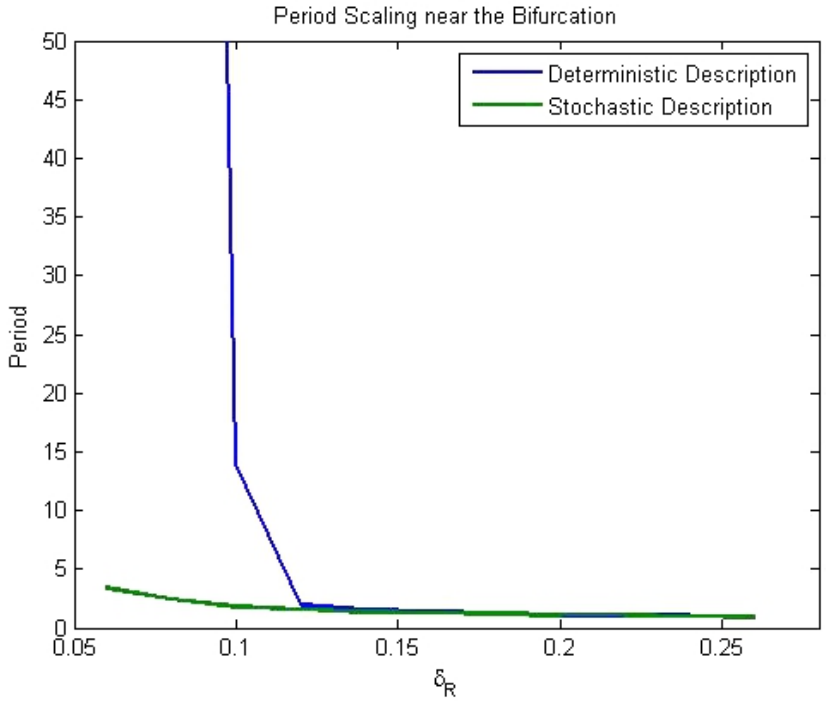

Figure 7.3: Period of stochastic and deterministic trajectories near a bifurcation to a stable steady state in the deterministic system. Note that the stochastic version continues to have oscillations over a much larger range of the bifurcation parameter. Note also that the period is much too short (around 5, versus the needed 24) near the bifurcation point.

7.6 Conclusions

In this chapter we have shown that circadian rhythms use leaky mRNA production as a fail safe mechanism to improve the robustness of the oscillations. We also have shown that sloppiness can arise from origins other than permutation symmetries in the parameters. The findings in this chapter emphasize the key point that robustness in biological systems can have many origins, ranging from emergent properties of high dimensional dynamical systems and evolution of active feedback mechanisms, to engineering limitations (such as the inability to completely turn off translation as in the above) exploited as passive fail safe mechanisms.

Chapter 8

Conclusions

In this dissertation, I presented my findings on the evolution of the genetic code, robustness of circadian rhythms, and the emergence of fluctuation driven patterns in space and time in ecosystems and the visual cortex. Work on fluctuation driven pattern formation led me to introduce a new emergent state of pattern forming systems, the quasi-pattern state. I analyzed the quasi-pattern state in the context of a simple generic model of predator-prey populations, and then explored the emergence of this state in a much less trivial model of the visual cortex.

The presence of quasi-patterns in models of the brain, predator-prey interactions, and in the Brusselator model of oscillatory chemical reactions (worked out by Biancalani et al. in [112]) indicate that the quasi-pattern state is a long wavelength emergent state of matter that is independent of many of the idiosyncratic details of the underlying system. The quasi-pattern state emerges in the parameter regions far more likely to be important for experimental application than the standard pattern forming state present in mean field theories of pattern formation and can be distinguished from it experimentally, which led to the conjecture that many observed patterns previously interpreted as mean field patterns are actually quasi-patterns. This conjecture will be tested in future work.

8.1 Final thoughts on the role of theory in biology

In the introduction to his book on model ecosystems, the distinguished theoretical physicist turned ecologist Robert May makes a distinction between what he calls “tactical” modeling

of the specific interactions in a particular ecosystem of interest that is being concurrently subjected to field studies, and general "strategic modeling," carried out to obtain broader insight into the dynamics of ecosystems without attempt to match in detail the features of an individual ecosystem [11]. After noting his own primary interest in strategic modeling, May reflects, or perhaps more accurately laments, that in ecology, tactical modeling has been more successful than strategic modeling.

Most of the work contained in this dissertation is firmly in the strategic mold, so given the picture of biological modeling in the previous paragraph it is not entirely surprising that when I interviewed for jobs, I encountered one prominent biophysicist, who upon hearing my lecture on the emergence of quasi-patterns said to me: "All of this is quite interesting, but I don't think you can do much with these kinds of toy models." I disagreed with him, and think I answered to his satisfaction by identifying experimental systems, such as pattern forming plant resource systems [15] that I believed could not be understood at the "tactical" level without consideration of the quasi-pattern phenomena.

The grounds for my confidence that broad strategic models are essential to understand real ecosystems or biological systems more generally, is based on the rather different history of theory in physics. Contrary to ecology, the renaissance in condensed matter physics that started in the 1950's resulted on the theoretical side primarily from results on strategic models such as Onsager's 1944 exact solution of the two dimensional Ising model [155], and the BCS theory of superconductivity [156]. These models are so firmly in the strategic class that not only do the above cited papers on them only mention actual compounds in passing, but they have been applied to hundreds of different materials. In physics, it is clear that contrary to ecology or biology more generally, "strategic" models have been far more successful than tactical models and have led directly to an impressive synthesis of theory and experiment envied but not duplicated in any other field of science, including biology.

With the lessons of the renormalization group, scaling, and universality in hand, a program to replicate the success in physics of minimal models in biology is not difficult to

imagine. It will require asking the type of detail independent questions in the lab and in the field that can be answered by minimal models in a sharp and quantitative manner. Additionally, minimal models must, as the Onsager solution and the BCS theory from condensed matter physics do, make sharp experimental predictions valid at length and time scales that liberate the data from extreme dependence on the idiosyncratic details of the system being studied. Examples of minimal model based analysis applied quantitatively to experimental data in biology are emerging at an increasing rate (a personal favorite is the observation of a percolation transition in plague epidemics in great gerbil populations [108]). Such predictions are made in the chapter of my thesis on the statistical mechanics of the Turing mechanism, and future work will apply these results to satellite images of plant patterns.

But biology isn't the same as physics. Taking the example of plant patterns, I recall from the introduction the observation that to an ecologist interested in the ecosystem of stripe forests in Colorado, it may be precisely the details my analysis ignores that make the system of interest to her. Such differences in priority and interest should not be discouraging to the physicist. It is just a reminder that in biology, there is richness in the diversity of details as well as in the universal emergent states.

Appendix A

MSR response function formalism

We briefly review the Martin Siggia Rose response function formalism (MSR formalism) below [89, 90]. For a more complete discussion, see for example, [74]. The MSR formalism converts a traditional Langevin equation into a path integral. A first order langevin equation is given by

$$\frac{\partial \boldsymbol{S}(x,t)}{dt} = F(\boldsymbol{S}(x,t)) + \boldsymbol{h}(x,t) + \boldsymbol{\xi} \tag{A.1}$$

where $\langle \xi_i(x,t)\xi_j(x',t')\rangle = B_{ij}\delta(t-t')\delta(x-x')$, i.e. $\boldsymbol{\xi}$ is standard delta correlated gaussian white noise. To convert this equation into a path integral, we write the equation as a functional integral over a delta function

$$\int \mathcal{D}\boldsymbol{S}\delta(\partial_t \boldsymbol{S} - F(\boldsymbol{S}) - \boldsymbol{h} - \boldsymbol{\xi}) \tag{A.2}$$

Functions of the solution can now be written

$$G(\boldsymbol{S}) = \int \mathcal{D}\boldsymbol{S'}G(\boldsymbol{S'})\delta(\partial_t \boldsymbol{S'} - F(\boldsymbol{S'}) - \boldsymbol{h} - \boldsymbol{\xi}) \tag{A.3}$$

and in particular,

$$\boldsymbol{S} = \int \mathcal{D}\boldsymbol{S'}\boldsymbol{S'}\delta(\partial_t \boldsymbol{S'} - F(\boldsymbol{S'}) - \boldsymbol{h} - \boldsymbol{\xi}) \tag{A.4}$$

We next write the delta funtion as a functional integral over the purely imaginary fields $\tilde{\boldsymbol{S}}$, called "response fields" (ignoring multiplicative constants)

$$\delta(\partial_t \boldsymbol{S} - F(\boldsymbol{S}) - \boldsymbol{h} - \boldsymbol{\xi})$$
$$= \int \mathcal{D}\tilde{\boldsymbol{S}} \exp\left[- \int dt d^d x \tilde{\boldsymbol{S}} \cdot (\partial_t \boldsymbol{S} - \boldsymbol{F}(\boldsymbol{S}) - \boldsymbol{h} - \boldsymbol{\xi}) \right] \tag{A.5}$$

The noise is Gaussian distributed, with the probability of a given realization proportional to

$$P(\boldsymbol{\xi}) = \exp\left[- \int dt d^d x \frac{1}{2} \boldsymbol{\xi}^T B^{-1} \boldsymbol{\xi} \right] \tag{A.6}$$

Averaged quantities are therefore given by

$$\langle G(\boldsymbol{S}) \rangle =$$
$$\int \mathcal{D}\boldsymbol{\xi} \mathcal{D}\boldsymbol{S'} \mathcal{D}\tilde{\boldsymbol{S}} G(\boldsymbol{S'}) exp\Big(- \int dt d^d x \tilde{\boldsymbol{S}} \cdot (\partial_t \boldsymbol{S'} - \boldsymbol{F}(\boldsymbol{S'}) - \boldsymbol{h} - \boldsymbol{\xi}) - \int dt d^d x \frac{1}{2} \boldsymbol{\xi}^T B^{-1} \boldsymbol{\xi} \Big) \tag{A.7}$$

This expression can be further simplified by noting that we can explicitly carry out the integral over the noise using the Gaussian integration identity (up to multiplicative constants)

$$\int \mathcal{D}\boldsymbol{\xi} \exp\left[- \int dt d^d x \frac{1}{2} \boldsymbol{\xi}^T B^{-1} \boldsymbol{\xi} - \boldsymbol{A} \cdot \boldsymbol{\xi} \right] = \exp \frac{1}{2} \boldsymbol{A}^T B A \tag{A.8}$$

Applying this identity and adding a linear coupling $\boldsymbol{J}$ for generating moments of $\mathbf{S}$, we obtain a path integral representation of the Langevin equation A.1

$$Z[\boldsymbol{J}, \boldsymbol{h}] = \int \mathcal{D}\boldsymbol{\xi} \mathcal{D}\boldsymbol{S} \mathcal{D}\tilde{\boldsymbol{S}} \exp\left[- \int dt d^d x \left(\tilde{\boldsymbol{S}} \cdot \left(\partial_t \boldsymbol{S} - \boldsymbol{F}(\boldsymbol{S}) - \boldsymbol{h} - 1/2 B \tilde{\boldsymbol{S}} \right) + \boldsymbol{J} \cdot \boldsymbol{S} \right) \right] \tag{A.9}$$

This expression can now be used to look at the behavior of the system using standard field theoretic techniques such as renormalization group and diagrammatic perturbation theory. Note that the external field $\boldsymbol{h}$ allows us to use the expression as a generating functional for moments of the response field $\tilde{\boldsymbol{S}}$ as well as for $\boldsymbol{S}$.

As a final example of how the formalism works, we look at the response of $\langle S_i \rangle$ to one component of the external field $\boldsymbol{h}$

$$\left.\frac{\delta \langle S_i \rangle}{\delta h_j}\right|_{h=0} = \left.\frac{\delta^2 Z[\boldsymbol{J}, \boldsymbol{h}]}{\delta h_j \delta J_i}\right|_{\substack{h=0 \\ j=0}} = \langle S_i \tilde{S}_j \rangle \tag{A.10}$$

This calculation indicates the origin of the term response fields for the fields $\tilde{\boldsymbol{S}}$ in that they are linearly coupled to the external field and their moments determine the structure of the physical fields $\boldsymbol{S}$ response to external fields.

Appendix B

Continuum limit for reaction-diffusion master equations

This appendix briefly reviews the continuum limit for the master equation for processes with diffusion. For concreteness, consider a two dimensional lattice of fixed size, with edge length L and M sites($M = (L/\delta)^2$). The probability in the time dt that a given organism will move to its nearest neighbor site is defined as τ, so the master equation is

$$\partial_t P(N_i) = -4\tau N_i P(N_i) + \tau \sum_\delta N_{i+\delta} P(N_{i+\delta}) \tag{B.1}$$

where δ are the nearest neighbor vectors of magnitude δ. We take the continuum limit ($\delta \to 0$) holding volume constant and recognize the discrete form of the laplacian in Eq. B.1 to obtain

$$\partial_t P(N, x, t) = \tau \delta^2 \nabla^2 P(N, x, t) \tag{B.2}$$

We have let the spatial index i go to the continuum index x. The diffusion constant is $D = \tau\delta^2$ which is naïvely 0 in the continuum limit. To avoid this, τ must grow as the dimensionless variable $(L/\delta)^2$.

We now obtain the mean field theory. Since the average number of organisms at a site is given by

$$\langle N_x \rangle = \sum_j P(j, x, t) j \tag{B.3}$$

we write

$$\sum_j j\partial_t P(j,x,t) = \sum_j j\tau\delta^2\nabla^2 P(j,x,t) \tag{B.4}$$

Since the derivatives don't depend on j, we move the sums through the derivatives to obtain the standard mean field theory of diffusive motion

$$\partial_t N_x = \tau\delta^2\nabla^2 N_x \tag{B.5}$$

Note that the mean field and full theory are the same equation here. This is because there are no two body interactions in this simple system and we have not considered higher moments.

Appendix C

Some explicit calculations for the Levin-Segel model

In this appendix we provide highly detailed summaries of a few key calculations from the analysis of the field theory for the Levin-Segel predator-prey model.

C.1 Predator competition term in second quantized Hamiltonian

For the predator competition term, the master equation is

$$\begin{aligned} \partial_t P(m) &= d(n+1)nP(n+1) - dn(n-1)P(n) \\ \rightarrow \partial_t|\psi\rangle &= d\sum\left[(n+1)nP(n+1) - n(n-1)P(n)\right]|n\rangle \end{aligned} \tag{C.1}$$

To extract the operator representation, let $n' = n+1 \rightarrow n = n'-1$

$$\begin{aligned} d\sum &(n-1)nP(n+1)|n-1\rangle \\ &= d\hat{a}a^2|\psi\rangle \end{aligned} \tag{C.2}$$

The next term can be computed as

$$\begin{aligned} d\sum &n(n-1)P(n)|n\rangle \\ &= d\hat{a}^2a^2|\psi\rangle \end{aligned} \tag{C.3}$$

combining gives

$$\partial_t|\psi\rangle = d\,(1-\hat{a})\,\hat{a}a^2 \tag{C.4}$$

C.2 System size expansions

Additional terms in the system size expansion

For the birth term

$$\begin{aligned}
&\frac{\beta}{V}z\rho(1-e^{-\hat{\rho}}) \\
&= b\left(\varphi+\frac{\eta}{\sqrt{V}}\right)(V\phi+\sqrt{V}\xi)\left(\frac{\hat{\rho}}{\sqrt{V}}-\frac{\hat{\rho}^2}{2V}\right) \\
&= b\left(\varphi+\frac{\eta}{\sqrt{V}}\right)\left(\sqrt{V}\phi\hat{\rho}-\frac{\phi\hat{\rho}^2}{2}+\hat{\rho}\xi\right) \\
&= b\left(\sqrt{V}\varphi\phi\hat{\rho}-\frac{\varphi\phi\hat{\rho}^2}{2}+\varphi\xi\hat{\rho}+\eta\phi\hat{\rho}\right)
\end{aligned} \tag{C.5}$$

For the predation term

$$\begin{aligned}
&\frac{p}{V}z\rho(1-e^{\hat{z}-\hat{\rho}}) \\
&= -p\left(\varphi+\frac{\eta}{\sqrt{V}}\right)\left(V\phi+\sqrt{V}\xi\right)\left(\frac{\hat{z}-\hat{\rho}}{\sqrt{V}}+\frac{(\hat{z}-\hat{\rho})^2}{2V}\right) \\
&= -p\left(\varphi+\frac{\eta}{\sqrt{V}}\right)\left(\sqrt{V}\phi(\hat{z}-\hat{\rho})+\frac{\phi(\hat{z}-\hat{\rho})^2}{2}+\xi(\hat{z}-\hat{\rho})\right) \\
&= -p\left(\sqrt{V}\phi\varphi(\hat{z}-\hat{\rho})+\frac{\phi\varphi(\hat{z}-\hat{\rho})^2}{2}+\varphi\xi(\hat{z}-\hat{\rho})+\eta\phi(\hat{z}-\hat{\rho})\right)
\end{aligned} \tag{C.6}$$

For predator death

$$\begin{aligned}
&\frac{d}{V}z^2(1-e^{-\hat{z}}) \\
&= d\left(\varphi+\frac{\eta}{\sqrt{V}}\right)(V\varphi+\eta)\left(\frac{\hat{z}}{\sqrt{V}}-\frac{\hat{z}^2}{2V}\right) \\
&= d\left(\varphi+\frac{\eta}{\sqrt{V}}\right)\left(\sqrt{V}\varphi\hat{z}-\frac{\hat{z}^2\varphi}{2}+\eta\hat{z}\right) \\
&= d\left(\sqrt{V}\varphi^2\hat{z}-\frac{\hat{z}^2\varphi^2}{2}+\eta\varphi\hat{z}\right)
\end{aligned} \tag{C.7}$$

C.3 Mean field analysis

To obtain the fixed points of the Levin-Segel model without space the equations are solved with the time derivative set to 0

$$\begin{aligned}
0 &= p\varphi\phi - d\varphi^2 \\
\rightarrow p\phi &= d\varphi
\end{aligned} \tag{C.8}$$

the next equation yields

$$\begin{aligned}
0 &= \alpha\phi + e\phi^2 - c\phi\varphi \\
\rightarrow 0 &= \alpha + e\phi - c\varphi \\
&= \alpha + \frac{de}{c}\varphi - c\varphi \\
\left(c-\frac{de}{c}\right)\varphi &= \alpha \\
\varphi &= \frac{c\alpha}{c^2-de}
\end{aligned} \tag{C.9}$$

Thus, the fixed point values at coexistence are

$$\begin{aligned} \varphi &= \frac{c\alpha}{c^2 - de} \\ \phi &= \frac{d\alpha}{c^2 - de} \end{aligned} \tag{C.10}$$

The fixed point values can be used to simplify the Jacobian matrix substantially. A_{11} can be simplified as

$$\begin{aligned} p\phi - 2d\varphi &= p\phi - pc\phi \\ &= -p\phi \end{aligned} \tag{C.11}$$

A_{22} can be simplified as

$$\begin{aligned} & \alpha + 2e\phi - p\varphi \\ = & \frac{\alpha p^2 - \alpha ed + 2ed\alpha - p^2\alpha}{p^2 - ed} \\ = & \frac{\alpha ed}{p^2 - ed} \\ = & e\phi \end{aligned} \tag{C.12}$$

In simplified form, the matrix $\boldsymbol{A}$ is given by

$$\boldsymbol{A} = \begin{pmatrix} -\nu k^2 - c\phi & c\varphi \\ -c\phi & -\mu k^2 + e\phi \end{pmatrix} \tag{C.13}$$

Now we evaluate the determinant of the ODE stability matrix (J above, and equal to $\boldsymbol{A}$

with space removed) and the trace

$$\begin{aligned} det(J) &= -ce\phi^2 + c^2\phi\varphi \\ &= c\phi(c\varphi - e\phi) = c\phi\frac{c^2\alpha - de\alpha}{c^2 - de} \\ &= c\phi\alpha \end{aligned} \tag{C.14}$$

The trace is simpler to compute

$$Tr(J) = (e - c)\phi \tag{C.15}$$

References

[1] Kalin Vetsigian, Carl Woese, and Nigel Goldenfeld. Collective evolution and the genetic code. *Proc Natl Acad Sci U S A*, 103(28):10696–10701, Jul 2006.

[2] A. J. McKane and T. J. Newman. Predator-Prey Cycles from Resonant Amplification of Demographic Stochasticity. *Phys. Rev. Lett.*, 94(21):218102, 2005.

[3] Carlos Lugo and A. J. McKane. Quasicycles in a spatial predator-prey model. *Phys. Rev. E*, 78, 2008.

[4] S J Freeland and L D Hurst. The genetic code is one in a million. *J Mol Evol*, 47(3):238–248, 1998.

[5] Shelley D Copley, Eric Smith, and Harold J Morowitz. A mechanism for the association of amino acids with their codons and the origin of the genetic code. *Proc. Natl. Acad. Sci. USA*, 102:4442–4447, 2005.

[6] Jose M. G. Vilar, Hao Yuan Keuh, Naama Barkai, and Stanislas Liebler. Mechanisms of noise-resistance in genetic oscillators. *Proc. Natl. Acad. Sci. USA*, 99(9):5988–5992, 2002.

[7] Thomas Butler, Nigel Goldenfeld, Damien Mathew, and Zaida Luthey-Schulten. Extreme genetic code optimality from a molecular dynamics calculation of amino acid polar requirement. *Phys. Rev. E*, 79(6):060901, Jun 2009.

[8] Paul C. Bressloff, Jack D. Cowan, Martin Golubitsky, Peter J. Thomas, and Matthew C. Wiener. What geometric visual hallucinations tell us about the visual cortex. *Neural Computation*, 14(3):473–491, 2002.

[9] Bruce Bower. Visions on the rocks. *Science News*, 150(14):216–217, 1996.

[10] J. C. Dunlap. Molecular bases for circadian clocks. *Cell*, 96:271–290, 1999.

[11] Robert M May. *Stability and complexity in model ecosystems*. Princeton University Press, 1973.

[12] William G Wilson, Susan P Harrison, Alan Hastings, and Kevin McCann. Exploring stable pattern formation in models of tussock moth populations. *J. Anim. Ecol.*, pages 94–107, 1999.

[13] Thomas Butler and Nigel Goldenfeld. Robust ecological pattern formation induced by demographic noise. *Phys. Rev. E*, 80(3):030902, Sep 2009.

[14] Nigel D Goldenfeld. *Lectures on phase transitions and the renormalisation group*. Westview press, 1992.

[15] Max Reitkerk and Johan van de Koppel. Regular pattern formation in real ecosystems. *TREE*, 23(3):169–175, 2008.

[16] A M Turing. The chemical basis of morphogenesis. *Phil. Trans. Roy. Soc. B*, 237:37–72, 1953.

[17] John Beggs and Dietmar Plenz. Neuronal avalanches in neocortical circuits. *Journal of Neuroscience*, 23, 2003.

[18] C R Woese and G E Fox. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proceedings of the National Academy of Sciences of the United States of America*, 74(11):5088–5090, 1977.

[19] C R Woese. The universal ancestor. *Proc. Natl. Acad. Sci. USA*, 95:6854–6859, June 1998.

[20] C R Woese. On the evolution of the genetic code. *Proc Natl Acad Sci U S A*, 54(6):1546–1552, Dec 1965.

[21] C R Woese. Evolution of the genetic code. *Naturwissenschaften*, 60(10):447–459, Oct 1973.

[22] Matthew T G Holden and et. al. Complete genomes of two clinical staphylococcus aureus strains: Evidence for the rapid evolution of virulence and drug resistance. *Proc. Natl. Acad. Sci. USA*, 101:9786–9791, 2004.

[23] H. Ochman, J. Lawrence, and E. Groisman. Lateral gene transfer and the nature of bacterial innovation. *Nature*, 405:299–304, 2000.

[24] Michael L Arnold. *Evolution Through Genetic Exchange*. Oxford University Press, 2006.

[25] E Cohen, D A Kessler, and H Levine. Recombination dramatically speeds up evolution of finite populations. *Phys. Rev. Lett.*, 94:098102, 2005.

[26] David Haig and Laurence D. Hurst. A quantitative measure of error minimization in the genetic code. *J. Mol. Evol.*, 33:412–417, 1991.

[27] CR Woese. Evolution of the genetic code. *Naturwissenschaften*, 60(10):447–459, 1973.

[28] S. Osawa. *Evolution of the genetic code*. Oxford Univ. Press, Oxford, 1995.

[29] R. Knight, S. Freeland, and L. Landweber. Rewiring the keyboard: evolvability of the genetic code. *Nat. Rev. Genet.*, 2:49–58, 2001.

[30] E Zuckerkandl and L Pauling. Evolutionary divergence and convergence in proteins. In V. Bryson and H. J. Vogel, editors, *Evolving Genes and Proteins*, pages 97–166, New York, 1965. Academic Press.

[31] T. Sonneborn. Degeneracy of the genetic code: Extent, nature and genetic implications. In V. Bryson and H. J. Vogel, editors, *Evolving Genes and Proteins*, pages 277–297, New York, 1965. Academic Press.

[32] C Alff-Steinberger. The genetic code and error transmission. *Proc. Natl. Acad. Sci. USA*, 64:584–591, 1969.

[33] C R Woese, D H Dugre, W C Saxinger, and S A Dugre. The molecular basis for the genetic code. *Proc Natl Acad Sci U S A*, 55(4):966–974, Apr 1966.

[34] C R Woese, D H Dugre, S A Dugre, M Kondo, and W C Saxinger. On the fundamental nature and evolution of the genetic code. *Cold Spring Harb Symp Quant Biol*, 31:723–736, 1966.

[35] R. D. Knight. *The origin and evolution of the Genetic Code: Statistical and experimental Investigations*. PhD thesis, Princeton University, June 2001.

[36] S Itzkovitz and U Alon. The genetic code is nearly optimal for allowing additional information within protein-coding sequences. *Genome Res.*, 17:405, 2007.

[37] Jun Sun and Michael W. Deem. Spontaneous emergence of modularity in a model of evolving individuals. *Physical Review Letters*, 99(22):228107, 2007.

[38] D. Ardell and G. Sella. No accident: genetic codes freeze in error-correcting patterns of the standard genetic code. *Phil. Trans. R. Soc. Lond. B*, 357:1625–1642, 2002.

[39] C R Woese. Order in the genetic code. *Proc Natl Acad Sci U S A*, 54(1):71–75, Jul 1965.

[40] R Grantham. Amino acid difference formula to help explain protein evolution. *Science*, 185:862–864, 1974.

[41] D C Mathew and Z Luthey-Schulten. On the physical basis of the amino acid polar requirement. *J Mol Evol*, 66:519–528, 2008.

[42] J C Phillips, R Braun, W Wang, J Gumbart, E Tajkhorshid, E Villa, C Chipot, R D Skeel, L Kale, and K Schulten. Scalable molecular dynamics with NAMD. *J Comp Chem*, 26:1781–1802, 2005.

[43] A. D. MacKerell, Jr. and N. Banavali. All-atom empirical force field for nucleic acids: 2) Application to molecular dynamics simulations of DNA and RNA in solution. *J Comp Chem*, 21:105–120, 2000.

[44] A. D. MacKerell, Jr., D. Bashford, M. Bellott, R.L. Dunbrack Jr., J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, III, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M Karplus. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B*, 102:3586–3616, 1998.

[45] W Humphrey, A Dalke, and K Schulten. VMD - Visual Molecular Dynamics. *J Molec Graphics*, 14:33–38, 1996.

[46] M.P. Allen and D.J. Tildesley. *Computer Simulation of Liquids*. Clarendon Press, Oxford, 1987.

[47] S J Freeland, R D Knight, L F Landwebber, and L D Hurst. Early fixation of an optimal genetic code. *Mol Biol Evol*, 17(4):511–518, 2000.

[48] S.J. Freeland, T. Wu, and N. Keulmann. The case for an error minimizing standard genetic code. *Origins of Life and Evolution of Biospheres*, 33(4):457–477, 2003.

[49] Thomas Butler and David Reynolds. Predator-prey quasicycles from a path integral formalism. *Phys. Rev. E*, 79:032901, 2009.

[50] R.D. Knight, S.J. Freeland, and L.F. Landweber. Rewiring the keyboard: evolvability of the genetic code. *Nature Reviews Genetics*, 2(1):49–58, 2001.

[51] RD Knight, L. Landweber, and M. Yarus. Tests of a stereochemical genetic code. *Translation mechanisms*, page 115, 2003.

[52] M. Yarus, J. Gregory Caporaso, and R. Knight. Origins of the genetic code: The escaped triplet theory. *Annual review of biochemistry*, 74:179–198, 2005.

[53] S. Itzkovitz and U. Alon. The genetic code is nearly optimal for allowing additional information within protein-coding sequences. *Genome research*, 17(4):405, 2007.

[54] R. Knight. *The origin and evolution of the genetic code: statistical and experimental investigations*. PhD thesis, Princeton University, 2001.

[55] Eugene V Koonin and Artem S Novozhilov. Origin and evolution of the genetic code: The universal enigma. *arXiv:0807.4749v2 [q-bio.GN]*, 2008.

[56] J Gregory Caporaso, Michael Yarus, and Rob Knight. Error minimization and coding triplet/binding site associations are independent features of the canonical genetic code. *J Mol Evol*, 61:597–607, 2005.

[57] S. Osawa and T. H. Jukes. Codon capture (codon capture) in evolution. *Journal of Molecular Evolution*, 28:271–278, 1989.

[58] D W Schultz and M Yarus. Transfer rna mutation and the malleability of the genetic code. *J. Mol. Bio.*, 235:1377–1380, 1994.

[59] R Knight, S J Freeland, and L F Landweber. Selection, history and chemistry: the three faces of the genetic code. *Trends Biochem. Sci.*, 24, 1999.

[60] D. H. Ardell and G. Sella. On the evolution of redundancy in genetic codes. *J. Mol. Evol.*, 53(4-5):269–281, 2001.

[61] E. N. Trifonov. Consensus temporal order of amino acids and evolution of the triplet code. *Gene*, 261(1):139 – 151, 2000.

[62] Jose M Montoya, Stuart L Pimm, and Ricard V Sole. Ecological networks and their fragility. *Nature*, 442:259–264, 2006.

[63] Jennifer A. Dunne, Richard J. Williams, and Neo D. Martinez. Food-web structure and network theory: The role of connectance and size. *Proceedings of the National Academy of Sciences of the United States of America*, 99(20):12917–12922, 2002.

[64] Nicolas E Humphries et al. Environmental context explains levy and brownian movement pattersn of marine predators. *Nature*, 465:1066–1069, 2010.

[65] Ottar N Bjornstadt and Bryan T Grenfell. Noisy clockwork: Time series analysis of population fluctuations in animals. *Science*, 293:638–643, 2001.

[66] A. J. McKane and T. J. Newman. Stochastic models in population biology and their deterministic analogues. *Phys. Rev. E*, 70:041902, 2004.

[67] Steven H. Strogatz. *Nonlinear Dynamics and Chaos*. Westview Press, 1994.

[68] Martin A. Nowak. *Evolutionary Dynamics*. Belknap/Harvard Press, 2006.

[69] Alan A Berryman. The origins and evolution of predator-prey theory. *Ecology*, 73(5), 1992.

[70] Andronov A and Pontrjagin L. Systemes grossiers. *Dokl. Akad. Nauk., SSSR*, 14, 1937.

[71] R. A. Fisher. The wave of advance of advantageous genes. *Annals of Eugenics*, 7(355-369), 1937.

[72] J. Maynard Smith. *Models in ecology*. Cambridge University Press, Cambridge, 1974.

[73] N. G. Van Kampen. *Stochastic Processes in Physics and Chemistry*. Elsevier, New York, 1992.

[74] John Cardy. *Scaling and Renormalization in Statistical Physics*. Cambridge University Press, 1996.

[75] A.J. McKane, J. D. Nagy, T. J. Newman, and M. O. Stefanini. Amplified biochemical oscillations in cellular systems. *J. Stat. Phys.*, 128(1/2):165–191, July 2007.

[76] M Mobilia, I. T. Georgiev, and U. C. Tauber. Fluctuations and correlations in lattice models for predator-prey interaction. *Phys. Rev. E*, 73:040903(R), 2006.

[77] M Mobilia, I. T. Georgiev, and U. C. Tauber. Phase transitions and spatio-temporal fluctuations in stochastic lattice lotka-volterra models. *J. Stat. Phys.*, 128(1/2):447–483, 2007.

[78] N. Boccara, O. Roblin, and M. Roger. Automata network predator-prey model with pursuit and evasion. *Phys. Rev. E*, 50(6):4531–4541, 1994.

[79] T Antal and M Droz. Phase transitions and oscillations in a lattice predator-prey model. *Phys. Rev. E*, 63:056119, 2001.

[80] Masao Doi. Second quantization representation for classical many particle system. *J. Phys. A.*, 9:1465–1479, 1976.

[81] A S Mikhailov. Path integrals in chemical kinetics i. *Phys. Lett.*, 85(4):214–216, 1981.

[82] L Peliti. Path integral approach to birth death processes on a lattice. *PJ. Physique*, 46:1469–1483, Sep 1985.

[83] N Goldenfeld. Kinetics of a model for nucleation-controlled polymer crystal growth. *J. Phys. A*, 17:2807–2821, 1984.

[84] Mark J Washenberger, Mauro Mobilia, and Uwe C Tauber. Influence of local carrying capacity restrictions on stochastic predator-prey models. *J. Phys. Condens. Matter*, 19:065139, 2007.

[85] D Mattis and M. L. Glasser. The uses of quantum field theory in diffusion limited reactions. *Rev. Mod. Phys.*, 70(3):979–1001, July 1998.

[86] John Cardy. Renormalization group approach to reaction-diffusion problems. *arXiv:cond-mat/9607163v2*, 1996.

[87] Hans Karl Janssen and Uwe C. Tauber. The field theory appraoch to percolation processes. *Annals of Physics*, 315(1):147–192, 2005.

[88] E E Holmes, M A Lewis, J E Banks, and R R Veit. Partial differential equations in ecology: Spatial interactions and population dynamics. *Ecology*, 75:17–29, 1994.

[89] P. C. Martin, E. D. Siggia, and H. A. Rose. Statistical dynamics of classical systems. *Phys. Rev. A*, 8:423–437, 1973.

[90] R. Bausch, H. K. Janssen, and H. Wagner. Renormalized field theory of critical dynamics. *Z. Phys. B.*, 24:113–127, 1976.

[91] M. C. Cross and P. C. Hohenberg. Pattern formation outside of equilibrium. *Rev. Mod. Phys.*, 65(3):851, Jul 1993.

[92] Simon Levin. The problem of pattern and scale in ecology. *Ecology*, 73:1943–1967, 1992.

[93] A. J. Koch and H. Meinhardt. Biological pattern formation: from basic mechanisms to complex structures. *Rev. Mod. Phys.*, 66(4):1481–1507, Oct 1994.

[94] J. D. Murray. A pre-pattern formation mechanism for animal coat markings. *Journal of Theoretical Biology*, 88(1):161 – 199, 1981.

[95] T. Lu, D. Karig, and R. Weiss. Emergent pattern formation in a synthetic bacterial system. 2010. submitted.

[96] V. Castets, E. Dulos, J. Boissonade, and P. De Kepper. Experimental evidence of a sustained standing turing-type nonequilibrium chemical pattern. *Phys. Rev. Lett.*, 64(24):2953–2956, Jun 1990.

[97] Johan van de Koppel, Joanna C. Gascoigne, Guy Theraulaz, Max Rietkerk, Wolf M. Mooij, and Peter M. J. Herman. Experimental Evidence for Spatial Self-Organization and Its Emergent Effects in Mussel Bed Ecosystems. *Science*, 322(5902):739–742, 2008.

[98] Lord Rayleigh. On convection currents in a horizontal layer of fluid when the higher temperature is on the underside. *Philosophical Magazine Series 6*, 32(192):529–546, 1916.

[99] Simon A Levin and Lee A Segel. Hypothesis for the origin of planktonic patchiness. *Nature*, 259:659, 1976.

[100] Horst Malchow, Frank M Hilker, Ivo Siekmann, Sergei Petrovski, and Alexander B Medvinsky. Mathematical models of pattern formation in planktonic predation-diffusion systems: A review. *Aspects of Mathematical Modelling*, pages 1–26, 1998.

[101] Cabell S Davis, Scott M Gallager, and Andrew R Solow. Microaggregation of oceanic plankton observed by towed video microscopy. *Science*, 257:230–232, 1992.

[102] Edward R Abraham. The generation of planktonic patchiness by turbulent stirring. *Nature*, 391:577–580, 1998.

[103] John L Maron and Susan Harrison. Spatial pattern formation in an insect host-parasitoid system. *Science*, 278:1619–1621, 1997.

[104] Thomas C Butler, Jack D. Cowan, Marc Benayoun, Edward Wallace, and Wim van-Drongelen. Fluctuation driven patterns on the visual cortex. *In preparation*, 2010.

[105] M Mimura and J D Murray. On a diffusive prey-predator model which exhibits patchiness. *J. Theor. Biol.*, 75:249–262, 1978.

[106] Martin Baurmann, Thilo Gross, and Ulrike Feudel. Instabilities in spatially extended predator-prey systems: Spatio-temporal patterns in the neighborhood of turing-hopf bifurcations. *J. Theor. Biol.*, 245:220–229, 2007.

[107] M Katori, S Kizaki, Y Terui, and T Kubo. Forest dynamics with canopy gap expansion and the stochastic ising model. *Fractals*, 6:81–86, 1998.

[108] S Davis, P Trapman, H Leirs, M Begon, and J A P Heesterbeek. The abundance threshold for plague as a critical percolation phenomenon. *Nature*, 454:634–637, 2008.

[109] Mingming Wu, Guenter Ahlers, and David S. Cannell. Thermally induced fluctuations below the onset of rayleigh-bénard convection. *Phys. Rev. Lett.*, 75(9):1743–1746, Aug 1995.

[110] Mario Pineda-Krch, Hendrik J Blok, and Michael Doebeli. A tale of two cycles - distinguishing quasi-cycles and limit cycles in finite predator-prey populations. *Oikos*, 116:53–64, 2007.

[111] Alex J. Bladon, Tobias Galla, and Alan J. McKane. Evolutionary dynamics, intrinsic noise, and cycles of cooperation. *Phys. Rev. E*, 81(6):066122, Jun 2010.

[112] Tommaso Biancalani, Duccio Fanelli, and Francesca Di Patti. Stochastic turing patterns in the brusselator model. *Phys. Rev. E*, 81(4):046215, Apr 2010.

[113] Frank Courchamp, Tim Clutton-Brock, and Brian Grenfell. Inverse density dependence and the allee effect. *TREE*, 14(10):405–410, 1999.

[114] William G. Wilson. Resolving discrepancies between deterministic population models and individual-based simulations. *The American Naturalist*, 151(2):116–134, 1998.

[115] J. García-Ojalvo, A. Hernández-Machado, and J. M. Sancho. Effects of external noise on the swift-hohenberg equation. *Phys. Rev. Lett.*, 71(10):1542–1545, Sep 1993.

[116] O Carrillo, Santos M A, Garcia-Ojalvo J, and J M Sancho. Spatial coherence resonance near pattern-forming instabilities. *Europhys. Lett.*, 65:452, 2004.

[117] Michael Sieber, Horst Malchow, and Lutz Schimansky-Geier. Constructive effects of environmental noise in an excitable prey-predator plankton system with infected prey. *Ecological complexity*, 4:223–233, 2007.

[118] Brian Dennis and Robert F Costantino. Analysis of steady state populations with the gamma abundance model: Application to tribolium. *Ecology*, 69:1200–1213, 1988.

[119] B. E. Saether, J. Tufto, S. Engen, K. Jerstad, O. W. Rostad, and J. E. Skatan. Population dynamical consequences of climate change for a small temperate songbird. *Science*, 287:854–856, 2000.

[120] Daniel T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22(4):403 – 434, 1976.

[121] Andre W. Visser and Thomas Kiorbee. Plankton motility patterns and encounter rates. *Oecologia*, 148:538–546, 2006.

[122] W. K. Li. Macroecological patterns of phytoplankton in the northwestern north atlantic ocean. *Nature*, 419:154–157, 2002.

[123] Joseph Kane. Zooplankton trends on george's bank, 1977-2004. *ICES journal of marine science*, 64:909–919, 2007.

[124] B T Grenfell, O N Bjornstadt, and J Kappey. Travelling waves and spatial hierarchies in measles epidemics. *Nature*, 414:716–724, 2001.

[125] Derek A T Cummings, Rafael A Irizarry, Norden E Huang, Timothy P Endy, Ananda Nisalak, Kumnuan Ungchusak, and Donald S Burke. Travelling waves in the occurence of dengue haemorrhagic fever in thailand. *Nature*, 427:344–347, 2004.

[126] G. B. Ermentrout and J. D. Cowan. A mathematical theory of visual hallucination patterns. *Biol. Cybernetics*, 34:137–150, 1980.

[127] H. Kluver. *Mescal and Mechanisms of Hallucinations.* University of Chicago Press, 1966.

[128] P. C. Bressloff, J. D. M. Golubitsky, P. J. Thomas, and M. C. Wiener. Geometric visual hallucinations, euclidean symmetry and the functional architecture of striate cortex. *Phil. Trans. R. Soc. Lond.*, 356:299–330, 2001.

[129] P. C. Bressloff and J. D. Cowan. An amplitude equation approach to contextual effects in visual cortex. *Neural Computation*, 14:493–525, 2002.

[130] C. D. Gilbert and T. N. Wiesel. Clustered intrinsic connections in cat visual cortex. *J. Neurosci.*, 3:1116–1133, 1983.

[131] J. D. Cowan. Stochastic neurodynamics. *Advances in Neural Information Processing Systems*, 3:62–69, 1990.

[132] M. Buice and J. D. Cowan. Field-theoretic approach to fluctuation effects in neural networks. *Physical Review E*, 75, 2007.

[133] W. van Drongelen M. Benayoun, J. D. Cowan and E. Wallace. Avalanches in a stochastic model of spiking neurons. *PLOS Computational Biology*, 2010. (In Press).

[134] J. D. Cowan. Statistical mechanics of nervous nets. *In E.R. Caianiello, Editor, Neural Networks*, pages 181–188, 1968.

[135] J. D. Cowan. A statistical mechanics of nervous activity. *In M. Gerstenhaber, Editor, Some Mathematical Questions in Biology*, 1971.

[136] D. H. Hubel and T. N. Wiesel. Uniformity of monkey striate cortex: A parallel relationship between field size, scatter and magnification factor. *J. Comp. Neurol.*, 158:295–306, 1974.

[137] C. F. Stevens. How cortical interconnectedness varies with network size. *Neural Computation*, 1:473–479, 1989.

[138] R. J. Douglas and K. A. C. Martin. A functional microcircuit for cat visual cortex. *J. Physiol.*, 440:735–769, 1991.

[139] H. Markham, M. Toledo-Rodriguez, Y. Wang, A. Gupta, G. Silverberg, and C. Wu. Interneurons of the neocortical inhibitory system. *Nature Reviews, Neuroscience*, 5(10):793–807, 2004.

[140] E. Bullmore and O. Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews: Neuroscience*, 10:186–198, 2009.

[141] D. J. Watts and S. H. Strogatz. Collective dynamics of "small-world" networks. *Nature*, 393:440–442, 1998.

[142] H. R. Wilson and J. D. Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical J.*, 12:1–22, 1972.

[143] H. R. Wilson and J. D. Cowan. A mathematical thoery of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik*, 13:55–80, 1973.

[144] G. B. Ermentrout and J. D. Cowan. Large scale spatially organized activity in neural nets. *SIAM J. Appl. Math.*, 38:1–21, 1980.

[145] Uri Alon. *Introduction to Systems Biology.* Chapman Hall/CRC, 2007.

[146] Johan Paulsson. Summing up the noise in gene networks. *Nature*, 427:415–418, 2004.

[147] Christopher V. Rao, Denise M. Wolf, and Adam P. Arkin. Control, exploitation and tolerance of intracellular noise. *Nature*, 420:231–237, 2002.

[148] Yan-Jiun Daniels, Bryan C. Chen and James P. Sethna. Sloppiness, robustness and evolvability in systems biology. *Current Opinion in Biotechnology*, 19:389–395, 2008.

[149] Hiroaki Kitano. Toward a theory of biological robustness. *Molecular systems biology*, 3, 2007.

[150] Hiroaki Kitano. Biological robustness. *Nature reviews: Genetics*, 5, 2004.

[151] Ryan N. Gutenkunst, Joshua J. Waterfall, Fergal P. Casey, Kevin S. Brown, Christopher R. Myers, and James P. Sethna. Universal sloppy parameter sensetivities in systems biology models. *PLoS Comp. Bio.*, 3:1871–1878, 2007.

[152] Joshua J. Waterfall, Fergal P. Casey, Ryan N. Gutenkunst, Kevin S. Brown, Christopher R. Myers, Piet W. Brouwer, Viet Elser, and James P. Sethna. Sloppy model universality class and the vandermonde matrix. *Phys. Rev. Lett.*, 97:150601, 2006.

[153] Kevin S. Brown and James P. Sethna. Statistical mechanical approaches to models with many poorly known paramaters. *Phys. Rev. E*, 68:021904, 2003.

[154] Bela Novak and John J. Tyson. Design principles of biological oscillators. *Nature reviews: Molecular cell biology*, 9:981–991, 2008.

[155] Lars Onsager. Crystal statistics. i. a two-dimensional model with an order-disorder transition. *Phys. Rev.*, 65(3-4):117–149, Feb 1944.

[156] J. Bardeen, L. N. Cooper, and J. R. Schrieffer. Theory of superconductivity. *Phys. Rev.*, 108(5):1175–1204, Dec 1957.

Vita

Thomas Butler was born in Seattle Washington in 1980. He attended Brigham Young University starting in 1998, and took a two year break to work with his Church as a volunteer in economically and socially depressed neighborhoods of Chicago from 2000-2002. He graduated in late 2004 with a B.S. in physics and minor in mathematics, and after a year of graduate school at UCLA, he transferred to the University of Illinois. At Illinois, he has been recognized as an outstanding teacher several times, and has also had his research recognized by a Drickhamer fellowship.