

**BROKEN SYMMETRY AND CRITICAL PHENOMENA IN
POPULATION GENETICS: THE STEPPING-STONE MODEL.**

By

Timothy Lee Duty

B.S., Virginia Polytechnic Institute and State University, 1990

M.Sc., University of British Columbia, 1993

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES
DEPARTMENT OF PHYSICS AND ASTRONOMY

We accept this thesis as conforming
to the required standard

.....
.....
.....
.....

THE UNIVERSITY OF BRITISH COLUMBIA

September 2000

© Timothy Lee Duty, 2000

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Physics and Astronomy
The University of British Columbia
6224 Agricultural Road
Vancouver, B.C., Canada
V6T 1Z1

Date:

Abstract

In this thesis, I study the behaviour of the Stepping-Stone model: a stochastic model from theoretical population genetics. It was introduced by Kimura and Weiss [1][2] as a simple model to investigate the interplay of the evolutionary processes of random genetic drift, mutation, migration and selection. In particular, they were interested in the behaviour of spatially-structured populations when these processes were mediated by local interactions. From the point of view of statistical physics, the Stepping-Stone model can be viewed as a 2-species, non-equilibrium reaction-diffusion model with spatial degrees of freedom and unique fluctuations that arise from irreversible processes at the microscale.

My thesis begins with a brief overview of the theory of stochastic processes that includes both the classical treatment using master equations, the Fokker-Planck equation and the Langevin equation, and modern formalisms that map these equations to operators and functional integrals. This is followed by a discussion of the steady-state and critical phenomena in the Wright-Fisher and Moran models, the non-spatial predecessors of the Stepping-Stone model. Critical phenomena in these models is shown to be associated with the breaking of a discrete symmetry that results in a discontinuous order parameter. Next, the Stepping-Stone model is introduced and reformulated using operators and path-integrals.

Kimura and Weiss were able to solve for the steady-state of the Stepping-Stone model, under certain conditions, but the dynamics of the model remained elusive. This model is related however, to the “Voter” model, which has been well-studied and is known to have a critical spatial dimension of 2, in that only for $d \leq 2$ does the system asymptotically reach one of the 2 degenerate absorbing states.

I extend the results of Kimura and Weiss and obtain exact results for the dynamics

of the Stepping-Stone model. Two regimes of the model are analyzed: 1) the neutral regime, where selection has vanished and mutation is viewed as the control parameter. Here a steady state exists, characterized by a correlation length that diverges as the mutation rate, μ becomes zero; and 2) the broken symmetry or “fixed” state, where selection is small but finite, and mutations so rare that the relevant description concerns the dynamics of “avalanches” or “cascades” of new alleles induced by the initial mutation and perturbing the system from the absorbing state.

A unique kind of dynamical critical phenomenon occurs when selective advantage and mutation become negligible. Like the Voter model, it is qualitatively different both above and below a critical dimension $d_c = 2$. For spatial dimension $d \leq 2$, the critical behaviour is associated with the breaking of a discrete symmetry and corresponds to fixation of one of the two alleles (genotypes). Symmetry breaking in the Stepping-Stone model is shown to be a consequence of the asymptotic return probability of a random walk—identically one for $d \leq 2$ and strictly less than one for $d > 2$. In addition to the correlation length, the steady-state of the neutral regime is further characterized by a measure of variance defined as the amplitude of the two-point correlation function evaluated at vanishing separation. In genetics this measure of variance is known as F_{ST} , the fixation index. Exact results are derived for both the steady-state value of F_{ST} , and its asymptotic time-dependence at the critical point, which approaches 1 in both $d = 1$ and $d = 2$. At the critical point, $1 - F_{ST} \sim t^{-1/2}$ for $d = 1$. For $d = 2$, a much slower decay is found, $1 - F_{ST} \sim 1/\ln(t)$. The $d = 3$ critical behaviour of F_{ST} is that it approaches a constant $C < 1$ as $(C - F_{ST}) \sim t^{-1/2}$. The constant C is non-universal and related to the return probability of a random walk. It approaches 1 for very large values of the dimensionless constant $\kappa = \frac{\Lambda}{2\pi^2 D n \tau_g}$, where Λ^{-1} is the spatial scale of the interaction, τ_g is the generation time, D is the diffusion constant and n is the population density.

A related infinite-alleles model was studied by both Sawyer[3] and Nagalaki[4]. These authors found similar asymptotics for the probability that two randomly chosen individuals are genetically identical under certain assumptions. Sawyer[3] also proved that fixation in the infinite alleles model occurs iff the random walk followed by the individuals is recurrent.

The broken symmetry or “fixed” regime of the Stepping-Stone model has been explored from the point of view of survival of rare mutant alleles, here parameterized by a coefficient of selection s . The exponent ν , governing the divergence of the relaxation time as $s \rightarrow 0$ is calculated and found to be $\nu = 2$ for $d = 1$, while for $d \geq 2$ it is given by the mean field value of $\nu = 1$. Two other exponents for critical spreading processes are determined and scaling arguments are presented and used to find the decay exponents that characterize the time-dependent survival probability and its asymptotic value for very long times. These results also establish the upper critical dimension of the model, $d_c = 2$.

Finally, these results are rederived and supplemented by a dynamical renormalization group (RG) analysis. The critical behaviour in $d = 1$ is found in both regimes to be controlled by non-trivial fixed points. The critical behaviour of the broken symmetry regime for $d \geq 2$ and its RG fixed point is that of a critical branching process. For the neutral regime, however, the renormalization group flow is qualitatively different in $d = 2$ and 3, reflecting the existence of broken symmetry for $d = 2$. The RG flow for the $d = 3$ neutral regime contains a line of fixed points with the effective description at large scales given by a Gaussian version of the time-dependent Landau-Ginsburg model. Finally, the renormalization results are found to be valid to all orders of perturbation theory.

Table of Contents

Table of Contents	v
List of Figures	x
1 Introduction	1
2 Stochastic Processes.	6
2.1 The Master Equation.	6
2.2 The Fokker-Planck Equation.	10
2.2.1 The forward equation.	10
2.2.2 Solution of the Fokker-Planck equation.	12
2.2.3 The backward equation.	13
2.2.4 System size expansions: from master equation to Fokker-Planck equation.	15
2.3 Langevin Equations.	16
2.3.1 A single variable Langevin equation.	17
2.3.2 Continuum Langevin equations.	18
2.4 Operators and Path Integrals.	20
2.4.1 Response functional formalism.	21
2.4.2 The master equation in occupation number representation.	25
2.4.3 The Fokker-Planck equation using operators.	38
2.4.4 System size expansions, again!	44

3	Spontaneous Symmetry Breaking in the Wright-Fisher/Moran Model	49
3.1	The Wright-Fisher and Moran Models.	50
3.2	The Fokker-Planck Equation for Moran's Model.	51
3.2.1	Expansion of the master equation.	53
3.2.2	Operator forms of Moran's model.	55
3.3	Heterogeneous diffusion in the Wright-Fisher and Moran models.	57
3.4	Mean-field theory.	60
3.5	The Effective Potential.	61
3.6	The Broken Symmetry State: Fate of new mutants.	65
3.7	Discussion.	72
4	The Stepping-Stone Model Revisited	74
4.1	Introduction.	74
4.1.1	The Exchange Coupled Stepping-Stone Model (ECSS).	77
4.1.2	The Diffusively Coupled Stepping-Stone Model (DCSS).	77
4.2	The Neutral Regime.	78
4.2.1	Fourier transformation to wave-vector space.	79
4.2.2	Equation of motion for the two-point correlation function.	81
4.2.3	The steady-state.	83
4.3	Broken Symmetry or $d \leq 2$	86
4.4	Critical dynamics of the two-point correlation function.	86
4.4.1	The decay of heterozygosity.	86
4.4.2	The infinite alleles model.	91
4.5	Alternative Formulations of the Stepping-Stone Model.	93
4.5.1	The Fokker-Planck equation and operators.	95
4.5.2	The path-integral.	95

4.5.3	The response functional method.	96
4.5.4	The neutral regime.	97
4.6	The Broken-Symmetry State: Fate of a single mutant gene.	98
4.6.1	Diagrammatic Methods in Field Theory.	102
4.6.2	Perturbation theory for the relaxation rate Σ	103
4.6.3	The effective relaxation rate and “screening” of selection in low dimensions.	107
4.6.4	Survival probabilities: scaling arguments.	110
5	Renormalization Group Analysis	115
5.1	Introduction.	115
5.2	A Simple Example of Renormalization: The biased random walk.	116
5.3	Perturbations from the Random Walk Fixed Point.	119
5.3.1	n -particle to m -particle processes.	119
5.3.2	Relevance of independent decay and branching processes.	121
5.3.3	Behavior of coagulation, or annihilation processes.	121
5.3.4	A paradox.	121
5.4	The Critical Branching Fixed Point.	122
5.4.1	Perturbations from the Critical Branching fixed point.	124
5.5	The Gaussian Fixed Point.	125
5.5.1	Perturbations from the Gaussian Fixed Point.	127
5.6	Summary of non-interacting fixed points.	128
5.7	Renormalization of the Stepping-Stone Model.	130
5.7.1	The Neutral Regime.	130
5.7.2	The Broken-Symmetry State.	135
5.8	Determining critical exponents from RG eigenvalues.	138

6	Summary of Results and Conclusions	141
6.1	The Wright-Fisher and Moran Models.	141
6.2	The Stepping-Stone Model.	142
6.2.1	The neutral regime.	142
6.2.2	The broken symmetry state.	144
6.3	Renormalization Group For Reaction-Diffusion and Related Stochastic Processes.	145
6.3.1	Renormalization to all orders.	146
6.3.2	Modern Theory of Stochastic Processes.	146
	Bibliography	148
A	Random Walks on the Linear, Square and Cubic Lattices.	152
A.1	Random Walk Green's Functions.	152
A.2	Relation of the first-passage time to the Green's function.	155
A.3	The Critical Dimension of the Random Walk.	157
A.4	Asymptotic Behaviour of the Random Walk Green's Function $G_\alpha(\mathbf{0},t)$	157
A.5	The Laplace Transform of $G_\alpha(\mathbf{0},t)$	158
A.5.1	The Linear Lattice.	159
A.5.2	The Square Lattice.	159
A.5.3	The Cubic Lattice.	160
B	Computer Simulations	162
B.1	Stochastic Differential Equations.	162
B.2	Continuous-time Reaction Diffusion Processes.	162
C	The Branching Process	165
C.1	The model.	165

C.2	Survival probability.	166
C.2.1	The method of characteristics.	167
C.2.2	The backward equation.	169
C.3	Size distribution.	171

List of Figures

3.1	Schematic diagram of the Wright-Fisher model.	52
3.2	Stochastic Processes in the Moran Model.	58
3.3	The mean-field potential $V(x)$	62
3.4	Comparison of the stochastic “effective potential” $U(x)$ with the mean-field potential $V(x)$	66
3.5	Simulations of the Wright-Fisher model using stochastic differential equations.	67
3.6	Discontinuous behaviour of the order parameter at the critical point of the Wright-Fisher model.	68
4.1	Computer simulations of the $d = 2$ Stepping Stone model steady state.	87
4.2	Numerical solution for the heterozygosity $h(t) = 1 - F_{ST}(t)$	92
4.3	Vertices representing the action for propagation of new mutant alleles.	101
4.4	Feynman diagram representing the 1-loop correction to the two-point propagator	104
4.5	Perturbation expansion for the two-point propagator.	105
4.6	Perturbation expansion and Dyson equation for Σ	106
4.7	Divergence of the relaxation time for $d = 1, 2$ and 3	109
4.1	Simulation results for the time-dependent survival probability for $d = 1$	114
5.2	Vertices representing the Lagrangian of the Neutral regime	131
5.3	Bubble diagrams contributing to the renormalization of v and u	132

5.4	Diagram representing the 1-loop integral that contributes to the differential RG recursion relation.	133
5.5	Renormalization group flow for the neutral theory.	136
5.6	RG flow for the broken symmetry state.	139
C.1	A typical tree from a branching process	171

Chapter 1

Introduction

“Evolution is a stochastic process of change in gene frequencies in natural populations.”

—Motoo Kimura [5]

Population geneticists have long recognized the importance of random or stochastic processes in producing and maintaining genetic variation but have always disagreed on their significance to the mechanisms of evolution. The work of both Sewall Wright [6] and R. A. Fisher [7] marked the beginning of a quantitative theory of the role of genetic drift in the evolutionary process that was outlined by Darwin much earlier in his *Origin of Species*. What has become known as the Wright-Fisher model now occupies a central role of this theory and forms a starting point for the application of molecular population genetics to natural populations. The Wright-Fisher model has been further generalized and aspects of it find their way into many areas of the biological sciences. Surprisingly, Wright argued that random processes were essential to every aspect of evolution while Fisher took the opposite view that evolution worked by mass selection alone, with random processes playing the minor role of generating rare new variations. Fisher was responsible for “The Fundamental Theorem of Natural Selection”, a theory of evolution based upon deterministic equations that has influenced generations of evolutionary biologists.

The controversy surrounding the importance of random genetic drift reached a culmination with the “Neutral Theory of Molecular Evolution” outlined by Kimura [8][9]. He proposed that most if not all the observed genetic variation at the molecular level was

due to recent mutations of selectively neutral alleles balanced by fixation due to random genetic drift. In this thesis I show that from the standpoint of modern statistical physics, his Neutral Theory in the limit of rare mutations can be considered a theory of critical phenomena in population genetics. The significance of this finding is not merely academic as observed mutation rates *are* in fact small ($\sim 10^{-9} - 10^{-10}$ per nucleotide per generation), and selective differences, at least at the molecular level are tiny or non-existent [8].

Kimura also investigated the influence of spatial structure on genetic drift by formulating the so-called Stepping-Stone model, following Wright's work on what became known as "isolation by distance" or equivalently, correlation of nearby individuals. In the Wright-Fisher model, a predecessor to the Stepping-Stone model, each individual of the population is equally likely to interact with any other. This gives rise to questions concerning the implications of purely local interactions on the evolutionary process (in genetics this is often referred to as spatial population structure). Wright, argued for the necessity of such spatial subdivision of populations in order to facilitate the "crossing of adaptive valleys" in the fitness landscape. Physicists tend to think of this as "hopping over a potential barrier". His theory, which he christened the "Shifting Balance Theory", was first proposed in 1931 and has met with a stern array of opponents including (of course) Fisher, who down-played the role of small population size and genetic drift in favour of his mass selection. Debate over the relevance of the shifting balance theory for the evolutionary processes continues to rage and despite some recent criticisms of Wright's theory, the question is far from being resolved [10] [11][12]. Interestingly enough, Kimura's neutral theory seems to sidestep Wright's Shifting Balance Theory by providing a very different picture of the evolutionary process at the molecular level. However, in both the theories of Kimura and Wright, random processes are fundamental and therefore unfavoured by the opposing camp of determinists.

Some of the questions concerning spatial structure were answered by Kimura and Weiss [1][2] with their analysis of the Stepping-Stone model. Others, including those related to the shifting balance theory were left unresolved. The work carried out in this thesis describes an attempt to extend their results using some of the ideas and techniques from modern statistical physics and as such, does not directly address the debate over the shifting balance theory¹. On the other hand, several new results have been obtained in this study that unambiguously highlight the importance of local interactions and spatial dimensionality on large scale gene flow in distributed populations. These new results are even more relevant for the neutral theory. Hopefully, the extension of the theory to more complicated systems will succeed in addressing ideas such as Wright's shifting balance theory and further elaborating Kimura's neutral theory. With regard to understanding complex many-body populations, it is hoped that the techniques for treating stochastic processes described in this thesis take us a few steps further along our own fitness landscape that represents our understanding of many-body problems in stochastic population models.

Kimura realized the need for new mathematical techniques in 1957,

“In his theories of evolution Wright put forward an important concept of ‘balance,’ especially of balance between directional factors such as selection, mutation, and migration and non-directional or stochastic factors such as random sampling of gametes and random fluctuation of environmental conditions. *It appears that new methods of stochastic processes will be needed for a satisfactory treatment of Wright's theory of evolution.*—Motoo Kimura [13]

The research contained in this thesis employs several alternative, yet equivalent theoretical frameworks for treating stochastic processes. These are briefly reviewed in

¹That would require a multi-locus, perhaps many allele model in contrast to the single-locus, two-allele model studied in the following chapters.

chapter 2. Each theoretical approach has its particular advantages and disadvantages and therefore contributes in complementary ways to the overall understanding that has emerged from their synthesis. In this thesis, we will employ all of these to varying degrees in order to complete our understanding of some very fundamental models of population genetics. Perhaps the most intuitive, easiest to understand and oldest of these frameworks is the Fokker-Planck equation which describes the explicit time evolution of the probability density. Unfortunately, its generalization to many degrees of freedom has limited usefulness. The operator formalism for either discrete or continuous variables is much less intuitive and quite unwieldy for describing probability densities. On the other hand, its utility is found in deriving and solving equations for moments and correlation functions. Finally, a functional integral can be obtained from either the response functional method or the operator approach and is analogous to the partition function in equilibrium statistical mechanics or the path-integral in quantum field theory. The path-integral description of stochastic processes is invaluable for setting up a consistent perturbation theory and renormalization group analysis.

In chapter 3 I review the Wright-Fisher and Moran models as a warm-up for understanding the behaviour of the Stepping-Stone model. The next two chapters contain the main results of this thesis. In chapter 4 exact results for the Stepping-Stone model are derived and discussed. In chapter 5, I re-examine the Stepping-Stone model from the point of view of the Renormalization Group (RG). The results obtained from the RG are shown to correspond qualitatively and quantitatively to the exact results of chapter 4.

We are fortunate that the Stepping-Stone model admits a few exact results and is amenable to the different approaches mentioned above. This situation creates a much firmer foundation for the application of these methods to the many, more difficult problems in stochastic population dynamics, reaction-diffusion theory and other non-equilibrium models. Indeed, the Stepping Stone model may be somewhat unique in

that complementary to the exact results, a non-trivial renormalization group analysis to all orders is possible. The benefits afforded by this observation form a two-way street. Not only can the RG help us understand the Stepping Stone model, the Stepping-Stone model provides a deeper understanding of the application of renormalization group methods to non-equilibrium processes in general, and sheds light on the conceptual foundation of the RG, itself.

Chapter 2

Stochastic Processes.

This chapter is intended as an overview of stochastic processes and serves to introduce the reader to the terminology and notation used throughout this thesis. First I discuss the classic approaches to working with stochastic processes: the master equation, the Fokker-Planck and Langevin equations. I also point out how these are related and highlight the various equivalences between formalisms. Next we describe several more modern methods of stochastic processes that map the classic approaches to operators, functional integrals or both. These latter techniques, while more abstract and somewhat harder to grasp, are much more valuable for the many-body problems encountered in the analysis of lattice and continuum models. Readers with experience in the description of stochastic processes may only need to skim this chapter to familiarize themselves with the notation. Newcomers to stochastic processes should first read the entire chapter and may need to consult some of the references, even though the presentation is meant to be self-contained. Those whose are familiar with the more traditional methods, the master equation for discrete variables and Fokker-Planck and Langevin equations for continuous variables, should review section 2.4 on the mapping of these equations to operators descriptions and path-integrals.

2.1 The Master Equation.

The master equation offers the most flexible description for stochastic processes. Consider a stochastic model where the state or configuration space of the system is the set

$\{C\}$. The generality of the master equation is that C can represent almost any kind of variable: a configuration of spins on a lattice, a m -tuple (n_a, n_b, \dots, n_m) , for example describing the numbers of a m -species system, a set of vertices and edges, a set of binary numbers and so on. The continuous time master equation is a linear equation for the time derivative of the probability $P_C(t)$ —the probability that the system is in configuration C at time t . It is written in terms of the transition rates $w(C'/C)$ that give the probability per unit time of jumping from C to C' ,

$$\frac{\partial}{\partial t} P_C(t) = \sum_{C'} [w(C/C') P_{C'}(t) - w(C'/C) P_C(t)], \quad (2.1)$$

and simply expresses the conservation of probability: The first term represents the net flow of probability into configuration C and the second, the net flow out of C . We will be concerned with stochastic processes that have the Markov property. This means that the transition rates $w(C'/C)$ depend only on the current configuration as opposed to earlier ones and furthermore are not explicitly time-dependent.

As a simple example, we consider the master equation for a nearest-neighbor random walk on the integer line or linear lattice. Here the state space is \mathbb{Z} (the integers) and we label a particular configuration as n , the position of our walker. The master equation reads

$$\frac{\partial}{\partial t} P_n(t) = \alpha \{P_{n-1}(t) + P_{n+1}(t) - 2P_n(t)\}, \quad (2.2)$$

where α is the jump or migration rate. The special solution of this equation for the initial condition $P_n(0) = \delta_{n,m}$ is denoted the random walk *Green's function* or *propagator* $G_\alpha(n, m, t)$. Note that this quantity is really a *conditional probability* which is commonly denoted $P(n, t|m, 0)$ in probability textbooks. I, however, will use the physics convention here of calling it G . The translational invariance of Eq. 2.2 tells us that the dependence of G on m and n can only be through their difference, that is $G_\alpha(n, m, t) = G_\alpha(n - m, t)$. If Q_m is the probability that the walker is initially at m , then the general solution of Eq.

(2.2) is given by

$$P_n(t) = \sum_m G_\alpha(n-m, t) Q_m.$$

Another example of a master equation would be that describing radioactive decay. Here the configuration space is the positive integers \mathbb{N} , and the master equation given by

$$\frac{\partial}{\partial t} P_n(t) = \lambda \{ (n+1) P_{n+1}(t) - n P_n(t) \}, \quad (2.3)$$

having a single parameter λ , the decay rate. Radioactive decay is a special case of *one-step* or *birth and death* processes. These stochastic processes can be modeled as taking place one event at a time so that the transition rates have the form

$$w(n'/n) = \lambda(n) \delta_{n',n-1} + \gamma(n) \delta_{n',n+1},$$

where $\lambda(n)$ is the death rate and $\gamma(n)$ the birth rate. If n is restricted to the positive integers, one usually has what is called a “natural boundary” at $n = 0$,

$$\lambda(n) = 0.$$

In this case the steady state is readily found, for if $\partial_t P_n(t) = 0$ for all n , one must have

$$\lambda(n) P_n = \gamma(n-1) P_{n-1}.$$

One can therefore express P_n in terms of P_0 through

$$P_n = \frac{\gamma(n-1) \gamma(n-2) \cdots \gamma(0)}{\lambda(n) \lambda(n-1) \cdots \lambda(1)} P_0,$$

and determine P_0 from normalization

$$\sum_n P_n = 1.$$

Birth and death processes are usually solved (when possible) by making use of the *generating function*

$$F(z, t) = \sum_{n=0}^{\infty} P_n(t) z^n. \quad (2.4)$$

From the generating function, we can find the moments by taking derivatives with respect to z and then setting $z = 1$. For example, the expectation of n is

$$\langle n \rangle = \left. \frac{\partial F(z, t)}{\partial z} \right|_{z=1}.$$

Often the moments

$$\phi_m \equiv \langle n(n-1)(n-2) \cdots (n-m+1) \rangle = \left. \frac{\partial^m F(z, t)}{\partial z^m} \right|_{z=1} \quad (2.5)$$

are referred to as the *factorial moments*. From the factorial moments, one can always get back the probabilities

$$P_n = \frac{1}{n!} \sum_{m=n}^{\infty} \frac{(-1)^m}{m!} \phi_m$$

To illustrate the use of the generating function, we take our example of radioactive decay above and multiply both sides of the master equation in Eq. (2.3) by z^n . Next we sum over n to produce the equation

$$\frac{\partial}{\partial t} F(z, t) = -\lambda(z-1) \frac{\partial}{\partial z} F(z, t). \quad (2.6)$$

Equation (2.6) can be explicitly solved for arbitrary initial conditions using the method of characteristics (see appendix C, section 2.1 for an example of this method). For the present purposes, however, we'll simply look at one particular solution. Suppose we start with a Poisson distribution of particles so that

$$P_n(0) = \frac{n_0^n e^{-n_0}}{n!},$$

and hence

$$F(z, 0) = \exp[n_0(z-1)].$$

In this case the solution of Eq. (2.6) is easily verified

$$F(z, t) = \exp[n_0 e^{-\lambda t} (z-1)].$$

From this we see that

$$\langle n(t) \rangle = n_0 e^{-\lambda t},$$

and

$$\sigma \equiv \sqrt{\langle n^2(t) \rangle - \langle n(t) \rangle^2} = n_0 e^{-\lambda t}.$$

Another simple, yet important birth-death process, the branching process, is solved in detail in appendix C and further illustrates the use of generating functions.

2.2 The Fokker-Planck Equation.

The Fokker-Planck equation is a generalized “diffusion” partial differential equation (p.d.e) that comes in two slightly different forms. The most common form describes the time evolution of the probability distribution of continuous random variable, which we shall denote as $W(x, t)$. This equation was first used by Fokker [14] and Planck [15] to describe the Brownian motion of particles. A very thorough reference on applications and methods for solving the Fokker-Planck equation is the book by Risken [16]. Risken also shows how the Fokker-Planck equation can be regarded as a special form of the master equation for continuous random variables.

2.2.1 The forward equation.

The more general form of the Fokker-Planck equation is also known as the *forward Kolmogorov equation* and is written not for $W(x, t)$, but in terms of the *conditional probability density* $P(x, t|x_0, t_0)$, which is the probability density for x at time $t \geq t_0$, given the initial condition $W(x, t_0) = \delta(x - x_0)$. The distinction between these two forms is necessary at this point because in the discussion that follows we will introduce the *backward equation*—another equation for the conditional probability $P(x, t|x_0, t_0)$.

Following Risken, we write the forward equation as

$$\frac{\partial}{\partial t} P(x, t|x_0, t_0) = -\frac{\partial}{\partial x} A(x) P(x, t|x_0, t_0) + \frac{1}{2} \frac{\partial^2}{\partial x^2} B(x) P(x, t|x_0, t_0), \quad (2.7)$$

or defining the linear differential operator

$$\mathcal{L}_{\text{FP}}(x) = -\frac{\partial}{\partial x} A(x) + \frac{1}{2} \frac{\partial^2}{\partial x^2} B(x), \quad (2.8)$$

we can write

$$\frac{\partial}{\partial t} P(x, t|x_0, t_0) = \mathcal{L}_{\text{FP}}(x) P(x, t|x_0, t_0), \quad (2.9)$$

We will often refer to the operator \mathcal{L}_{FP} as the *Liouville operator* of the Fokker-Planck equation. Multiplying Eq. (2.9) by $W(x_0, t_0)$, integrating over x_0 , and recognizing that

$$W(x, t) = \int P(x, t|x_0, t_0) W(x_0, t_0) dx_0,$$

produces the more commonly known Fokker-Planck equation for $W(x, t)$,

$$\frac{\partial}{\partial t} W(x, t) = \mathcal{L}_{\text{FP}}(x) W(x, t), \quad (2.10)$$

The first term on the right hand side of Eq. (2.8), containing $A(x)$, is referred to as the “drift” term. The second term, containing $B(x)$, is called the “diffusion” term. In the most simple diffusion processes $B(x)$ is simply a constant.

Loosely speaking, the two terms have a relatively simple interpretation. Suppose we start with the initial state above, $W(x, t_0) = \delta(x - x_0)$. $A(x_0)$ describes the instantaneous rate of change of the expectation of $x - x_0$ (*i.e.* $\langle x \rangle - x_0$), while $B(x)$ gives the instantaneous rate of change of the variance of x (*i.e.* $\langle x^2 \rangle - x_0^2$). A steady state describing fluctuations around x_0 is achieved if $A(x_0)$ is such that it focuses trajectories back to x_0 , counteracting the spreading behaviour of trajectories caused by the diffusion term.

The Fokker-Planck equation can be generalized to M variables $x_1 \dots x_M = \{x\}$ and has the form

$$\frac{\partial}{\partial t} W(\{x\}, t) = \left[- \sum_{i=1}^M \frac{\partial}{\partial x_i} A_i(\{x\}) + \sum_{i,j=1}^M \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} B_{ij}(\{x\}) \right] W(\{x\}, t). \quad (2.11)$$

In general, the drift vector $A_i(\{x\})$ and diffusion tensor $B_{ij}(\{x\})$ depend on all the variables $\{x\}$.

2.2.2 Solution of the Fokker-Planck equation.

The Fokker-Planck equation is obviously linear in $W(\{x\}, t)$, however, two classes of Fokker-Planck equations are often discussed in the literature. These go by the names “linear” and “non-linear”. This terminology can be somewhat confusing since the two categories are not mutually inclusive and refer to different aspects of the Fokker-Planck equation. This distinction is important to make so it is included in this brief overview.

A *linear* Fokker-Planck equation of the form of Eq. (2.11) is one where the drift vector $A_i(\{x\})$ is linear in the variables $x_1 \dots x_M = \{x\}$ and the diffusion tensor $B_{ij}(\{x\})$ is constant (*i.e.* independent of $\{x\}$, but still a tensor). Linear Fokker-Planck equations have Gaussian distributions for the stationary as well as non-stationary solutions and this is true even if $A_i(\{x\})$ is explicitly time-dependent. A *non-linear* Fokker-Planck equation has a diffusion tensor $B_{ij}(\{x\})$ that is explicitly dependent on at least one of the variables of $\{x\}$. A more appropriate term for this type of stochastic process is *heterogeneous diffusion*.

Fokker-Planck equations describing heterogeneous diffusion are difficult to solve except for the single variable case, where a change of variables produces a constant diffusion term and a modified drift term. Conceptually one can think of the heterogeneous

diffusion as giving rise to “fluctuation generated forces”. For the single-variable Fokker-Planck equation of Eq. (2.10), the steady state distribution $W_S(x)$ can always be found,

$$W_S(x) = \frac{C}{B(x)} \exp \left[2 \int^x \frac{A(y)}{B(y)} dy \right], \quad (2.12)$$

where C is some normalization constant. This allows us to define a potential $U(x)$,

$$U(x) = \ln B(x) - 2 \int^x \frac{A(y)}{B(y)} dy$$

so that

$$W_S(x) = C \exp(-U(x)).$$

2.2.3 The backward equation.

The backward Kolmogorov equation, or *adjoint equation*, as it is sometimes called, is also an equation for the conditional probability $P(x, t|x_0, t_0)$. However, the partial derivatives are with respect to the initial time t_0 and initial “position” x_0 . We write it as

$$\frac{\partial}{\partial t_0} P(x, t|x_0, t_0) = \mathcal{L}_{\text{FP}}^+(x_0) P(x, t|x_0, t_0), \quad (2.13)$$

and the adjoint operator $\mathcal{L}_{\text{FP}}^+(x_0)$ corresponding to $\mathcal{L}_{\text{FP}}(x)$ defined in Eq. (2.8) is given by

$$\mathcal{L}_{\text{FP}}^+(x_0, t_0) = A(x_0) \frac{\partial}{\partial x_0} + \frac{1}{2} B(x_0) \frac{\partial^2}{\partial x_0^2}. \quad (2.14)$$

The adjoint operator is obtained by: (i) changing the sign of the drift term, (ii) reversing the order of differentiation and multiplication by $A(x_0)$ and $B(x_0)$, (iii) changing the partial derivatives with respect to x to those with respect to x_0 . We should point out that the backward equation is not peculiar to stochastic processes described by a Fokker-Planck equation. The master equation 2.1 is a *forward equation*. One can also construct a backward equation. However, the relation between forward and backward

operators for the master equation is slightly more convoluted than the relation of $\mathcal{L}_{\text{FP}}^+(x_0)$ to $\mathcal{L}_{\text{FP}}(x)$ for the Fokker-Planck equation given above.

The backward equation is very useful for *first-passage time* problems. The first-passage time can be defined in a couple of different ways. The first is as the time at which the stochastic variable $x(t)$ first leaves a given domain [17][18]. For example, suppose the domain is the interval $x \in [x_1, x_2]$. For realizations of $x(t)$ that start at $t = 0$, the first-passage time is the time when $x(t)$ reaches either x_1 or x_2 for the first time. The second definition is more relevant for systems with absorbing states. Absorbing states are present if both $A(x)$ and $B(x)$ vanish for some value of x . In this case, one usually defines the first-passage time as the time at which $x(t)$ hits a particular absorbing state. As an example of the latter, suppose x_1 and x_2 are absorbing states. The the probability $u(x_0, t)$ of hitting one of the absorbing states in a time interval t satisfies an equation similar to Eq. (2.13),

$$\frac{\partial}{\partial t} u(x_0, t) = \mathcal{L}_{\text{FP}}^+(x_0) u(x_0, t). \quad (2.15)$$

The probability for hitting x_1 is obtained by solving Eq. (2.15) with the boundary conditions

$$u(x_1, t) = 1,$$

$$u(x_2, t) = 0.$$

Conversely, the probability of hitting x_2 is obtained from the boundary conditions

$$u(x_1, t) = 0,$$

$$u(x_2, t) = 1.$$

2.2.4 System size expansions: from master equation to Fokker-Planck equation.

I have yet to describe how one arrives at a Fokker-Planck description for a given stochastic process. In the following section I discuss one method that relies on the equivalence of the Langevin equation to the Fokker-Planck equation. This approach, however, does not go beyond a merely phenomenological description of the fluctuations. Ideally, one should start with a master equation, written in terms of microscopic configurations and transition probabilities, and systematically derive a Fokker-Planck equation for some sort of coarse-grained variables. Along these lines, van Kampen [19][20][21] has formulated a general technique for treating the birth-death and one-step processes described in the previous section which involves an expansion of the master equation in inverse powers of Ω , which is a measure of the system size (Ω is usually the volume).

The Ω -expansion of van Kampen begins with the ansatz that fluctuations in a microscopic variable n should be $\propto \sqrt{\Omega}$. This ansatz is quite reminiscent of the Central Limit Theorem for random variables. This situation for a single variable is described by writing

$$n = \Omega\phi(t) + \sqrt{\Omega}x, \quad (2.16)$$

where $\phi(t)$ is the non-fluctuating “macroscopic” variable and x , a fluctuating continuous variable with probability distribution $W(x, t)$. Van Kampen shows that to $\mathcal{O}(\Omega^{-1})$, $\phi(t)$ obeys a deterministic ordinary differential equation that is precisely the one obeyed by the expectation $\langle n(t) \rangle / \Omega$ under the stochastic process. The fluctuating part described by $W(x, t)$ satisfies a linear Fokker-Planck equation. Furthermore, the expansion is systematic in that its validity is readily apparent. The appropriateness of the expansion can be briefly summarized by saying that stability of the differential equation for $\phi(t)$ in the standard dynamical systems sense implies the validity of the expansion for large

enough system size. The reader should consult the references above to determine just what “large enough” means [19][20][21]. The important point to make here is that van Kampen’s method fails near the instabilities, or critical points as they are called in statistical physics. This is unfortunate since these points of the phase diagram are precisely those we would like to study.

So the question remains, does there exist a Fokker-Planck representation of the master equation which is asymptotic in system size but retains an accurate description of the fluctuations, even near critical points? By accurate we mean both qualitatively in that the phase diagram is preserved, and quantitatively in that the critical exponents characterizing the instabilities are unchanged. The answer is emphatically yes, as was pointed out by Horsthemke and Brenig [22][23]. Their work was based on the mathematically rigorous results of Kurtz [24][25] who proved similar results for the asymptotic representation of stochastic processes by stochastic differential equations (s.d.e.’s). Horsthemke and Brenig’s results are simple to state, in contrast to the change of variables of Eq. (2.16), one makes the change of variables

$$n = \Omega x.$$

In other words, we transform to the intensive variable $x = n/\Omega$. Expansion of the master equation now produces a non-linear Fokker-Planck equation that preserves the character of the fluctuations near unstable critical points. We will illustrate this type of expansion in section 3.2.1 when we derive a non-linear diffusion approximation for Moran’s model.

2.3 Langevin Equations.

Many stochastic processes are formulated as stochastic differential equations (s.d.e.’s). Such processes are most easily visualized in terms of ordinary differential equations (o.d.e.’s) subject to random perturbations or “noise”. In the physics literature first

order s.d.e.'s are often referred to as *Langevin equations* due to the work of Langevin [26], who also investigated Brownian motion.

2.3.1 A single variable Langevin equation.

A single variable Langevin equation is usually written in the physics literature as

$$\frac{dx(t)}{dt} = A(x) + \sqrt{B(x)}\eta(t), \quad (2.17)$$

where $A(x)$ is the deterministic force and $\eta(t)$, a Gaussian distributed random impulse with zero mean and auto-correlation

$$\langle \eta(t) \eta(t') \rangle = \delta(t - t'). \quad (2.18)$$

The main advantage of this formulation is its conceptual simplicity. It is easy to simulate a Langevin equation by first integrating the deterministic part for a time interval Δt , followed by adding the random “kick” of $\sqrt{B(x)}\eta(t) \Delta t$. The disadvantages of using Langevin equations are twofold. First of all, Eq. (2.17) as written, is ambiguous. This point is discussed in detail by van Kampen [27]. The problem is that $\eta(t)$ is a stochastic process with discrete jumps. When $B(x)$ is not a constant, one must choose what value of x is used to compute the noise. Should it be the value before or after the jump? Or possibly an average of the two? The Itô interpretation [28] is to use the value before the jump,

$$x(t + \delta t) = x(t) + A(x) \delta t + \sqrt{B(x(t))} \int_t^{t+\delta t} \eta(t') dt',$$

and is typically appropriate for internal fluctuations, like those of birth-death processes. One can show that this interpretation of Eq. (2.17) is equivalent to the Fokker-Planck equation

$$\frac{\partial}{\partial t} W(x, t) = -\frac{\partial}{\partial x} A(x) W(x, t) + \frac{1}{2} \frac{\partial^2}{\partial x^2} B(x) W(x, t).$$

Alternatively, one can use the Stratonovich interpretation [29] and take the average

$$x(t + \delta t) = x(t) + A(x) \delta t + \sqrt{B\left(\frac{x(t + \delta t) + x(t)}{2}\right)} \int_t^{t+\delta t} \eta(t') dt'.$$

The latter interpretation is appropriate for external noise, for example, a system coupled to an environment where the Gaussian noise is never really “white” (delta-function correlated), but has a non-zero relaxation time. One can show that this interpretation leads to the Fokker-Planck equation

$$\frac{\partial}{\partial t} W(x, t) = -\frac{\partial}{\partial x} A(x) W(x, t) + \frac{1}{2} \frac{\partial}{\partial x} \sqrt{B(x)} \frac{\partial}{\partial x} \sqrt{B(x)} W(x, t).$$

The second limitation of the Langevin equation formulation is that the form of the noise term $B(x)$ must come from a more detailed knowledge of the stochastic process. As pointed out by van Kampen [27], both of these problems relating to specification of the noise are resolved if one starts with a microscopic description in terms of the master equation and proceeds using the system size expansion discussed in the previous section to arrive at a Fokker-Planck equation. One can then use the equivalence described above to get the appropriate Langevin equation and interpretation. This procedure is convenient for simulating a Fokker-Planck equation using the Langevin formulation.

2.3.2 Continuum Langevin equations.

The Langevin equation defined by Eq. (2.17) describes a single variable, or degree of freedom, and is easily generalized to several degrees of freedom by using a system of o.d.e.’s supplemented with appropriate auto and cross correlators for the noise terms. This procedure can be further extended to continuous systems with infinitely many degrees. For concreteness, a large class of continuum Langevin equations for a field $\phi(\mathbf{r}, t)$ in d spatial dimensions has the form

$$\frac{\partial}{\partial t} \phi(\mathbf{r}, t) = F[\phi(\mathbf{r}, t)] + \sqrt{B[\phi(\mathbf{r}, t)]} \eta(\mathbf{r}, t). \quad (2.19)$$

Here the force term $F[\phi(\mathbf{r}, t)]$ and the noise strength $B[\phi(\mathbf{r}, t)]$ are *functionals* of $\phi(\mathbf{r}, t)$. The force term $F[\phi(\mathbf{r}, t)]$ is the *functional derivative* of a potential

$$F[\phi(\mathbf{r}, t)] = -\frac{\delta U[\phi(\mathbf{r}, t)]}{\delta \phi(\mathbf{r}, t)}, \quad (2.20)$$

and the Gaussian noise term $\eta(\mathbf{r}, t)$ is specified by

$$\langle \eta(\mathbf{r}, t) \eta(\mathbf{r}', t') \rangle = \delta(r_1 - r'_1) \delta(r_2 - r'_2) \cdots \delta(r_d - r'_d) \delta(t - t'). \quad (2.21)$$

In this thesis, $\phi_{\mathbf{r}}(t)$ will be synonymous with $\phi(\mathbf{r}, t)$. This choice of notation serves to emphasize that \mathbf{r} is to be regarded as an index for the stochastic variables $\phi_{\mathbf{r}}(t)$. For convenience we will often write $\phi_{\mathbf{r}}$ and the t -dependence will be implied. The same notation will be used when \mathbf{r} is a vector of some d -dimensional lattice. In this case the functional derivative in Eq. (2.20) will simply be a partial derivative, *i.e.*

$$F(\phi_{\mathbf{r}}) = -\frac{\partial U(\phi_{\mathbf{r}})}{\partial \phi_{\mathbf{r}}} \quad (2.22)$$

For many stochastic models, especially those found in statistical physics, the magnitude of the noise is independent of $\phi(\mathbf{r}, t)$, *i.e.* $B[\phi(\mathbf{r}, t)]$ equals a constant which I denote here Γ . For these models, the steady-state (t -independent) distribution of $\phi_{\mathbf{r}}$, $W[\phi_{\mathbf{r}}]$ can be shown to be

$$W[\phi_{\mathbf{r}}] = Z^{-1} \exp(-U[\phi_{\mathbf{r}}]/\Gamma), \quad (2.23)$$

and the normalization Z , in the continuum limit is given by the *path-integral or functional integral* obtained by integrating over all possible realizations of $\phi_{\mathbf{r}}$. In this thesis the notation $\int [d\phi]$ will be used to represent functional integrals. Z is also known as the *partition function* and is given by

$$Z = \int [d\phi] \exp(-U[\phi_{\mathbf{r}}]/\Gamma).$$

As a concrete example we consider the so-called ϕ^4 model of equilibrium statistical physics. This model correctly describes the critical behaviour of Ising-like systems. In this case U is the Landau-Ginzburg free energy

$$U[\phi_{\mathbf{r}}(t)] = \int d^d\mathbf{r} \left[\frac{D}{2} (\nabla\phi_{\mathbf{r}})^2 + \tilde{t}\phi_{\mathbf{r}}^2 + u\phi_{\mathbf{r}}^4 \right], \quad (2.24)$$

where \tilde{t} is a linear function of the temperature and u is a constant. It turns out that there is more than one way to extend dynamics to this model. Here we discuss “model A” from the literature of equilibrium critical phenomena [30][31] (see also chapter 10 of [32]). The thermodynamic force is

$$F[\phi_{\mathbf{r}}] = D\nabla^2\phi_{\mathbf{r}} - 2\tilde{t}\phi_{\mathbf{r}} - 4u\phi_{\mathbf{r}}^3, \quad (2.25)$$

and to be consistent with equilibrium statistical mechanics, we must set

$$\Gamma = k_b T.$$

2.4 Operators and Path Integrals.

The previous sections of this chapter reviewed classical methods for the formulation and solution of stochastic processes. These methods are adequate for the treatment of stochastic models having only a few degrees of freedom but become extremely cumbersome for the analysis of stochastic processes defined on the lattice or in the continuum. Beginning in the 1970's, two different formalisms emerged for the treatment of these types of many-body problems. The response functional formalism is perhaps the most conceptually straightforward and takes as its starting point the lattice or continuum Langevin equation discussed in the preceding section. The Langevin equation is recast as a functional integral from which a consistent perturbation theory and other diagrammatic methods can be applied [33]. For the study of spatial versions of birth-death

or one-step processes introduced in section 2.1, a more exotic field-theoretic method is available that uses creation and annihilation operators similar to those of non-relativistic quantum field theory. We give a brief overview of these two approaches in this section. In addition, a third formulation is discussed that involves writing the Fokker-Planck equation in terms of operators that are similar to the creation and annihilation operators for birth-death processes. In contrast to birth-death processes, the operators of this theory represent continuous stochastic variables. This method has not been extensively discussed in the literature and represents the investigation of the author. We show that this is an alternative to the response functional method for obtaining a functional integral but has the added advantage of retaining the operator character when needed.

2.4.1 Response functional formalism.

The response functional formalism was introduced by Martin, Siggia and Rose [34] who applied it to the problem of fluid critical behaviour. It was further developed for renormalization group calculations of equilibrium critical dynamics by Janssen [35] and Bausch, Janssen and Wagner [36]. A nice introduction to this method can be found in chapter 10 of the book by Cardy [32]. An extremely thorough exposition of this method and the underlying mathematics is found in the book by Zinn-Justin [33].

Derivation of the functional integral.

One starts with a Langevin equation like that given in Eq. (2.19) but now we add a source term $h_{\mathbf{r}}(t)$ and for simplicity we will consider the case of constant noise. We have

$$\frac{\partial}{\partial t} \phi_{\mathbf{r}}(t) = F[\phi_{\mathbf{r}}(t)] + h_{\mathbf{r}}(t) + \sqrt{\Gamma} \eta(\mathbf{r}, t) \quad (2.26)$$

and use this to construct a functional integral. The construction of the functional integral can be viewed as an application of the transformation of random variables. Our

aim is to calculate averages and correlation functions of quantities depending on $\phi_{\mathbf{r}}(t)$. For example we may want to calculate $G_{\phi\phi}$, the two-point correlation function which is given by the expectation

$$G_{\phi\phi}(\mathbf{r}, t; \mathbf{r}', t) = \langle \phi_{\mathbf{r}}(t) \phi_{\mathbf{r}'}(t') \rangle. \quad (2.27)$$

This means we need to evaluate path-integrals over $\phi_{\mathbf{r}}(t)$. Furthermore, we need to include only those values of $\phi_{\mathbf{r}}(t)$ that satisfy Eq. (2.26). We move the terms on the right hand side of Eq. (2.26) to the left hand side and call the resulting expression $C(\mathbf{r}, t)$. This expression must vanish for all values of $\phi_{\mathbf{r}}(t)$ that satisfy the Langevin equation. We can accomplish this by writing $C(\mathbf{r}, t)$ as the argument of a delta function which results in the dynamic functional

$$Z[h_{\mathbf{r}}] = \int [d\phi] \delta[C(\mathbf{r}, t)] = \int [d\phi] [d\bar{\phi}] \exp\left(-\int \bar{\phi}_{\mathbf{r}}(t) C(\mathbf{r}, t) d^d\mathbf{r} dt\right) \quad (2.28)$$

This requires a *functional integral* because the delta function constraint has to be applied for each value of \mathbf{r} and t . We have introduced the $\bar{\phi}_{\mathbf{r}}(t)$ field for the delta function and the $\bar{\phi}_{\mathbf{r}}(t)$ integrals are taken along the imaginary axis. Now we can average Z over the noise as this is a Gaussian integral for each point \mathbf{r}, t . We multiply Z by $\exp(-\eta_{\mathbf{r}}^2/2\Gamma)$ and average over $\eta_{\mathbf{r}}$ at each point where we have

$$\int \exp\left(\eta_{\mathbf{r}} \bar{\phi}_{\mathbf{r}} - \frac{\eta_{\mathbf{r}}^2}{2\Gamma}\right) d\eta_{\mathbf{r}} \propto \exp\left(\frac{1}{2}\Gamma \bar{\phi}_{\mathbf{r}}^2\right),$$

and therefore we have produced the functional integral

$$Z[h_{\mathbf{r}}] = \int [d\phi] [d\bar{\phi}] \exp\left(-\int \left[\bar{\phi}_{\mathbf{r}}(t) C(\mathbf{r}, t) - \frac{1}{2}\Gamma \bar{\phi}_{\mathbf{r}}^2(t)\right] d^d\mathbf{r} dt\right) \quad (2.29)$$

The result is that expectations are evaluated with respect to an ensemble weighted by $\exp(-S[\phi_{\mathbf{r}}, \bar{\phi}_{\mathbf{r}}])$, where we have defined the *action* as

$$S[\phi_{\mathbf{r}}, \bar{\phi}_{\mathbf{r}}] = \int \left(\bar{\phi}_{\mathbf{r}} \partial_t \phi_{\mathbf{r}} - \bar{\phi}_{\mathbf{r}} F[\phi_{\mathbf{r}}] - \bar{\phi}_{\mathbf{r}} h_{\mathbf{r}} - \frac{1}{2}\Gamma \bar{\phi}_{\mathbf{r}}^2\right) d^d\mathbf{r} dt. \quad (2.30)$$

(For simplicity we have dropped the explicit time-dependence of the fields). Another term that is often used instead of action is the *Lagrangian density*, or simply *Lagrangian*, which is the quantity in Eq. (2.30) that is integrated over space and time.

Correlation and response functions.

We will need to evaluate expectations of ϕ and $\bar{\phi}$ fields for arbitrary times and positions. For example, the most informative correlation is the *density-density correlation function*, which we also called the two-point correlation function in Eq. (2.27). The term *spatial auto-correlation* is often found in the ecological literature and is synonymous with the two-point correlation function. Another fundamental quantity is the $G_{\phi\bar{\phi}}$, the *response function*, which is defined by the expectation

$$G_{\phi\bar{\phi}}(\mathbf{r}, t; \mathbf{r}', t') = \langle \phi_{\mathbf{r}}(t) \bar{\phi}_{\mathbf{r}'}(t') \rangle, \quad (2.31)$$

and for many-body systems is also called the *two-point Green's function*, *two-point propagator*, or simply *propagator*. The response function or propagator of Eq. (2.31) describes the response of the density field $\phi_{\mathbf{r}}(t)$ to a delta function perturbation at the point (\mathbf{r}', t') . This gives us a nice interpretation for the $\bar{\phi}_{\mathbf{r}}$ field introduced in the derivation of the functional integral; it is often called the response field.

The expectation of any number of ϕ or $\bar{\phi}$ fields can, at least formally if not practically, be found if we introduce another source term, $\phi_{\mathbf{r}}\bar{h}_{\mathbf{r}}$, to the action in Eq. (2.30),

$$S[\phi_{\mathbf{r}}, \bar{\phi}_{\mathbf{r}}] = \int d^d\mathbf{r}dt \left(\bar{\phi}_{\mathbf{r}}\partial_t\phi_{\mathbf{r}} - \bar{\phi}_{\mathbf{r}}F[\phi_{\mathbf{r}}] - \bar{\phi}_{\mathbf{r}}h_{\mathbf{r}} - \phi_{\mathbf{r}}\bar{h}_{\mathbf{r}} - \frac{1}{2}\Gamma\bar{\phi}_{\mathbf{r}}^2 \right). \quad (2.32)$$

Expectations can now be found by successive functional differentiation of $Z[h_{\mathbf{r}}, \bar{h}_{\mathbf{r}}]$,

$$\begin{aligned} G_{m,n} &= \langle \phi_{\mathbf{x}_1}(t_1) \phi_{\mathbf{x}_2}(t_2) \dots \phi_{\mathbf{x}_m}(t_m) \bar{\phi}_{\mathbf{y}_1}(t'_1) \bar{\phi}_{\mathbf{y}_2}(t'_2) \dots \bar{\phi}_{\mathbf{y}_n}(t'_n) \rangle \\ &= \frac{\delta}{\delta \bar{h}_{\mathbf{x}_1}(t_1)} \frac{\delta}{\delta \bar{h}_{\mathbf{x}_2}(t_2)} \dots \frac{\delta}{\delta \bar{h}_{\mathbf{x}_m}(t_m)} \frac{\delta}{\delta h_{\mathbf{y}_1}(t'_1)} \frac{\delta}{\delta h_{\mathbf{y}_2}(t'_2)} \dots \frac{\delta}{\delta h_{\mathbf{y}_n}(t'_n)} Z[h_{\mathbf{r}}, \bar{h}_{\mathbf{r}}] \Big|_{h, \bar{h}=0}, \end{aligned} \quad (2.33)$$

where $h, \bar{h} = 0$ means we have set all h and \bar{h} fields to zero everywhere. Knowledge of the full set of correlation and response functions is equivalent to a full solution of the Langevin equation combined with the required boundary conditions. The full solution is a probability functional that even in the best case, that of an explicit closed form solution, is usually much too cumbersome for any practical calculation. On the other hand, the correlation and response functions, especially the two-point correlation function and propagator, are invaluable for understanding the behaviour of the stochastic model.

Many-body perturbation theory.

In general, the action $S[\phi_{\mathbf{r}}, \bar{\phi}_{\mathbf{r}}]$ for a given Langevin equation is not sufficiently simple to put $Z[h_{\mathbf{r}}, \bar{h}_{\mathbf{r}}]$ in a form useful calculating correlation and response functions using Eq. (2.33). The usual method in field theory is to separate $S[\phi_{\mathbf{r}}, \bar{\phi}_{\mathbf{r}}]$ into a bi-linear, or Gaussian part, S_0 and the remaining or “interacting” part S_I . For a particular correlation or response function, one obtains a perturbation expansion containing correlation and response functions that are evaluated with respect to S_0 as prescribed by Wick’s theorem. To simplify the book-keeping, each term of the expansion can be represented by so-called Feynman diagrams. I will not be able to describe this procedure in detail in this thesis. The reader is encouraged to find the details in one of the numerous standard texts covering statistical field theory or zero-temperature, non-relativistic quantum field theory. Some recommendations are the books by Binney, *et. al.* [37] and Goldenfeld [38]. The reader is cautioned, however, that these books concern the time-independent theory of equilibrium statistical mechanics—there are some subtle differences between this and the field theory described in this section. A much more appropriate treatment for this thesis is the book by Zinn-Justin [33] mentioned in the introduction of this section. A more entertaining exposition of Feynman diagrams can be found in the book by Mattuck [39].

I do illustrate the application of field-theoretic methods for the Stepping-Stone model

in more detail in section 6 of chapter 4. It is unfortunate that such heavy machinery is needed to develop the diagrammatic techniques that in the end are quite simple and almost fun to use.

2.4.2 The master equation in occupation number representation.

Master equations for stochastic birth-death processes can be described in terms of an occupation number representation that has been given several other names in the literature including “second quantization representation”, “Fock-space methods” and “quantum field methods”. These names emphasize the fact that the occupation number representation borrows heavily from the mathematical framework developed for the quantum-mechanical many-body problem. In addition to stochastic population models, these techniques are used for the closely related stochastic models of reaction-diffusion processes. The occupation number representation was introduced by Doi [40] and used to develop the theory of diffusion-controlled reactions and in particular the reaction-diffusion model dubbed $A + A \rightarrow B$ [41]. This model was also studied by Zel’dovich and Ovchinnikov using similar techniques [42]. The formulation used by these authors was further elaborated by Grassberger and Scheunert who applied it to several birth-death models [43]. Their article presents a good review of the underlying mathematical structure of the theory. The theory has subsequently been applied to many reaction-diffusion models. A recent review of this theory can be found in the article by Mattis and Glasser [44] who also provide an extensive bibliography of many applications and related works. Here we give a brief review of the formalism. Although the formalism is more appropriate for lattice or spatial models having many degrees of freedom, for simplicity of presentation we initially discuss the formalism as applied to a single discrete variable representing the number of particles or individuals of a stochastic birth-death process and follow this with the extension to many degrees of freedom.

The single-variable master equation.

In section 2.1 we introduced birth-death processes, where the state of the system is specified by the discrete probabilities, P_n , of having n -individuals. The main idea of the occupation number formalism is to represent the state of the system specified by the P_n 's as a vector $|\psi\rangle$ in an abstract space spanned by the basis states $|n\rangle$ that correspond to exactly n individuals,

$$|\psi\rangle = \sum_n P_n |n\rangle. \quad (2.34)$$

The $|n\rangle$ basis states are normalized by defining the scalar product as

$$\langle n|m\rangle = n! \delta_{n,m}, \quad (2.35)$$

so that

$$P_n = \frac{1}{n!} \langle n|\psi\rangle. \quad (2.36)$$

The choice of scalar product in Eq. (2.35) will be convenient for stochastic processes describing identical particles. Using the vector space we have constructed, the master equation

$$\frac{\partial}{\partial t} P_n(t) = \sum_m \{w(n/m) P_m(t) - w(m/n) P_n(t)\} \quad (2.37)$$

takes the form

$$\frac{\partial}{\partial t} |\psi\rangle = \mathcal{L} |\psi\rangle, \quad (2.38)$$

where we have introduced the Liouville operator \mathcal{L} which, using the occupation number representation, is expressed in terms of creation and annihilation operators, a^\dagger and a . The a^\dagger and a operators are defined by their action on the $|n\rangle$ basis states,

$$\begin{aligned} a |n\rangle &= n |n-1\rangle \\ a^\dagger |n\rangle &= |n+1\rangle. \end{aligned} \quad (2.39)$$

They obey the “canonical” commutation relations

$$[a, a^\dagger] = 1,$$

where the commutator $[\mathcal{A}, \mathcal{B}]$ of two operators, \mathcal{A} and \mathcal{B} , is defined as it is in quantum mechanics

$$[\mathcal{A}, \mathcal{B}] = \mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A}.$$

We see that in terms of these operators, $|n\rangle = (a^\dagger)^n |0\rangle$ and the state of zero particles is defined by

$$a |0\rangle = 0.$$

The normalization requirement, $\sum_n P_n = 1$, takes the form

$$\langle S | \psi \rangle = 1, \tag{2.40}$$

where $|S\rangle$ is defined by

$$\begin{aligned} |S\rangle &= \exp(a^\dagger) |0\rangle \\ &= \sum_n \frac{(a^\dagger)^n}{n!} |0\rangle. \end{aligned} \tag{2.41}$$

In statistical physics and quantum mechanics, $|S\rangle$ is known as a “coherent state” or “Glauber state”.

The normalization requirement of Eq. (2.40) places a constraint on the form of \mathcal{L} . Since $\langle S | \psi \rangle = 1$ must be preserved by the dynamics, *i.e.*

$$\frac{\partial}{\partial t} \langle S | \psi \rangle = \langle S | \mathcal{L} | \psi \rangle = 0 \tag{2.42}$$

for an arbitrary state $|\psi\rangle$. Therefore, one has the requirement

$$\langle S | \mathcal{L} = 0. \tag{2.43}$$

This implications of this requirement can be found by considering a process that takes n particles into m particles. Obviously we need a term with $w (a^\dagger)^m a^n$, where w is the transition rate. The requirement of Eq. (2.43) can be satisfied if we add to this a term with same number of creation and annihilation operators so that the Liouville operator for this process becomes

$$\mathcal{L}_{m,n} = w [(a^\dagger)^m a^n - (a^\dagger)^n a^n]. \quad (2.44)$$

These two terms correspond to the two terms in the master equation, Eq. (2.37) that represent the flow of probability into and out of the state of n particles. The form of Eq. (2.44) can be verified to satisfy the requirement of Eq. (2.43) by noting that

$$\langle S | a^\dagger = \langle S |. \quad (2.45)$$

Interestingly enough, using the above definitions of operators and states, Eq. (2.38) can be shown to simultaneously be a representation of the master equation, the forward equation *and* the backward equation.

The generating function.

We now make contact with the generating function introduced in Eq. (2.4). One can show that

$$F(z) = \langle z | \psi \rangle, \quad (2.46)$$

where the coherent state $|z\rangle$ is defined by

$$|z\rangle = \exp(za^\dagger) |0\rangle. \quad (2.47)$$

Coherent states are also eigenstates of a ,

$$a |z\rangle = z |z\rangle.$$

The normalization condition, $\langle S|\psi\rangle = 1$ becomes

$$F(z)|_{z=1} = \sum_n P_n = 1.$$

Using coherent states and the associated generating function, the creation and annihilation operators have a simple representation

$$a = \frac{\partial}{\partial z}, \quad (2.48)$$

and

$$a^\dagger = z. \quad (2.49)$$

This is a very practical result, for it allows one to very quickly derive the equation of motion for the generating function once \mathcal{L} is specified. For example, consider a branching process. This stochastic process is of fundamental importance for understanding stochastic population models. It consists of particles that can independently branch into two particles or decay and otherwise be removed from the system. The Liouville operator for the branching process is defined by

$$\mathcal{L}_{\text{BP}} = \lambda (a - a^\dagger a) + \gamma \left[(a^\dagger)^2 a - a^\dagger a \right], \quad (2.50)$$

where λ is the death rate and γ is the branching rate. By mere inspection, the generating function is found to satisfy the equation

$$\begin{aligned} \frac{\partial}{\partial t} F(z) &= \lambda \left(\frac{\partial}{\partial z} - z \frac{\partial}{\partial z} \right) F(z) + \gamma \left(z^2 \frac{\partial}{\partial z} - z \frac{\partial}{\partial z} \right) F(z) \\ &= (z-1)(\gamma z - \lambda) \frac{\partial}{\partial z} F(z). \end{aligned} \quad (2.51)$$

Expectations of operators.

A few remarks concerning expectations of operators are necessary at this point. For any operator \mathcal{A} , its expectation is given by a scalar product $\langle S|\mathcal{A}|\psi\rangle$. An important

operator is the number operator $\mathcal{N} = a^\dagger a$. Its expectation is the average number of particles

$$\begin{aligned}\langle n \rangle &= \langle S | a^\dagger a | \psi \rangle. \\ &= \langle S | a | \psi \rangle,\end{aligned}\tag{2.52}$$

where the last line follows from Eq. (2.45). Similarly, the k -th moment, μ_k is given by the expectation

$$\mu_k = \langle S | a^k | \psi \rangle.$$

The Schrödinger picture.

Just as in quantum mechanics, we can think of our state $|\psi\rangle$ as evolving in time and denote the time dependent state as $|\psi\rangle_t$. The formal solution of the master equation (2.37) is

$$|\psi\rangle_t = e^{\mathcal{L}(t-t_0)} |\psi\rangle_0.\tag{2.53}$$

The operator

$$\mathcal{Z}(t) = e^{\mathcal{L}t}$$

is the time-evolution operator that takes the initial state to the state at time t . If one can find the eigenvalue spectrum of \mathcal{L} , then the model is completely solved. Needless to say, this is only possible for the simplest of birth-death processes.

Heisenberg equations.

One can also view the time-evolution from the Heisenberg picture. Here the operators evolve in time and expectations are taken with respect to the initial state. The equation of motion for an operator \mathcal{A} is given by

$$\frac{\partial}{\partial t} \mathcal{A}(t) = [\mathcal{A}, \mathcal{L}],\tag{2.54}$$

and has the solution

$$\mathcal{A}(t) = e^{\mathcal{L}t} \mathcal{A}(0) e^{-\mathcal{L}t}. \quad (2.55)$$

An example of Heisenberg equations: the branching process.

Consider, once again, the branching process defined above by \mathcal{L}_{BP} in Eq. (2.50). First we calculate the average number of particles

$$\frac{d}{dt} a(t) = [a, \mathcal{L}]. \quad (2.56)$$

We evaluate the right-hand side of Eq. (2.56) by commuting the a operator all the way to the right. Using $[a, (a^\dagger)^2] = 2a^\dagger$, we find

$$\frac{d}{dt} a(t) = 2\gamma a^\dagger(t) a(t) - (\gamma + \lambda) a(t). \quad (2.57)$$

Taking the expectation, we find

$$\frac{d}{dt} \langle n(t) \rangle = (\gamma - \lambda) \langle n(t) \rangle, \quad (2.58)$$

which is trivially solved. For more complicated models, the Heisenberg equations couple higher moments. This does not occur for the average number of particles in the branching process since the particles behave independently.

Continuing with our example, the branching process, we would like to know $P_S(t)$, the time-dependent *survival probability*. This is the probability, $1 - P_0(t)$, that the system has not yet reached the state of zero particles. Recall that $P_0(t)$ is given by the product $\langle 0 | e^{\mathcal{L}_{\text{BP}}t} | \psi \rangle$. Suppose we start with only one particle. Then

$$\begin{aligned} P_0(t) &= \langle 0 | e^{\mathcal{L}_{\text{BP}}t} a^\dagger | 0 \rangle, \\ &= \langle 0 | e^{\mathcal{L}_{\text{BP}}t} a^\dagger e^{-\mathcal{L}_{\text{BP}}t} | 0 \rangle. \end{aligned} \quad (2.59)$$

Adding $e^{-\mathcal{L}_{\text{BP}}t}$ in the product of the last line of Eq. (2.59) is permitted since \mathcal{L}_{BP} is normal-ordered (all annihilation operators are to the right) and therefore

$$e^{-\mathcal{L}_{\text{BP}}t} |0\rangle = |0\rangle.$$

We recognize the operator in the middle of Eq. (2.59) as $a^\dagger(t)$, so

$$P_{\text{S}}(t) = 1 - P_0(t) = 1 - \langle 0 | a^\dagger(t) | 0 \rangle.$$

It is convenient to define the operator $\tilde{a} = a^\dagger - 1$. It has the same commutation relation with a as does a^\dagger ,

$$[a, \tilde{a}] = 1.$$

This gives us

$$P_{\text{S}}(t) = -\langle 0 | \tilde{a}(t) | 0 \rangle.$$

Rewriting \mathcal{L}_{BP} in terms of \tilde{a} ,

$$\mathcal{L}_{\text{BP}} = (\gamma - \lambda) \tilde{a}a + \gamma \tilde{a}^2a.$$

Using the Heisenberg equation for \tilde{a} , $d\tilde{a}/dt = [\tilde{a}, \mathcal{L}_{\text{BP}}]$, we find

$$\frac{d}{dt} P_{\text{S}}(t) = (\gamma - \lambda) P_{\text{S}}(t) - \gamma P_{\text{S}}^2(t).$$

The *ultimate survival probability* P_∞ is given by the asymptotic value of $P_{\text{S}}(t)$ for long times,

$$P_\infty = \frac{\gamma - \lambda}{\gamma}. \quad (2.60)$$

Branching processes have the interesting property that for positive growth rate $r = \gamma - \lambda$, there is a finite probability to go extinct, yet the average number grows exponentially! The time-dependence of $P_{\text{S}}(t)$ at the critical point, $\gamma - \lambda = 0$, is also interesting,

$$P_{\text{S}}(t) = \frac{1}{1 + \gamma t}, \quad (2.61)$$

which is a power law with exponent -1 . Away from the critical point, we find

$$P_S(t) = \frac{\gamma - \lambda}{\gamma - \lambda e^{-(\gamma - \lambda)t}}, \quad (2.62)$$

or an exponential decay to P_S on both sides of the critical point.

Many degrees of freedom.

The extension of the occupation number formalism to several variables and lattice models is quite straight-forward. All that is necessary is to give indices to the creation and annihilation operators introduced above. For lattice models, we will use the convention of indexing these operators by the vector \mathbf{r} which is decomposed into multiples of primitive lattice vectors. For example, on hypercubic lattices (linear, square, cubic, *etc.*)

$$\mathbf{r} = \sum_i^d n_i \mathbf{e}_i, \quad (2.63)$$

where n_i is an integer, and the primitive lattice vector \mathbf{e}_i is simply the lattice spacing l times the unit vector in the i th direction.

Commutators of operators at different sites are zero,

$$[a_{\mathbf{r}}, a_{\mathbf{r}'}^\dagger] = \delta_{\mathbf{r}, \mathbf{r}'}. \quad (2.64)$$

The normalization state $|S\rangle$ becomes

$$|S\rangle = \exp\left(\sum_{\mathbf{r}} a_{\mathbf{r}}^\dagger\right) |0\rangle,$$

and represents a coherent state at each site. As an example Liouville operator, consider a collection of random walkers that hop to nearest-neighbour sites with migration (jump) rate m . (Usually in the physics literature, m is used for indices. Here, we bend to the population genetics convention and use m to denote the migration or hop rate). The

Liouville operator for a random walk is given by

$$\mathcal{L}_{\text{RW}} = m \sum_{\mathbf{r}} \sum_{\mathbf{e}} a_{\mathbf{r}+\mathbf{e}}^\dagger a_{\mathbf{r}} + a_{\mathbf{r}-\mathbf{e}}^\dagger a_{\mathbf{r}} - 2a_{\mathbf{r}}^\dagger a_{\mathbf{r}}, \quad (2.65)$$

where the sum over \mathbf{e} is a shorthand for the sum over primitive lattice vectors. This model is solved in detail in appendix A using Heisenberg equations to find the Green's function or two-point propagator. One should note that in contrast to the single particle random walk master equation (2.2) presented in section 2.1, the Liouville operator \mathcal{L}_{RW} describes an arbitrary number of random walkers. (In the absence of other interactions, this number is constant of the motion).

For lattice models with many sites, and especially in the thermodynamic (infinite lattice) limit, one is usually not interested in obtaining the explicit probability of an arbitrary configuration, nor is this possible for a general \mathcal{L} . For the same reasons outlined in the discussion of the response functional formalism, one needs to find correlation functions and propagators. In the present context, correlation functions are expectations of the number of particles at various sites, *e.g.* the two-point correlation function,

$$G_{\phi\phi}(\mathbf{r}, t; \mathbf{r}', t') = \langle a_{\mathbf{r}}(t) a_{\mathbf{r}'}(t') \rangle. \quad (2.66)$$

Propagators for birth-death processes describe the response of the system at one point caused by the addition of new particles somewhere else, *e.g.* the two-point propagator defined by

$$G_{\phi\bar{\phi}}(\mathbf{r}, t; \mathbf{r}', t') = \langle a_{\mathbf{r}}(t) a_{\mathbf{r}'}^\dagger(t') \rangle. \quad (2.67)$$

Here we are using the Heisenberg picture and the correspondence of ϕ ($\bar{\phi}$) with a (a^\dagger) will become clear shortly. For translationally invariant systems, the spatial dependence in Eq.(2.66) and Eq.(2.67) can only be through the difference $\mathbf{r} - \mathbf{r}'$. This is also true for the time-dependence once the steady-state has been reached.

The correlation functions and propagators for interacting systems cannot in general be found in closed form. This is the same situation discussed with regard to the response functional formalism. One way of dealing with this problem is to pass from operators to a path-integral description and follow the same procedure as discussed for the response functional. I discuss this in the following section. An equivalent alternative is to stick with the operators and use the formalism of *time-ordered operators*, which also leads to diagrammatic expansions. I will not discuss this method in this thesis, but refer the reader to one of the many texts on quantum-field theory. A perennial classic is the book by Abrikosov, Gorkov and Dzyaloshinski [45]. Their discussion of zero-temperature methods is most relevant to the operator formalism discussed here.

Before we discuss the path-integral, we point how one can take the continuum limit using the operator formalism. In this limit the creation and annihilation operators obey the standard “Bose” commutation relations

$$\begin{aligned} [a_{\mathbf{r}}, a_{\mathbf{r}'}^\dagger] &= \delta^d(\mathbf{r} - \mathbf{r}'), \\ [a_{\mathbf{r}}, a_{\mathbf{r}'}] &= [a_{\mathbf{r}}^\dagger, a_{\mathbf{r}'}^\dagger] = 0, \end{aligned} \tag{2.68}$$

and the normalization state is given by

$$|S\rangle = \exp\left(\int d^d\mathbf{r} a_{\mathbf{r}}^\dagger\right) |0\rangle.$$

From operators to path-integrals.

We can construct a functional integral that resembles that of the response functional formalism by considering the time evolution operator using the coherent state representation. We will illustrate this for a single degree of freedom as the extension to many degrees of freedom does not involve any new concepts. A more detailed exposition of this construction in the context of quantum many-body theory can be found in the book

by Negele and Orland [46]. First, we need to know some additional properties of coherent states. We recall from Eq. (2.47) that a coherent state $|z\rangle$ is an eigenstate of the annihilation operator with eigenvalue z ,

$$a|z\rangle = z|z\rangle. \quad (2.69)$$

It is also the left eigenstate of the creation operator,

$$\langle z|a^\dagger = \langle z|\bar{z},$$

where \bar{z} is the complex conjugate of z . Coherent states are orthogonal, but overcomplete. Their scalar product is given by

$$\langle z'|z\rangle = e^{z'\bar{z}},$$

and therefore the resolution of the identity operator becomes

$$\mathcal{I} = \int \frac{d^2z}{\pi} e^{-z\bar{z}} |z\rangle \langle z|,$$

and the integral is taken over the complex plane.

The state at time t_f given the initial state at time t_i can be written

$$|\psi\rangle_{t_f} = e^{\mathcal{L}(t-t_0)} |\psi\rangle_{t_i}. \quad (2.70)$$

Using the coherent state representation, this relates the generating function at t_f to that at t_i by

$$F(z', t_f) = \int \frac{d^2z}{\pi} e^{-z\bar{z}} \mathcal{Z}(z', t_f; z, t_i) F(z, t_i),$$

with the time evolution kernel

$$\mathcal{Z}(z', t_f; z, t_i) = \langle z' | e^{\mathcal{L}(t_f-t_i)} | z \rangle.$$

If we divide the time interval $t_f - t_i$ into M equal steps of size $\varepsilon = \frac{t_f - t_i}{M}$, we can write

$$\langle z' | (e^{\varepsilon \mathcal{L}})^M | z \rangle = \mathcal{N}^{-1} \int \prod_{k=1}^{M-1} dz_k d\bar{z}_k e^{-z_k \bar{z}_k} \langle z_M | e^{\varepsilon \mathcal{L}} | z_{M-1} \rangle \langle z_{M-1} | e^{\varepsilon \mathcal{L}} | z_{M-2} \rangle \dots \langle z_1 | e^{\varepsilon \mathcal{L}} | z \rangle, \quad (2.71)$$

where $z_M = z$ and $\mathcal{N} = \pi^M$. Since \mathcal{L} is normal-ordered, *i.e.* all the creation operators are to the left, and the annihilation operators are to the right, we can therefore let them act on their respective eigenstates in each term of Eq. (2.71) with the result

$$\begin{aligned} \mathcal{Z}(z', t_f; z, t_i) &= \mathcal{N}^{-1} \lim_{M \rightarrow \infty} \prod_{k=1}^{M-1} dz_k d\bar{z}_k e^{-z_k \bar{z}_k} \\ &\times \exp \left[- \sum_{k=1}^{M-1} \bar{z}_k \left(\frac{z_k - z_{k-1}}{\varepsilon} \right) - L(\bar{z}_k, z_k) \right], \end{aligned} \quad (2.72)$$

where $L(\bar{z}_k, z_k)$ is obtained from the normal-ordered \mathcal{L} by replacing all creation operators by the field \bar{z}_k and annihilation operators by the field z_k . Taking the limit we have the functional integral

$$\mathcal{Z}(z', t_f; z, t_i) = \int_{z(t_i)=z}^{\bar{z}(t_f)=z'} [d\bar{z}] [dz] e^{-z' \bar{z}'} \exp(-S[\bar{z}(t), z(t)]), \quad (2.73)$$

with the action

$$S[\bar{z}(t), z(t)] = \int_{t_i}^{t_f} dt \{ \bar{z}(t) \partial_t z(t) - L[\bar{z}(t), z(t)] \}. \quad (2.74)$$

This result combined with the initial state given by $F(z, t_i)$ tells us more than we really need to know in most circumstances. We can imagine turning on the interactions at some finite time, or perturbing the system in some other way and let $t_f \rightarrow \infty$, and $t_i \rightarrow -\infty$, at which time the system will be found in the vacuum state or some other convenient steady-state. This allows us to ignore the initial state at t_i and the $e^{-z' \bar{z}'}$ term at t_f . The only complication is getting the correct “ground state”, or steady-state as the case may be. I will not discuss this technical but important problem in this introductory chapter.

If we now add the extra source terms as was done for the response functional formalism, all correlation functions can be found from the functional integral

$$\mathcal{Z}[h, \bar{h}] = \int [d\bar{z}] [dz] \exp(-S[\bar{z}(t), z(t)]), \quad (2.75)$$

with

$$S[\bar{z}(t), z(t)] = \int dt \{ \bar{z}(t) \partial_t z(t) - L[\bar{z}(t), z(t)] \}.$$

As we mentioned earlier, it is not much more difficult to repeat this derivation when one needs to formulate the path-integral for many degrees of freedom. For these problems, we use the notation $\bar{\phi}_{\mathbf{r}}(t)$ and $\phi_{\mathbf{r}}(t)$ for the fields that correspond to the creation and annihilation operators of the Liouvillian. To summarize the results, we find that we can construct a path integral description of a stochastic process specified by a Liouville operator \mathcal{L} , by introducing the action

$$S[\bar{\phi}_{\mathbf{r}}(t), \phi_{\mathbf{r}}(t)] = \int d^d \mathbf{r} dt \{ \bar{\phi}_{\mathbf{r}}(t) \partial_t \phi_{\mathbf{r}}(t) - L[\bar{\phi}_{\mathbf{r}}(t), \phi_{\mathbf{r}}(t)] \}, \quad (2.76)$$

where L is obtained from \mathcal{L} by substituting respectively, $\bar{\phi}_{\mathbf{r}}(t)$ and $\phi_{\mathbf{r}}(t)$ for the creation and annihilation operators $a_{\mathbf{r}}^{\dagger}$ and $a_{\mathbf{r}}$.

2.4.3 The Fokker-Planck equation using operators.

In this section, we develop an operator formalism for the Fokker-Planck equation in a similar manner as was done for the master equation. The operators we introduce here represent continuous random variables, in contrast to the discrete ones of the occupation number formalism. Once again, for simplicity of illustration, we initially consider a single degree of freedom. The extension to many degrees of freedom is obvious and will be discussed after we illustrate use of the formalism with an example.

The basic idea is to represent state of system described by the probability density $W(x)$ as a vector $|\psi\rangle$ in a suitably constructed Hilbert space. The probability density is

given by a scalar product,

$$W(x) = \langle x | \Psi \rangle, \quad (2.77)$$

so that

$$|\psi\rangle = \int dx |x\rangle \langle x | \psi \rangle = \int dx W(x) |x\rangle. \quad (2.78)$$

In this formalism, the basis states $|x\rangle$ correspond to δ -functions,

$$\langle x | y \rangle = \delta(x - y). \quad (2.79)$$

Normalization requires

$$\langle 1 | \psi \rangle = \int W(x) dx = 1, \quad (2.80)$$

where we have introduced the normalization state $|1\rangle$, which corresponds to a uniform distribution $Q(x) = 1$, *i.e.*

$$|1\rangle = \int dx |x\rangle.$$

The dynamics must preserve the projection of Eq. (2.80). Furthermore, in this operator formalism, the expectation of an operator \mathcal{A} is given by the product

$$\langle \mathcal{A} \rangle = \langle 1 | \mathcal{A} | \psi \rangle.$$

We define operators \hat{x}, \hat{p} , such that in x -representation

$$\langle x | \hat{x} | \psi \rangle = x W(x), \quad (2.81)$$

$$\langle x | \hat{p} | \psi \rangle = -\frac{\partial}{\partial x} W(x). \quad (2.82)$$

These operators obey the commutation relations

$$[\hat{x}, \hat{p}] = 1.$$

This identification afforded to us by Eq.(2.81) and (2.82) allows one to easily write the operator form of \mathcal{L}_{FP} . For continuous random variables, one often uses the characteristic

function

$$\widetilde{W}(k) = \int dx e^{-ikx} W(x),$$

(the Fourier transform of W). In the k -representation

$$\langle k | \hat{x} | \psi \rangle = \frac{\partial}{\partial k} \widetilde{W}(k), \quad (2.83)$$

$$\langle k | \hat{p} | \psi \rangle = k \widetilde{W}(k). \quad (2.84)$$

This gives us a quick method to write the partial differential equation for the characteristic function. Expectations of \hat{x} operators are the moments: the j -th moment of the distribution, μ_j is given by the expectation

$$\mu_j = \langle 1 | (\hat{x})^j | \psi \rangle.$$

An additional result that is useful for evaluating expectations is the identity

$$\langle 1 | \hat{p} = 0.$$

Using our Hilbert space, the Fokker-Planck equation reads

$$\frac{d}{dt} |\psi\rangle = \mathcal{L}_{\text{FP}} |\psi\rangle, \quad (2.85)$$

and has the formal solution

$$|\psi\rangle_t = \exp(\mathcal{L}_{\text{FP}} t) |\psi\rangle_0. \quad (2.86)$$

The path integral

At this point, we could derive a path-integral from the time-evolution operator. This derivation follows closely that of the occupation number formalism, so I only give the result. One adds the source terms $\bar{h}(t)x(t)$ and $h(t)p(t)$ to the resulting action and arrive at the functional integral

$$Z[\bar{h}, h] = \int [dp] [dx] \exp(-S[p, x]), \quad (2.87)$$

with action

$$S[p, x] = \int dt \{p(t) \partial_t x(t) - L[p(t), x(t)] - \bar{h}(t) x(t) - h(t) p(t)\}, \quad (2.88)$$

where we have obtained $L[p(t), x(t)]$ by substituting $x(t)$ for \hat{x} , and $p(t)$ for \hat{p} in \mathcal{L}_{FP} .

An example.

We now illustrate the formalism with an extended example. We start with a continuous random walk. It is described by the Fokker-Planck equation

$$\frac{\partial W(x)}{\partial t} = \frac{\gamma}{2} \frac{\partial^2}{\partial x^2} W(x),$$

and therefore

$$\mathcal{L}_{\text{FP}} = \frac{\gamma}{2} \hat{p}^2.$$

The action, excluding the source terms, of this process is simply

$$S[p, x] = \int dt \left[p(t) \partial_t x(t) - \frac{\gamma}{2} p^2(t) \right]. \quad (2.89)$$

We can carry out the path integral over $p(t)$ since at each point in time it is a Gaussian.

The path-integral becomes

$$Z = \int [dx] \exp(-S'[x]),$$

with the new action

$$S'[x] = \frac{(dx(t)/dt)^2}{2\gamma},$$

which one recognizes as the well-known “Wiener functional”. It is typically not convenient to perform the integration over $[dp]$. The only reason here was to make contact with a known result.

We can use Heisenberg equations to look at the first two moments.

$$\frac{d}{dt} \langle x \rangle = \langle [\hat{x}, \mathcal{L}_{\text{FP}}] \rangle$$

$$\begin{aligned}
&= \gamma \langle 1 | \hat{p} | \psi \rangle \\
&= 0,
\end{aligned}$$

and

$$\begin{aligned}
\frac{d}{dt} \langle x^2 \rangle &= \langle [\hat{x}^2, \mathcal{L}_{\text{FP}}] \rangle \\
&= \gamma \langle 1 | 1 + 2\hat{p}\hat{x} | \psi \rangle \\
&= \gamma.
\end{aligned}$$

As expected, the variance grows linearly with time.

For the next example, we add a damping term.

$$\frac{\partial W(x)}{\partial t} = \lambda \frac{\partial}{\partial x} x W(x) + \frac{\gamma}{2} \frac{\partial^2}{\partial x^2} W(x) \quad (2.90)$$

This Fokker-Planck equation describes diffusion in a parabolic potential. Alternatively, if one thinks of x as representing the velocity, Eq. (2.90) is a simple model of Brownian motion. We know from Eq. (2.12) that the steady-state distribution is Gaussian, *i.e.*

$$W_S(x) = \sqrt{\frac{\lambda}{2\pi\gamma}} \exp\left(-\frac{\lambda x^2}{\gamma}\right). \quad (2.91)$$

In terms of the operators,

$$\mathcal{L}_{\text{FP}} = -\lambda \hat{x} \hat{p} + \frac{\gamma}{2} \hat{p}^2.$$

Using the Heisenberg equations, we find equations for the first two moments.

$$\begin{aligned}
\frac{d\langle x \rangle}{dt} &= -\lambda \langle x \rangle, \\
\frac{d\langle x^2 \rangle}{dt} &= -2\lambda \langle x^2 \rangle + \gamma.
\end{aligned} \quad (2.92)$$

Solving for the steady state,

$$\begin{aligned}
\mu_1 &= \langle x \rangle = 0, \\
\mu_2 &= \langle x^2 \rangle = \frac{\gamma}{2\lambda}.
\end{aligned} \quad (2.93)$$

Equations for higher moments can also be found in a similar manner,

$$\frac{d\mu_j}{dt} = -j\lambda\mu_j + \gamma\frac{j(j-1)}{2}\mu_{j-2}. \quad (2.94)$$

Solving these for the steady-state, we have

$$\mu_j = \frac{\gamma}{2\lambda} (j-1) \mu_{j-2}. \quad (2.95)$$

Therefore, the odd moments vanish. If we use units such that $\frac{\gamma}{2\lambda} = 1$, we find $\mu_2 = 1$, $\mu_4 = 3$, $\mu_6 = 5 \cdot 3$, ..., $\mu_{2n} = (2n-1)!!$, which are precisely the moments of a Gaussian distribution.

The two-point correlation function (or auto-correlation), and two-point propagator can also be found for this model using Heisenberg equations, we find

$$\frac{d}{dt}\hat{p} = \lambda\hat{p}, \quad (2.96)$$

and

$$\frac{d}{dt}\hat{x} = -\lambda\hat{x} + \gamma\hat{p}. \quad (2.97)$$

These equations do not couple \hat{x} and \hat{p} to higher products so we can write the solution for the operators,

$$\hat{x}(t) = \hat{x}(0)e^{-\lambda t} - \frac{\gamma}{\lambda}\hat{p}(0)\left[\frac{e^{-\lambda t}}{2} - \sinh(\lambda t)\right], \quad (2.98)$$

$$\hat{p}(t) = \hat{p}(0)e^{\lambda t}. \quad (2.99)$$

Using the commutation relations, the fact that $\langle 1|\hat{p}(0) = 0$, and taking expectations results in

$$G_{\phi\bar{\phi}}(t) = \langle 1|\hat{x}(t)\hat{p}(0)|\psi\rangle = e^{-\lambda t}, \quad (2.100)$$

and

$$G_{\phi\phi}(t) = \langle 1|\hat{x}(t)\hat{x}(0)|\psi\rangle = \mu_2(0)e^{-\lambda t}. \quad (2.101)$$

In the steady-state, Eq. (2.101) becomes

$$G_{\phi\phi}(t) = \frac{\gamma}{2\lambda} e^{-\lambda t}.$$

These results can also be found using the functional integral since it can be transformed to the standard Gaussian form.

To conclude this section, we note that this formalism is easily extended to continuum models. For example, a spatial version of the example of Eq. (2.90) would be described by the continuum Fokker-Planck equation

$$\begin{aligned} \frac{\partial W[\phi_{\mathbf{r}}(t)]}{\partial t} &= \frac{\delta}{\delta\phi_{\mathbf{r}}(t)} \left\{ -D\nabla^2\phi_{\mathbf{r}}(t) + \lambda\phi_{\mathbf{r}}(t) \right\} W[\phi_{\mathbf{r}}(t)] \\ &+ \frac{\Gamma}{2} \frac{\delta^2}{\delta\phi_{\mathbf{r}}(t)\delta\phi_{\mathbf{r}}(t)} W[\phi_{\mathbf{r}}(t)]. \end{aligned} \quad (2.102)$$

Identifying $\frac{\delta}{\delta\phi_{\mathbf{r}}(t)}$ with $\bar{\phi}_{\mathbf{r}}$ produces the action

$$S[\bar{\phi}_{\mathbf{r}}, \phi_{\mathbf{r}}] = \int d^d\mathbf{r}dt \left[\bar{\phi}_{\mathbf{r}}\partial_t\phi_{\mathbf{r}} + D\bar{\phi}_{\mathbf{r}}\nabla^2\phi_{\mathbf{r}} - \lambda\bar{\phi}_{\mathbf{r}}\phi_{\mathbf{r}} + \frac{\Gamma}{2}\bar{\phi}_{\mathbf{r}}^2 \right]. \quad (2.103)$$

We find that for any given stochastic process, the operator formalism for the Fokker-Planck equation produces the same functional integral that we could derive using the response functional method. One advantage of the operator formulation is that one is not restricted to only using path-integral techniques. We can alternatively work with operator equations including time-ordered expansions. In addition, as was mentioned previously, we must specify the noise term of the Langevin equation using some additional information.

2.4.4 System size expansions, again!

In section 2.2.4, I introduced two system size expansions for birth-death master equations. These expansions provide systematic procedures for passing from a microscopic picture

in terms of discrete random variables to a macroscopic one employing continuous random variables. In other words, they transform stochastic variables representing absolute numbers of things (extensive variables) into stochastic variables that are densities (intensive variables). The key ingredient for both of these expansions is an appropriate scaling of the microscopic variables by the system size and proper application of the rules for transformation of random variables. The choice of system size scaling, $\Omega^{-1/2}$ vs. Ω^{-1} , is also what distinguishes the two expansions. Here we point out that both expansions can be carried out using the occupation number formalism for birth-death processes. This is a very practical result since one is usually concerned with the behaviour in the limit of large system size. We will also find an important connection between the occupation number formalism and the operator form of the Fokker-Planck equation that parallels the classic results of Horsthemke, *et. al.* [22][23], and Kurtz [24][25].

The system size expansion Horsthemke, *et. al.* produces a non-linear Fokker-Planck equation. The expansion can be carried out using creation and annihilation operators, for example a and a^\dagger , by the canonical transformation

$$\begin{aligned} a &\rightarrow \Omega c, \\ \tilde{a} &= a^\dagger - 1 \rightarrow \tilde{c}/\Omega. \end{aligned} \tag{2.104}$$

In the limit of large Ω , the c operator now becomes what we denoted as the \hat{x} operator of the Fokker-Planck equation and the \tilde{c} operator becomes the \hat{p} operator. The need for the shift to $\tilde{a} = a^\dagger - 1$ can be seen from action of a^\dagger and \hat{x} on their respective normalization states. Recall that for the occupation number representation, expectations are taken with respect to the normalization state

$$|S\rangle = e^{a^\dagger} |0\rangle, \tag{2.105}$$

which has the property

$$\langle S| a^\dagger = \langle S|. \quad (2.106)$$

On the other hand, the normalization state $|1\rangle$ for the Fokker-Planck Hilbert space represents a uniform distribution

$$|1\rangle = \int dx |x\rangle, \quad (2.107)$$

and therefore has the property

$$\langle 1| \hat{p} = 0. \quad (2.108)$$

This suggests the identification above since

$$\langle S| \tilde{a} = 0.$$

This method can be shown to duplicate the results obtained with the classical technique although a more rigorous argument showing how one passes from a finite vector space to the infinite dimensional Hilbert space is desirable. We should mention that Grassberger and Scheunert [43] have discussed many of the implications of using shifted operators like \tilde{a} and the choice of the scalar product of the vector space.

Van Kampen's expansion of the master equation [19][20] can also be implemented with operators. This system size expansion of Van Kampen produces a linear Fokker-Planck equation, and therefore Gaussian fluctuations around the mean-field solution. In terms of operators, it is carried out by the canonical transformation,

$$a \rightarrow \sqrt{\Omega}c \quad (2.109)$$

$$\tilde{a} = a^\dagger - 1 \rightarrow \tilde{c}/\sqrt{\Omega}.$$

The only complication is that when one makes the identification $c \rightarrow \hat{x}$, and $\tilde{c} \rightarrow \hat{p}$, one must also shift \hat{x} by an amount equal to $\langle x \rangle$, *i.e.* the mean-field solution in order to get the Fokker-Planck operator describing the fluctuations around this value. This is not a

problem, however, for in the limit of large Ω , one easily solves the Heisenberg equation of motion for $\langle x \rangle$, as we did in the example of 2.4.3. Although this expansion can be quite useful and informative, we will not discuss it further. As we pointed out in section 2.2.4, Van Kampens system size expansion breaks down near the instabilities or critical points of the stochastic process. This is particularly disastrous for models with spatial degrees of freedom.

An example: density dependent mortality.

Here I illustrate the system size expansion of Eq. (2.104) by extending the branching process discussed in our introduction to the occupation number formalism (section 2.4.2) to include density-dependent mortality. Density-dependent mortality is added to the branching model *via* the process $A + A \rightarrow A$, with rate w/Ω . The choice of Ω in the rate of this process insures that the death rate due to other individuals scales with the density and not (unreasonable) the total number of particles. The Liouville operator has terms for birth, or branching at rate γ , a finite lifetime λ , and the density dependent term with rate w/Ω ,

$$\begin{aligned} \mathcal{L} = & \gamma (a^\dagger a^\dagger a - a^\dagger a) + \lambda (a - a^\dagger a) \\ & + w (a^\dagger a^2 - a^\dagger a^\dagger a^2) / \Omega. \end{aligned} \quad (2.110)$$

First, we perform the shift to $\tilde{a} = a^\dagger - 1$. This produces

$$\begin{aligned} \mathcal{L} = & \gamma \tilde{a}^2 a + (\gamma - \lambda) \tilde{a} a \\ & - w (\tilde{a} a^2 + \tilde{a}^2 a^2) / \Omega. \end{aligned} \quad (2.111)$$

Now we implement the scaling transformation

$$\begin{aligned} a & \rightarrow \Omega c, \\ \tilde{a} & = a^\dagger - 1 \rightarrow \tilde{c} / \Omega, \end{aligned} \quad (2.112)$$

to find

$$\begin{aligned} \mathcal{L} &= (\gamma - \lambda) \tilde{c}c - w\tilde{c}c^2 \\ &+ (\gamma\tilde{c}^2c - w\tilde{c}^2c^2) / \Omega. \end{aligned} \tag{2.113}$$

The non-linear Fokker-Planck equation for this stochastic process is therefore

$$\begin{aligned} \frac{\partial W(x)}{\partial t} &= -\frac{\partial}{\partial x} [(\gamma - \lambda)x - wx^2] W(x) \\ &+ \frac{1}{\Omega} \frac{\partial^2}{\partial x^2} (\gamma x - wx^2) W(x), \end{aligned}$$

and is an equally valid description of the stochastic process originally formulated *via* the master equation. In section 3.2.1 we implement this type of system size expansion for a population genetics model. The reader should compare the ease of the approach using operators with that of working directly with the master equation.

This section completes our survey of stochastic processes and was included to highlight the relationship between the occupation number representation of birth-death processes and the operator form of the Fokker-Planck equation. I use these ideas again when we look at the equivalent representations of the Stepping-Stone model in chapter 4.

Chapter 3

Spontaneous Symmetry Breaking in the Wright-Fisher/Moran Model

R. A. Fisher [7] was the first to mathematically analyze the problem of genetic drift using a partial differential equation (p.d.e.) known as the Fokker-Planck, or forward Kolmogorov equation. In this chapter, I will primarily use a non-linear Fokker-Planck equation [16] to describe the time-evolution of the probability distribution for a variable $x \in [0, 1]$. The variable x represents the frequency of a given gene or allele in a fixed-size population. Fisher's original work employed an equation for $y = \cos x$, which transforms the non-linear FP equation to a linear one. (This is discussed briefly in section 2.5). In either approach, the Fokker-Planck equation is a continuous-time "diffusion" equation that describes the large system size, long-time limit of what has become known as Wright-Fisher model. This model is of fundamental importance in theoretical population genetics and provides a basis for the quantitative tools of modern phylogenetic analysis. Even though this model has been quite well studied over the years, understanding its behavior is an important prerequisite for further development of the Stepping-Stone model and stochastic population dynamics in general. Methods other than the FP equation can be used to study the Wright-Fisher model. However, the FP equation would be most familiar to physicists (it resembles the Schrödinger equation) and more closely related to developments in modern statistical mechanics.

3.1 The Wright-Fisher and Moran Models.

The Wright-Fisher model is discrete, both in time and state space. It consists of a finite population of N haploid individuals which are either of type ‘ A ’ or ‘ a ’¹. Using genetics terminology, this would be called a “one locus, two allele” model. The population evolves under non-overlapping generations with sequential stages of producing an essentially infinite number of gametes (replication), followed by sampling of these gametes to reform a population of size N . Thus the total population size is conserved. Genetic drift, or fluctuation in the allele frequencies, results from this random sampling which is easily shown to be of binomial form. A schematic of the Wright-Fisher model is shown in figure 3.1. In the limit of large population size, the state of the system is described by the gene frequency $x = \frac{n_a}{N}$ (where n_a is the number of “ a ” individuals and we equally well could have chosen A as the allele of interest. Both Fisher [48], and Wright [6] studied the effects of this type of binomial sampling on gene frequencies for finite populations within the context of the discrete model. Subsequently Wright [49] introduced the diffusion approximation which lead to a non-linear FP equation that has been well-studied in the theoretical population genetics literature.

A simpler alternative to the Wright-Fisher model was presented by Moran [50], and consists of overlapping generations. A schematic of Moran’s model is presented in figure 3.2. In this model, two haploid individuals are picked at random. One dies without reproducing while the other is replicated. The basic effect is the same as the binomial sampling in the Wright-Fisher model. That is, not all individuals contribute to the next generation. Moran’s model can be formulated in term of a birth-death master equation

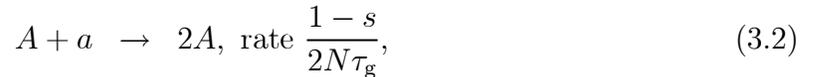
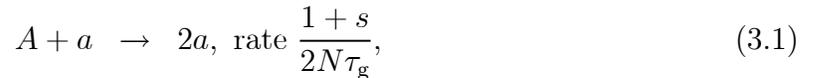
¹The convention in population genetics is to specify the population size as $2N$. The N refers to the number of *diploid* individuals, each formed by the union of 2 *haploid* gametes. Under random mating and after one generation, the diploid population will be in Hardy-Weinberg equilibrium, whereby the diploid population behaves as a *haploid* population with twice the size[47]. To simplify the calculations, we use N for the number of haploid individuals. One can convert the results to diploid populations under random mating by substituting $2N$ for N .

of the kind discussed in chapter 2. In the next section, start from this description of Moran’s model and derive a non-linear Fokker-Planck equation for Moran’s model which will be identical to that of the Wright-Fisher model.

3.2 The Fokker-Planck Equation for Moran’s Model.

In section 2.2.4, I presented an asymptotic expansion of the master equation that produces a non-linear Fokker-Planck equation. I use this expansion here derive a non-linear diffusion equation for Moran’s model. First I carry out the expansion with the classical technique. Although I use the system size expansion of Horsthemke and Brenig [22][23], the notation and style is closer to that of Van Kampen [19][20].

The dynamics of Moran’s model consist of two types of random events that occur independently in time. This means the stochastic process is Markovian and the time intervals between similar events are exponentially distributed. The first event type corresponds to random sampling of gametes with a selective advantage parameterized by s . Two individuals are picked at random. One produces an offspring identical to itself while the other is discarded without reproducing. This is illustrated by the ‘replacement’ processes



where N is the total number of haploids, $N = n_a + n_A$. The processes $2A \rightarrow 2A$ and $2a \rightarrow 2a$ do not change the configuration and hence do not enter the stochastic description. The N -dependence of the rates in Eqs. (3.1) and (3.2) is chosen so that the replacement rate per individual becomes independent of N . Hence the concept of a “generation time”, τ_g is meaningful. The second event type describes mutation. These

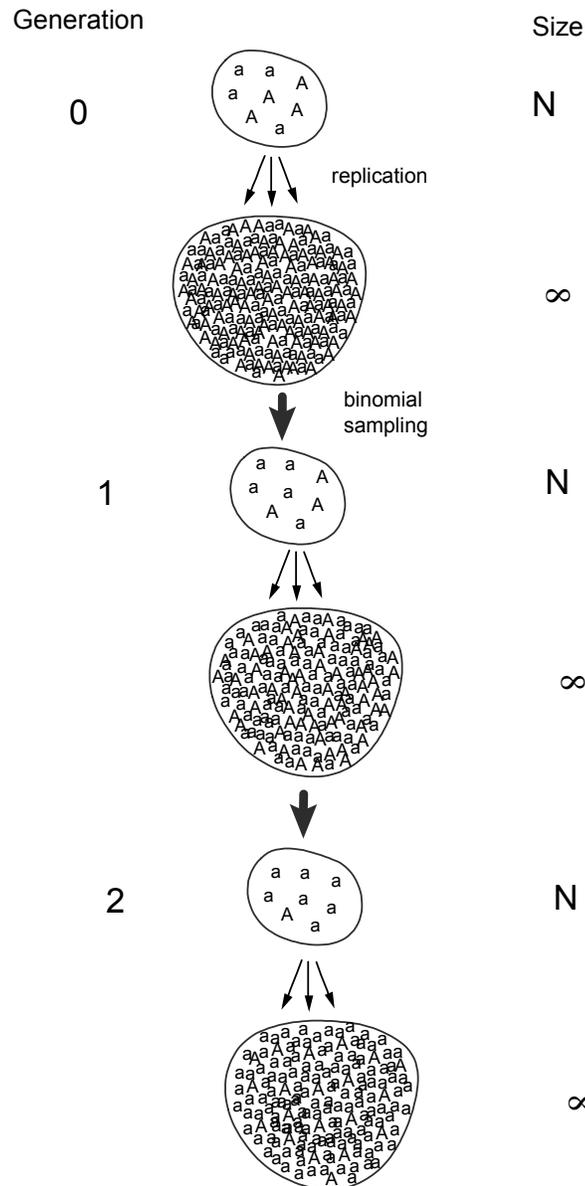


Figure 3.1: Schematic diagram of the Wright-Fisher model. Each generation of size N replicates into an infinite pool of gametes from which N are sampled to form the next generation.

events are described by the processes

$$a \rightarrow A,$$

$$A \rightarrow a.$$

We assume mutations are reversible and both events occur at rate μ .

In the expansion below, N will be the system size parameter we called Ω in chapter 2. In general, Ω is not the total number of particles, but the volume. For this model the total number is a conserved quantity and therefore we can take $\Omega = N$.

3.2.1 Expansion of the master equation.

With the stochastic events described above, the master equation for this process is

$$\begin{aligned} \frac{\partial P(n_a, n_A, t)}{\partial t} = & (N\tau_g)^{-1} \left\{ \frac{1+s}{2} (\mathcal{E}_A^{-1} \mathcal{E}_a - 1) + \frac{1-s}{2} (\mathcal{E}_a^{-1} \mathcal{E}_A - 1) \right\} n_a n_A P(n_a, n_A, t) \\ & + \mu \left\{ (\mathcal{E}_A^{-1} \mathcal{E}_a - 1) n_a + \mu (\mathcal{E}_a^{-1} \mathcal{E}_A - 1) n_A \right\} P(n_a, n_A, t) \end{aligned} \quad (3.4)$$

where \mathcal{E}_a is a “shift” operator defined by

$$\mathcal{E}_a f(n_a, n_A) = f(n_a + 1, n_A),$$

and

$$\mathcal{E}_a^{-1} f(n_a, n_A) = f(n_a - 1, n_A),$$

and likewise for the \mathcal{E}_A 's. Now we transform to intensive quantities

$$x_a = \frac{n_a}{N}, x_A = \frac{n_A}{N}.$$

The shift operators under this transformation become

$$\mathcal{E}_a = \sum_j \frac{1}{j!} N^{-j} \frac{\partial^j}{\partial x_a^j}, \quad \mathcal{E}_a^{-1} = \sum_j \frac{(-1)^j}{j!} \frac{\partial^j}{\partial x_a^j} N^{-j}. \quad (3.5)$$

Using the rules for transformation of random variables, we find the probability distribution

$$W(x_a, x_A; t) = P(n_a, n_A, t) / \left| \frac{\partial(x_a, x_A)}{\partial(n_a, n_A)} \right|, \quad (3.6)$$

where $\frac{\partial(x_a, x_A)}{\partial(n_a, n_A)}$ is the Jacobian of the transformation

$$\begin{aligned} \frac{\partial(x_a, x_A)}{\partial(n_a, n_A)} &= \begin{vmatrix} \frac{\partial x_a}{\partial n_a} & \frac{\partial x_a}{\partial n_A} \\ \frac{\partial x_A}{\partial n_a} & \frac{\partial x_A}{\partial n_A} \end{vmatrix} \\ &= N^{-2}. \end{aligned}$$

If we now write the equation $W(x_a, x_A; t)$ and use the shift operators for the expansion, we find no terms of order N^0 . The highest order terms are $\mathcal{O}(N^{-1})$, keeping only these gives us the equation

$$\begin{aligned} \frac{\partial W(x_a, x_A, t)}{\partial x} &= \frac{\mu}{2} \left(\frac{\partial}{\partial x_a} - \frac{\partial}{\partial x_A} \right) x_A W(x_a, x_A, t) + \frac{\mu}{2} \left(\frac{\partial}{\partial x_A} - \frac{\partial}{\partial x_a} \right) x_a W(x_a, x_A, t) \\ &\quad \frac{1}{2N\tau_g} \left\{ \frac{\partial^2}{\partial x_a^2} + sN \frac{\partial}{\partial x_a} - sN \frac{\partial}{\partial x_A} + \frac{\partial^2}{\partial x_A^2} - 2 \frac{\partial^2}{\partial x_A \partial x_a} \right\} x_a x_A W(x_a, x_A, t) \end{aligned}$$

but since

$$x_a + x_A = 1,$$

$$\partial x_a = -\partial x_A,$$

we can define $x \equiv x_a$, so that $\partial/\partial x_a = -\partial/\partial x_A = \partial/\partial x$, and the equation becomes

$$\begin{aligned} \frac{\partial W(x, t)}{\partial x} &= \mu \frac{\partial}{\partial x} (2x - 1) W(x, t) + s\tau_g^{-1} \frac{\partial}{\partial x} x(1-x) W(x, t) \\ &\quad + \frac{1}{2N\tau_g} \frac{\partial^2}{\partial x^2} x(1-x) W(x, t). \end{aligned} \quad (3.7)$$

which is also the diffusion equation for the Wright-Fisher model [47].

3.2.2 Operator forms of Moran's model.

Using the operator formalism discussed in section 4.2 of chapter 2, we can write the master equation for Moran's model in terms of creation and annihilation operators. This approach will be generalized in the next chapter as a reformulation of the Stepping-Stone model. We use a^\dagger and a to describe creation and annihilation of the 'a' individuals, and b^\dagger and b for the 'A' individuals. The Liouville operator reads

$$\begin{aligned} \mathcal{L} = \frac{1}{2N\tau_g} & \left[(1-s) (b^\dagger)^2 ab + (1+s) (a^\dagger)^2 ab - 2a^\dagger b^\dagger ab \right] \\ & + \mu (b^\dagger a - a^\dagger a) + \mu (a^\dagger b - b^\dagger b). \end{aligned} \quad (3.8)$$

Due to the conservation law, one might be tempted to replace the b operators with numbers,

$$b \rightarrow N - a, \quad (3.9)$$

$$b^\dagger \rightarrow 2 - a^\dagger.$$

and write a Liouville operator in terms of the a 's only. There is a problem with this, however. It does *not* represent a canonical transformation. The transformation of Eq. (3.9) is equivalent to operator transformations

$$d = a + b, \quad (3.10)$$

$$d^\dagger = a^\dagger + b^\dagger,$$

followed by the replacement $d \rightarrow N$, and $d^\dagger \rightarrow 2$. The first problem is that

$$[d, d^\dagger] = 2, \quad (3.11)$$

which is easily remedied. A much more serious problem arises from the commutator

$$[d, a^\dagger] = 1 \neq 0. \quad (3.12)$$

The correct way to handle the problem is illustrated in two successive transformations. We could combine the two, but i) it is more instructive to proceed in two steps, ii) we use the separate transformations in the next chapter when we generalize this model to spatially-distributed populations, and iii) the first step is carried out in a manner that facilitates the operator form of the system size expansion discussed in chapter 2.

First, we make a canonical transformation to sum and difference operators,

$$\begin{aligned} c &= (a - b) / N, & d &= a + b, \\ c^\dagger &= \frac{N}{2} (a^\dagger - b^\dagger), & d^\dagger &= \frac{1}{2} (a^\dagger + b^\dagger). \end{aligned} \quad (3.13)$$

The resulting Liouville operator is

$$\mathcal{L} = -2\mu c^\dagger c + \frac{1}{2N\tau_g} \left[-Nsd^\dagger c^\dagger c^2 + N^{-2} (c^\dagger)^2 d^2 + N^{-1} sc^\dagger d^\dagger d^2 - c^{\dagger 2} c^2 \right] \quad (3.14)$$

Now we can make the substitution $d \rightarrow N$, and $d^\dagger \rightarrow 1$. This leaves us with

$$\mathcal{L} = -2\mu c^\dagger c + \frac{1}{2N\tau_g} \left[-Nsc^\dagger c^2 + (c^\dagger)^2 + Nsc^\dagger - c^{\dagger 2} c^2 \right] \quad (3.15)$$

At this point, we could take N large and identify c and c^\dagger with Fokker-Planck operators and write a non-linear Fokker-Planck equation for the variable $z = 2x - 1$. We will skip this for now and make the (canonical) transformation to a and \tilde{a} ,

$$\begin{aligned} a &= \frac{c + 1}{2} \\ \tilde{a} &= 2c^\dagger. \end{aligned}$$

(We are re-using the notation; a here is actually N times the original a). The Liouville operator now reads

$$\mathcal{L} = -\mu a^\dagger (2a - 1) + s\tau_g^{-1} \tilde{a} a (1 - a) - \frac{1}{2N\tau_g} \tilde{a}^2 a (1 - a). \quad (3.16)$$

The system size expansion is obtained by taking a and \tilde{a} to represent Fokker-Planck operators (see section 2.4.3) rather than (scaled) occupation number operators, *i.e.*

$$\begin{aligned} a &\rightarrow \hat{x}, \\ \tilde{a} &\rightarrow \hat{p}. \end{aligned}$$

In what we have called the x -representation (see section 2.4.3), the operator \hat{x} is multiplication by x and the operator $-\hat{p}$ is partial differentiation with respect to x . In this manner we arrive at the same Fokker-Planck equation derived by the classical method and presented in Eq. (3.7).

3.3 Heterogeneous diffusion in the Wright-Fisher and Moran models.

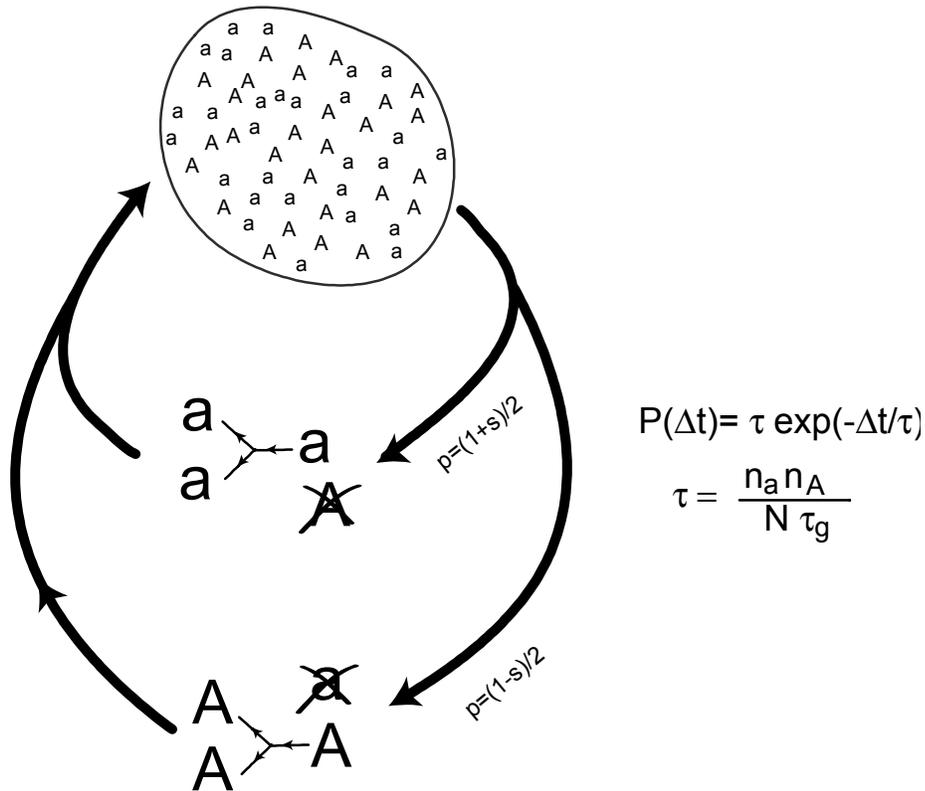
We now review the behaviour of the Wright-Fisher or Moran model as described by the associated Fokker-Planck equation. First, we consider the model in the absence of selective advantage or mutation. The Fokker-Planck equation for the allele probability density $W(x, t)$, in the limit of pure genetic drift is

$$\frac{\partial W(x, t)}{\partial t} = \frac{1}{2N} \frac{\partial^2}{\partial x^2} x(1-x), \quad (3.17)$$

where time is in units of the generation time τ_g .

One notices the effect of random sampling of gametes is to produce a singular, heterogeneous diffusion term—it vanishes at the boundaries $x = 0$ and $x = 1$. This results in a dynamical buildup of probability at the absorbing states $x = 0$ and $x = 1$, and correspondingly, an asymptotic decay of the probability density everywhere else. This arises not only because of the absorbing states, it is also due to the state-dependent or heterogeneous diffusion (see section 2.2 of chapter 2) given by the diffusion term $B(x) = x(1-x)/2N$. This term generates an effective force that drives the system toward the boundaries. In other words, the behaviour for vanishing s and μ is more

a) "replacement" with selection



b) mutation



Figure 3.2: Stochastic Processes in the Moran Model. The arrows represent random events. a) Replacement. This event occurs at time intervals Δt given by the distribution $P(\Delta t) \propto \exp(-\Delta t/\tau)$ with τ defined in the figure. The upper path is taken with probability $p = (1 + s)/2$, while the lower path is taken with probability $p = (1 - s)/2$. b) Reversible mutation at rate μ .

complicated than a random walk with “cliffs” at the boundary. We shall see that one can associate a potential with this force that is symmetric about $x = 1/2$. We can make this symmetry manifest by transforming to the variable $z \equiv 2x - 1$, $z \in [-1, 1]$. In the preceding section I used an operator version of this transformation and will use it again in the next chapter when I discuss the Stepping-Stone model. In terms of this new variable, the Fokker-Planck operator becomes

$$\mathcal{L}_{\text{drift}} = \frac{1}{2N} \frac{\partial^2}{\partial z^2} (1 - z^2),$$

which is obviously symmetric with respect to $z \rightarrow -z$. When mutation and selection vanish, this symmetry is “spontaneously broken”, meaning that the asymptotic probability distribution is

$$W(z) = \delta(z - 1), \text{ or } W(z) = \delta(z + 1).$$

Reversible mutation $A \xleftrightarrow{\mu} a$ and selective advantage in the sampling of gametes, as discussed in the preceding section, give rise to additional “drift” terms in the Fokker-Planck equation. If μ is non-zero, the absorbing states vanish. (Recall from section 2.2.2, that both the drift term and the diffusion term must vanish at the same point to produce an absorbing state.) A non-zero coefficient of selection, s , breaks the $z \rightarrow -z$ symmetry and makes one absorbing state preferable.

The steady-state distribution and the dynamics of the expectation and variance of the (discrete) Wright-Fisher model had been known since the work of Fisher and Wright. However, the full solution of the FP equation (??), in terms of eigenvalues and eigenfunctions for various regimes of selection and mutation was presented in a series of papers by Kimura [51][5][13]. Kimura gave a rather complete analysis, not only of Eq. (??), but also of the conjugate or backward equation [52]. It will not be necessary for us to review these details. Instead, we will look at this problem from another point of view,

that of the “effective potential”. For the moment, however, we will make a temporary digression and completely ignore the fluctuations to see how this system behaves using a rather naive type of mean field theory.

3.4 Mean-field theory.

Mathematical modeling of population systems is dominated by the dynamical systems approach [53]. This requires the underlying assumption that in the limit of large population size, fluctuations are irrelevant, and average out in such a way that the behaviour of the system is given by dynamical equations for the densities of the various species comprising the population (or sub-populations in spatial models). This amounts to a kind of ‘dynamical mean-field’ ansatz, and is the foundation of the “Fisherian” school of neo-Darwinian mass selection. Later we will see precisely how this assumption fails but first it will be instructive to proceed with this assumption.

Using the Wright-Fisher or Moran model defined above, we write an equation for the expectation value of x , which we can simply and without confusion denote by ‘ $x(t)$ ’ in this section. The dynamical equation reads

$$\frac{dx(t)}{dt} = -\mu(2x(t) - 1) + sx(t)(1 - x(t)), \quad (3.18)$$

where time is measured in units of the generation time τ_g . Alternatively, we can define a potential (or Lyapunov function)

$$V(x) = \mu(x^2 - x) - s\left(\frac{x^2}{2} - \frac{x^3}{3}\right), \quad (3.19)$$

and equate the time derivative of x , to the negative derivative of V ,

$$\frac{dx}{dt} = -\frac{\partial V}{\partial x}.$$

The steady state value of $x(t)$ will be given by the minimum of V . In fact, Eq. (3.18) is quadratic so its explicit time-dependent solution is easily found. Figure 3.3 shows the potential for several values of μ/s . Note that the potential varies quadratically about its minimum. Setting the right side of Eq. (3.18) to zero, one finds the steady state value of z ,

$$\hat{x} = \frac{1}{2} \left(1 \pm \sqrt{1 + \frac{4\mu^2}{s^2}} \right) - \frac{\mu}{s},$$

where the positive (negative) sign is chosen if $\mu/s > 0$ ($\mu/s < 0$). The steady state arises from a balance between the forces of mutation and selection and is known as *mutation-selection balance* in the population genetics literature [47].

We would also like to know about the stability of system based upon Eq. (3.18). This is given by the sign of the partial derivative that results from acting on the right-hand side of Eq. (3.18) with $\partial/\partial x$, and evaluating the result for $x = \hat{x}$. Its value is $-\sqrt{(s^2 + 4\mu^2)}$. So we see that the system is stable as long as either s or μ is non-zero. In much of what follows, mean-field theory will prove to be inadequate, yet this result will still hold, at least qualitatively in the thermodynamic limit, even in the presence of fluctuations—both in the Wright-Fisher model, as well as the Stepping-Stone model. This is somewhat surprising since it is generally not the case in stochastic population models.

3.5 The Effective Potential.

To go beyond the simple mean field picture presented in the previous section, I make use of an “effective potential” for our Fokker-Planck equation. For multivariate systems, such a potential does not, in general, exist. However, for any single variable Fokker-Planck equation and in particular, the one of Eq. (3.7), one can construct an “effective potential” [16].

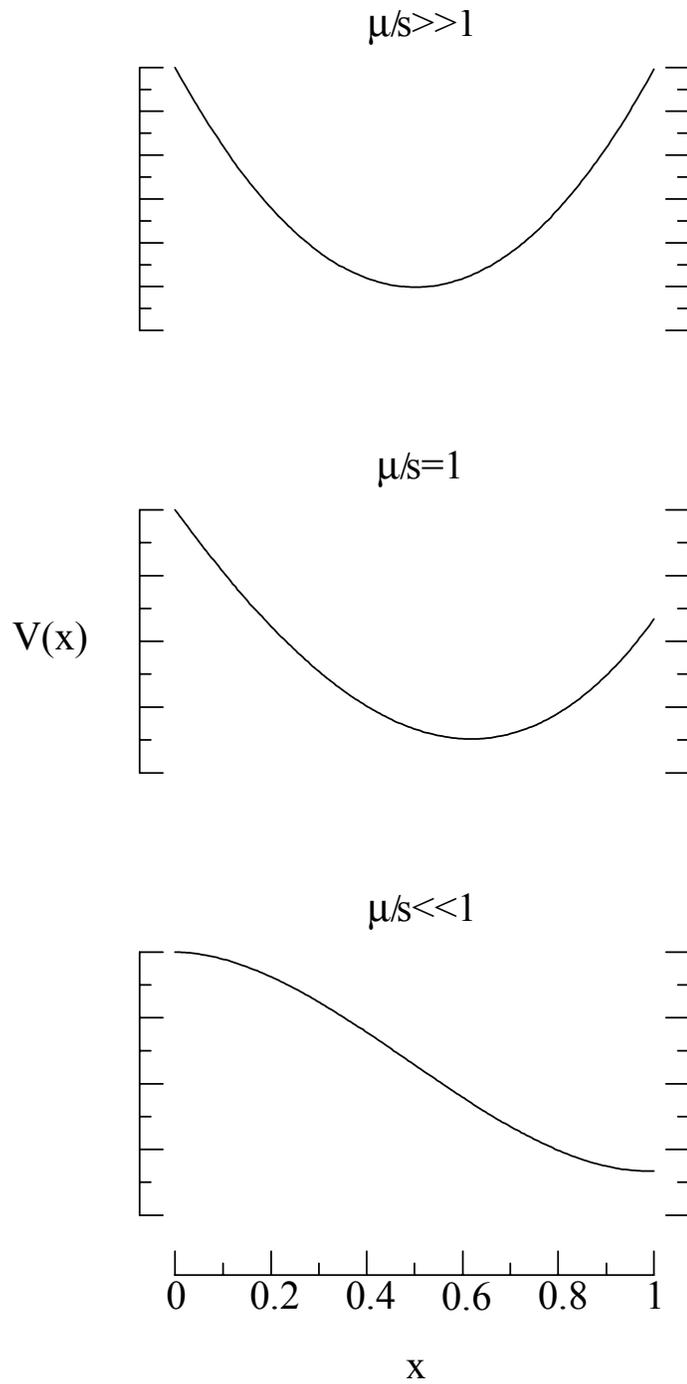


Figure 3.3: The mean-field potential $V(x)$ for several regimes of μ/s .

The steady-state distribution, $W_S(x)$ of the stochastic process defined by equations 3.7 is easily found since for any Fokker-Planck equation of this form, one has

$$W_S(x) = \frac{C}{B(x)} \exp \left[2 \int^x \frac{A(y)}{B(y)} dy \right],$$

where C is some normalization constant. This allows us to define a potential $U(x)$, such that

$$W_S(x) = C \exp(-U(x)).$$

If $B(x)$ were constant, which is the case for homogeneous diffusion, for example, in models where Gaussian noise is added to the mean field dynamics, this potential would be identical to that of the deterministic differential equation for the expectation value. In the present case, A and B are given by

$$\begin{aligned} A(x) &= -\mu(2x-1) + sx(1-x) \\ B(x) &= N^{-1}x(1-x), \end{aligned} \tag{3.20}$$

and where I have taken the units of time to be τ_g . $B(x)$ depends upon x , so the effective potential becomes

$$U(x) = (1 - 2N\mu) \ln(x(1-x)) - 2Nsx. \tag{3.21}$$

This gives us the steady-state distribution

$$W_S(x) = C e^{2Nsx} [x(1-x)]^{2N\mu-1}.$$

The normalization constant can be found from the intergral

$$\begin{aligned} C^{-1} &= \int_0^1 e^{2Nsx} [x(1-x)]^{2N\mu-1} dx \\ &= \frac{\Gamma^2(2N\mu)}{\Gamma(4N\mu)} F(2N\mu, 4N\mu; 2Ns), \end{aligned} \tag{3.22}$$

where $F(a, b; x)$ is the hypergeometric function. For $s = 0$, $W_S(x)$ is simply a symmetric Beta distribution, however, in the genetics literature, $W_S(x)$ is known as “Wright’s distribution”.

In contrast to the mean-field potential $V(x)$, the effective potential $U(x)$ in Eq. (3.21) is strongly anharmonic. More importantly, however, the effective potential changes sign when $2N\mu = 1$, and therefore is qualitatively different than its mean-field counterpart. Figure (3.4) shows a comparison of $U(x)$ and $V(x)$ for $\mu/s \simeq 1$. An important point to make here is that our new, effective potential tells us not only about the steady-state distribution, it also tells us something about the dynamics of the system. The point has been somewhat overlooked in the population genetics literature. In Figure (3.5) I show two simulations of the Wright-Fisher model for the two regimes, $2N\mu \gg 1$ and $2N\mu \ll 1$, with $s = 0$. These simulations were produced by integrating the stochastic differential equations (s.d.e.’s) associated with the Fokker-Planck Equation (3.7). For these simulations, the population size was $N = 1000$ and the mutation rate was adjusted so that $2N\mu = 20$ (top) and $2N\mu = 10^{-4}$ (bottom). The method of integration is discussed in section 1 of appendix B.

Let us consider what happens if we start with $s = 0$, and “tune” the mutation rate slowly to zero. First, as μ becomes less than $\frac{1}{2N}$, the effective potential flips. As μ is further reduced, we observe the behaviour shown in the bottom plot of Figure (3.5). In the limit $\mu \rightarrow 0$, the symmetry of $U(x)$ is broken as the system is trapped in one of the absorbing states. The manner in which this occurs, however, is quite different from what is common in equilibrium statistical mechanics. Figure (3.6) shows the effective potential for $2N\mu$ slightly larger than 1 (lower plot) and then slightly larger than 1 (upper plot). What one should notice, is that at this “critical point”, the order parameter (here the expectation value of x) changes discontinuously. But unlike first order transitions in equilibrium statistical mechanics, where the order parameter also

changes discontinuously, there is also critical behaviour near the transition.

In the next chapter, we will find that this unusual behavior has important ramifications for the scaling laws that relate critical exponents. In particular, the standard “hyperscaling” relation will have to be discarded and replaced with one that takes this effect into consideration.

3.6 The Broken Symmetry State: Fate of new mutants.

Let us suppose that the mutation rate is effectively zero and enough time has passed for the system to reach an absorbing state. We would like to know the response of the system to a small perturbation caused by a single new mutation to a completely new allele. One might imagine an individual of the population being hit by a cosmic ray, or subject to some other rare process that produces this new allele. Furthermore, the new allele could be selectively advantageous, disadvantageous, or just plain neutral. The ensuing dynamics of the system will be to produce a “cascade”, or “avalanche” of the new types in the population as a response to the new mutation. This is truly the Darwinian picture of the basic process of evolution, so it is quite desirable to know what can happen in this situation. One finds that even for an advantageous new allele, there is still a large probability that the resulting avalanche will die out. When the selective advantage is small, the fixation probability remains small regardless of the population size. Kimura has remarked that this fact has been never been truly appreciated by the mass selectionists, who imagine that every newly arisen beneficial allele becomes incorporated into the population [8].

Conversely, it is quite possible that the new allele will take over the population and reach fixation. For a neutral allele, the fixation probability *does* depend on the population size. It is arguable that the most important contribution by Kimura to

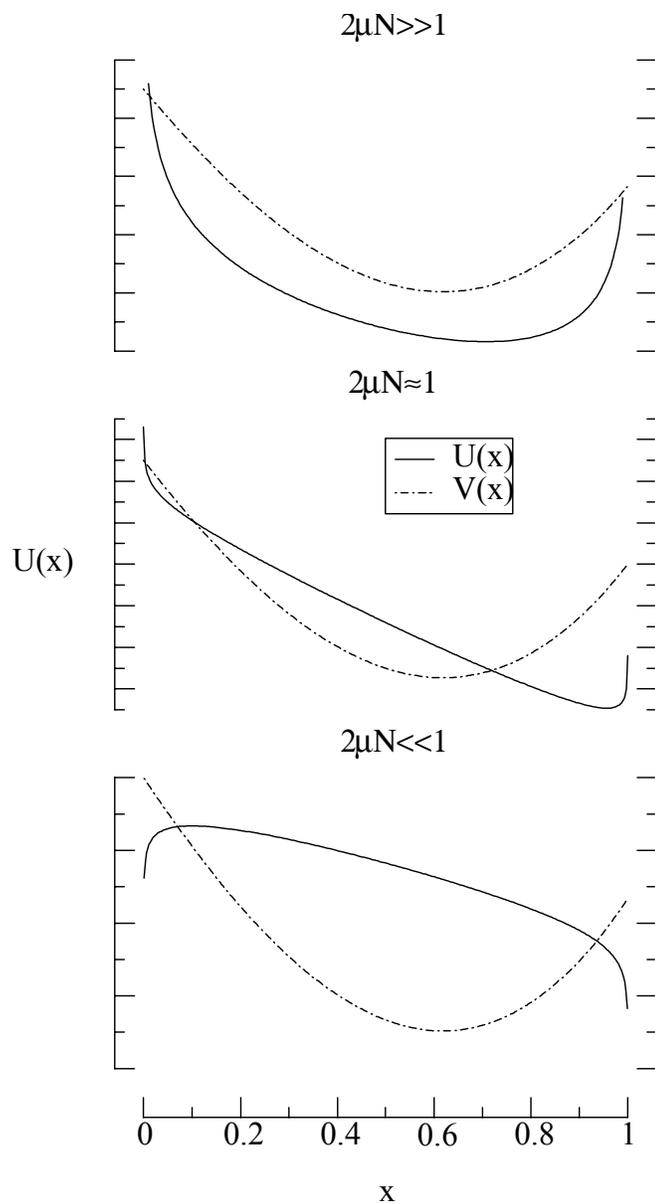


Figure 3.4: Comparison of the stochastic “effective potential” $U(x)$ with the mean-field potential $V(x)$ for different values of $2\mu N$. Top: $2\mu N \gg 1$, the minimum is shifted by an amount $\sim N^{-1}$. Center: $2\mu N \gtrsim 1$. Bottom: $2\mu N \ll 1$, the effective potential $U(x)$ has turned over.

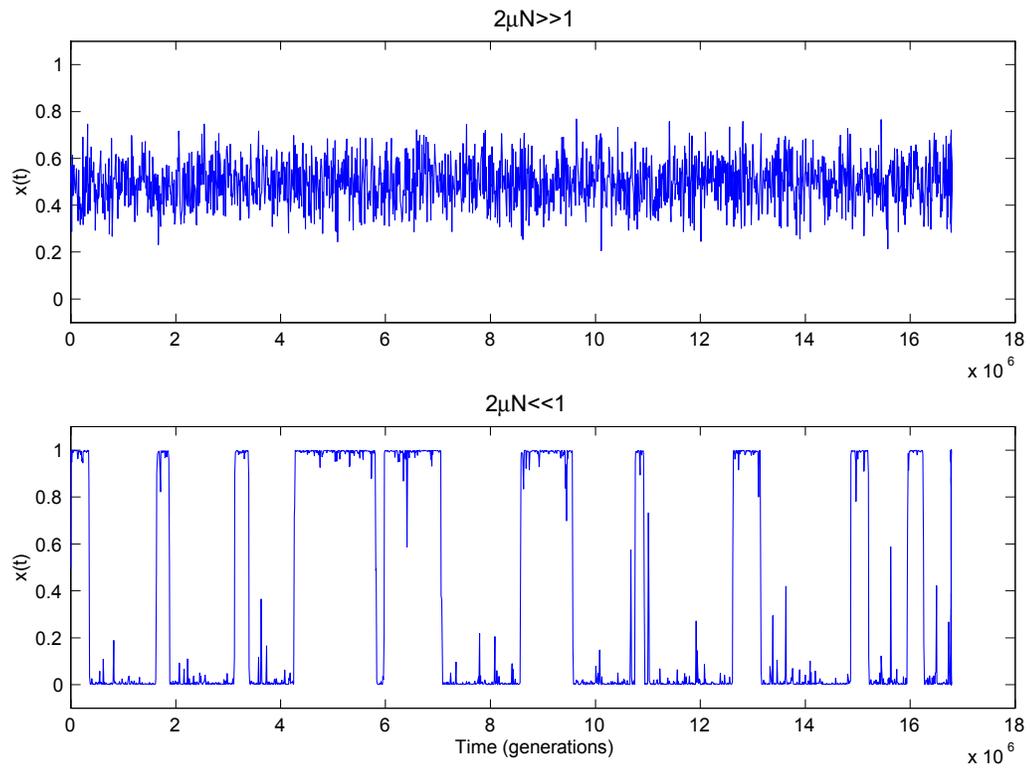


Figure 3.5: Simulations of the Wright-Fisher model using stochastic differential equations. For $2\mu N \gg 1$ (top), the fluctuations are approximately Gaussian around $x = 1/2$. For $2\mu N \ll 1$ (bottom), the system stays near the absorbing states.

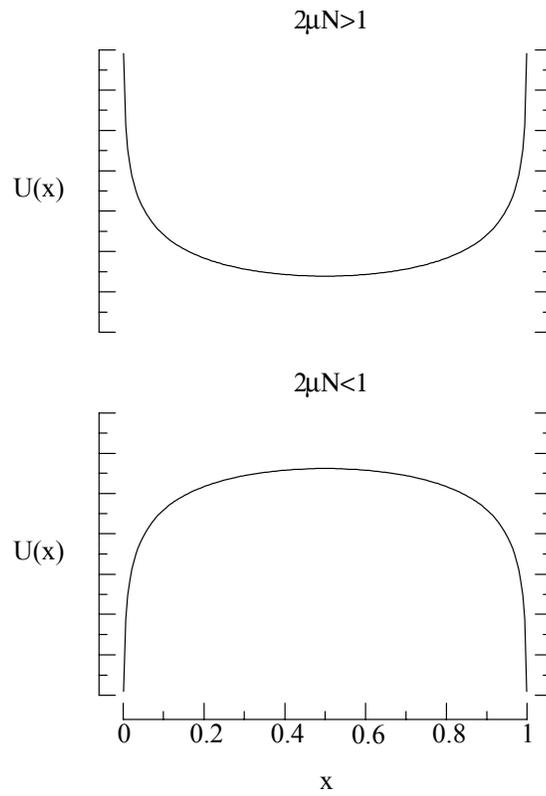


Figure 3.6: Discontinuous behaviour of the order parameter at the critical point of the Wright-Fisher model. The effective potential $U(x)$ changes sign at $2\mu N = 1$.

theoretical population genetics was to propose that most, if not all genetic variation observed in Nature is due to such neutral mutations [8][9]. This postulate, known simply as *The Neutral Theory* provides one of the few truly quantitative theories in Biology. It should not surprise us then, that as a predecessor to his Neutral Theory, Kimura has contributed the most to the simple problem of fixation probabilities of new mutant alleles discussed in this section [13][52].

The Fokker-Planck description we have been using so far is not well-suited for the type of calculation needed here, for what we want to calculate is effectively a first-passage type of problem. Using the Fokker-Planck equation to this end would require knowledge of the eigenvalues and eigenfunctions of the equation, which is indeed possible to derive [5]. Proceeding in this manner is much too cumbersome. The more general and also aesthetic solution is the use of the backward, or adjoint equation introduced in section 2.2.3.

$$\frac{\partial}{\partial t} u(x_0, t) = \mathcal{L}_{FP}^+(x_0) u(x_0, t). \quad (3.23)$$

This equation describes the fixation probability at time t , given that the system was initially at x_0 . This is precisely how Kimura proceeded to his solution [52]. He solved Eq. (3.23) when the left hand side is set equal to zero. This produces the asymptotic fixation or survival probability, $P_\infty = \lim_{t \rightarrow \infty} u(x_0, t)$. Kimura found

$$P_\infty = \frac{1 - e^{-4Nsx_0}}{1 - e^{-4Ns}}. \quad (3.24)$$

I will not review that calculation here. Instead a more illuminating exercise will be to find a first order approximation, valid for very large N , that gives us the desired results. I will show that in the thermodynamic, or large system size limit, the fate of a new mutant allele is described by a simple branching process. I will also show that when the new allele is selectively neutral, the branching process is critical. This limiting result is important to understand since many stochastic processes have rare event regimes that converge to

branching processes for large system sizes. Indeed, I show in chapter 4 that the critical branching process can be understood within the framework of the renormalization group as a fixed point that describes the critical behavior of many stochastic models above their upper critical dimension.

To show how this regime of the model emulates a branching process, we start with the master equation for Moran's model given in Eq. (3.3) and set $\mu = 0$. Then we use the conservation law, $N = n_a + n_A$, to eliminate n_A and simply use n for n_a . We are left with a master equation describing only the A 's. For clarity, and since we are not interested in a system size expansion, we do not use the shift operators introduced in Eq. (3.25) here. The master equation reads

$$\begin{aligned} \frac{\partial P_n}{\partial t} = & \frac{1+s}{2N} [(N-n+1)(n-1)P_{n-1} - (N-n)nP_n] \\ & + \frac{1-s}{2N} [(N-n+1)(n+1)P_{n+1} - (N-n)nP_n]. \end{aligned} \quad (3.25)$$

(For ease of notation, I have dropped the explicit time-dependence of P_n .) For very large N , we can ignore terms of $\mathcal{O}(N^{-1})$ in the master equation, which simplifies to

$$\begin{aligned} \frac{\partial P_n}{\partial t} = & \frac{1+s}{2} [(n-1)P_{n-1} - nP_n] \\ & + \frac{1-s}{2} [(n+1)P_{n+1} - nP_n]. \end{aligned} \quad (3.26)$$

We recognize Eq. (3.26) as a master equation that describes a branching process with branch rate

$$\gamma = \frac{1+s}{2}, \quad (3.27)$$

and decay rate

$$\lambda = \frac{1-s}{2}. \quad (3.28)$$

giving us the growth rate

$$\Delta = \gamma - \lambda = s.$$

I should mention that this result (3.26) can be very easily derived using the operator form of Moran's model (3.16).

Branching processes were discussed as an example in section 2.4.2 where we illustrated operator techniques. An alternative derivation using classical methods is presented in appendix C. We can use either results, Eq. (C.8) or Eq. (2.62). The time dependence of the survival probability is given by

$$P_S(t) = \frac{2s}{1+s - (1-s)e^{-\Delta t}}. \quad (3.29)$$

Its asymptotic value for $s > 0$, the "fixation" probability is,

$$P_\infty \equiv \lim_{t \rightarrow \infty} P_S(t) = \frac{\Delta}{\gamma} = \frac{2s}{1+s}, \quad (3.30)$$

which for small s tells us

$$P_\infty \simeq 2s. \quad (3.31)$$

This agrees with a very early calculation by Haldane [54], and the result due to Kimura, Eq.(3.24) with $x_0 = 1/2N$ (since Kimura uses the genetics convention of taking the population size to be $2N$),

$$P_\infty = \frac{1 - e^{-2s}}{1 - e^{-4Ns}} \simeq 2s, \quad (3.32)$$

for small s and $N \rightarrow \infty$.

If we repeat the above argument for a selectively disadvantageous new allele A , in other words, we let $s \rightarrow -s$, we find from Eq. (C.12) the size distribution, Q_n describing the total number n of new alleles produced in the avalanche. For large n ,

$$Q_n \simeq \frac{1}{2\sqrt{\pi}} n^{-3/2} e^{-n/\zeta} \quad (3.33)$$

with the characteristic size

$$\zeta = -\frac{1}{\ln(1-s^2)}. \quad (3.34)$$

When $s = 0$, the process is critical, ζ diverges and both Q_n and $P_S(t)$ follow power laws

$$P_n \sim n^{-3/2}, \quad (3.35)$$

and

$$P_S(t) = \frac{1}{1+t} \sim t^{-1}. \quad (3.36)$$

3.7 Discussion.

Taking into account the fluctuations, we have found that the effective potential is not the same as the mean-field potential or “fitness landscape” as it sometimes is called in the literature only concerned with a mean-field dynamical systems description. Even when $2N\mu \gg 1$, minimization of the effective potential reveals a shift of $\mathcal{O}(N^{-1})$ in the steady state value of x ,

$$\hat{x} = \frac{1}{2} \left(1 + \sqrt{1 + \frac{4}{s^2} \left(\mu - \frac{1}{2N} \right)} \right) - \frac{1}{s} \left(\mu - \frac{1}{2N} \right).$$

For $2N\mu < 1$, the situation is qualitatively different as the stability is no longer given by the mean-field result. In discussions of evolutionary models, there is frequent reference to “fitness” landscapes or “selective surfaces”. These approaches typically ignore the full implications of the particular type of state-dependent noise that arises from a microscopic description, and as we have shown are only valid in the limit $2N\mu \gg 1$. In fact, most models involving selective surfaces ignore mutation completely, making the concept of a fitness landscape quite irrelevant. This observation is the basis of Kimura’s neutral theory that postulates a model of evolution at the molecular level that is very different than the sort of mass selection envisioned by Fisher and other followers of the determinist school of thought.

We have re-examined the Wright-Fisher/Moran model from the point of view of the ‘effective potential’ and found that in the more realistic limit of small mutation rate, it

is this effective potential that is the relevant landscape upon which the system evolves. More germane to this thesis, however, the question arises as to whether or not the broken symmetry and critical behavior exhibited in the Wright-Fisher/Moran model persists when this model is generalized to include spatial structure, or in physics terminology, spatial degrees of freedom.

Chapter 4

The Stepping-Stone Model Revisited

4.1 Introduction.

Wright was the first to quantitatively examine the effects of random genetic drift for geographically distributed populations with his ‘Island’ model. According to this model, in which time is discrete, symmetric migration takes place between an infinite number of subpopulations or colonies, which otherwise behave according to the Wright-Fisher model discussed in chapter 2. As a generalization of the Island model, Kimura and Weiss [1] introduced the so-called “Stepping-Stone” model, which in addition to long-range migration, incorporated nearest-neighbour migration between the colonies, now arranged on a lattice. By solving the steady-state correlation function of this model, they were able to derive a number of interesting results concerning genetic differentiation among sub-populations.

Unfortunately, the approach described in the previous chapter of finding an effective potential does not help us much for the Stepping-Stone model. Such a potential can actually be derived, but it does not provide a convenient way to calculate the desired results. We must resort to other means, and will be able to infer the behavior of this potential with respect to symmetry breaking and critical properties.

We will consider two slightly different models in this chapter, both of which can be thought of as lattice generalizations of Moran’s model. We find that these models are equivalent to the original Stepping-Stone model of Kimura and Weiss [1] in the same sense

as the equivalence of Moran's model and the Wright-Fisher model discussed in the previous chapter. In this chapter, I formulate these models using the creation-annihilation operator formalism described in chapter 2. First, I reproduce results reminiscent of those from Kimura and Weiss[1][2] and Malécot[55] for the steady-state two-point correlation function. I then extend their work by solving the dynamics of the two-point correlation function for the neutral theory and also for small, but finite selection in the broken symmetry state. An alternative method for finding this correlation function, which is closer to the original work by Kimura and Weiss, would be to use the response function formalism or the operator form of the Fokker-Planck equation introduced in chapter 2. In section 5 of this chapter, I will give these alternative formulations of the Stepping-Stone model and will find that they are indeed equivalent to the occupation number description.

Selection was absent in the Stepping-Stone models considered by both Kimura and Weiss[1][2] and Malécot[55]. Several authors, however, have discussed the Stepping-Stone model with selection and provided theorems concerning the ergodic properties of the model for various cases of mutation and selection. Both Itatsu[56], and Shiga and Uchiyama[57] consider asymmetric mutation rates, that is, $A \rightarrow a$ with rate μ and $a \rightarrow A$ at rate v , and prove that a heterozygous steady-state exists when $\mu > 0$, $v > 0$ and s is arbitrary. If one of the mutation rates is identically 0 and the other strictly positive, then a heterozygous steady-state is found only for s greater than some critical value s_c . While Shiga[57] does not rule out the possibility that $s_c = 0$, Itatsu finds that $\frac{\mu}{1-\mu} \leq s_c \leq 1$. For the mean-field dynamical system, it is easy to show that for this case, $s_c = \mu$.¹ In section 4.6, I consider the case where $\mu = 0$, $v = 0$ and $s \neq 0$. For this case both Itatsu, and Shiga and Uchiyama find that the only stable steady state is the fixed state of the selectively advantageous allele. This finding is confirmed by the analysis in section 4.6.

¹It should be noted that Itatsu parametrizes selection in slightly different manner. His s is half of that used in this thesis.

In addition to this, the asymptotic dynamics of the approach to this steady-state.

The two models considered in this thesis have the same on-site processes, but differ in the way neighboring sites are coupled. In the first model, which I will refer to as the ‘Exchange Coupled Stepping-Stone (ECSS) Model’, the number of individuals at each site n , is fixed and independent of position in the lattice. Nearest neighbor sites are coupled by a random exchange of a pair of individuals, one from each site. Since all processes, both on-site and nearest-neighbor, conserve the number of individuals, we have only one independent field. The second model on the other hand, I refer to as the ‘Diffusively Coupled Stepping-Stone (DCSS) Model’, uses random diffusion or dispersal of individuals to neighbouring sites, does not conserve the number of individuals at each site and therefore has two independent fields. In the DCSS model for a homogeneously distributed population, however, the average number at each site is constant. I will show that the field representing the total number is independent of the gene-frequency field and therefore described by diffusive fluctuations.

First, I consider the on-site processes which are identical in both models. In the following, a -operators describe a -type individuals, while b -operators will describe A -type individuals. The derivation of the Stepping-Stone model here is a generalization of the operator form of Moran’s model presented in chapter 3. Symmetric mutation is the same as in Moran’s model and is described by one-body operators

$$\mathcal{L}_{\text{mut}} = \mu \sum_{\mathbf{r}} (a_{\mathbf{r}}^{\dagger} b_{\mathbf{r}} - b_{\mathbf{r}}^{\dagger} b_{\mathbf{r}} + b_{\mathbf{r}}^{\dagger} a_{\mathbf{r}} - a_{\mathbf{r}}^{\dagger} a_{\mathbf{r}}). \quad (4.1)$$

Genetic drift along with a selective advantage is given by two-body operators

$$\mathcal{L}_{\text{sel/drift}} = \frac{1}{2n} \sum_{\mathbf{r}} ((1-s) a_{\mathbf{r}}^{\dagger} a_{\mathbf{r}}^{\dagger} a_{\mathbf{r}} b_{\mathbf{r}} + (1+s) b_{\mathbf{r}}^{\dagger} b_{\mathbf{r}}^{\dagger} a_{\mathbf{r}} b_{\mathbf{r}} - 2a_{\mathbf{r}}^{\dagger} b_{\mathbf{r}}^{\dagger} a_{\mathbf{r}} b_{\mathbf{r}}). \quad (4.2)$$

Without coupling, we would have N (the number of sites) independent populations of size n , evolving under Moran’s dynamics.

4.1.1 The Exchange Coupled Stepping-Stone Model (ECSS).

Now the total Liouville operator is given by

$$\mathcal{L}_{\text{ECSS}} = \mathcal{L}_{\text{ex}} + \mathcal{L}_{\text{mut}} + \mathcal{L}_{\text{sel/drift}}, \quad (4.3)$$

with the exchange processes described by \mathcal{L}_{ex} , given by

$$\begin{aligned} \mathcal{L}_{\text{ex}} = & \frac{m}{2n} \sum_{\mathbf{r}} \sum_{\mathbf{e}} \left(b_{\mathbf{r}}^{\dagger} a_{\mathbf{r}+\mathbf{e}}^{\dagger} a_{\mathbf{r}} b_{\mathbf{r}+\mathbf{e}} - a_{\mathbf{r}}^{\dagger} b_{\mathbf{r}+\mathbf{e}}^{\dagger} a_{\mathbf{r}} b_{\mathbf{r}+\mathbf{e}} \right) \\ & + \frac{m}{2n} \sum_{\mathbf{r}} \sum_{\mathbf{e}} \left(a_{\mathbf{r}}^{\dagger} b_{\mathbf{r}+\mathbf{e}}^{\dagger} b_{\mathbf{r}} a_{\mathbf{r}+\mathbf{e}} - b_{\mathbf{r}}^{\dagger} a_{\mathbf{r}+\mathbf{e}}^{\dagger} b_{\mathbf{r}} a_{\mathbf{r}+\mathbf{e}} \right), \end{aligned}$$

where \mathbf{e} is a lattice vector of magnitude l , and the lattices can be chosen to be *e.g* the linear, square or cubic lattice in $d = 1, 2, 3$, respectively. Here m is the migration rate. Usually in the physics literature, m is used as an index. We break with that tradition here and bend to the convention used in population genetics. In the continuum limit, however, we denote by D , the diffusion constant, which is given by $D = ml^2$.

4.1.2 The Diffusively Coupled Stepping-Stone Model (DCSS).

For this model, we have

$$\mathcal{L}_{\text{DCSS}} = \mathcal{L}_{\text{diff}} + \mathcal{L}_{\text{mut}} + \mathcal{L}_{\text{sel/drift}},$$

the only difference from the ECSS model is the random walk term $\mathcal{L}_{\text{diff}}$, given by

$$\begin{aligned} \mathcal{L}_{\text{diff}} = & m \sum_{\mathbf{r}} \sum_{\mathbf{e}} \left(a_{\mathbf{r}+\mathbf{e}}^{\dagger} a_{\mathbf{r}} + a_{\mathbf{r}-\mathbf{e}}^{\dagger} a_{\mathbf{r}} - 2a_{\mathbf{r}}^{\dagger} a_{\mathbf{r}} \right) \\ & + m \sum_{\mathbf{r}} \sum_{\mathbf{e}} \left(b_{\mathbf{r}+\mathbf{e}}^{\dagger} b_{\mathbf{r}} + b_{\mathbf{r}-\mathbf{e}}^{\dagger} b_{\mathbf{r}} - 2b_{\mathbf{r}}^{\dagger} b_{\mathbf{r}} \right), \end{aligned} \quad (4.4)$$

which describes “hopping” to neighboring sites. Once again, in both models, $D = ml^2$ appears as the diffusion constant in the continuum limit.

4.2 The Neutral Regime.

Under vanishing selection, no allele is favoured and as was discussed in the previous chapter, if the mutation rate also vanishes, the system becomes critical. Moreover, even in the spatial model, the critical point remains at $\mu = 0$. This fact will become clear from inspection of the total Liouville operator—none of the interacting terms in $\mathcal{L}_{\text{sel/drift}}$ produce corrections to the single particle propagator (the Green's function). We will also be able to show this by examining the Heisenberg equations for the densities. These do not couple to higher correlations and are dynamically stable for non-zero μ . In this section we will obtain exact results for the two-point correlation function for both finite and vanishing mutation rate. We will also find that the critical behavior of the d -dimensional system for $\mu = 0$ is intimately tied to return probability of a random walk in d dimensions. Results similar to those of this section were found by Sawyer[3] and Nagalaki[4] for an infinite alleles model and are discussed below.

First, we make a canonical transformation to sum and difference operators of the form used for the Moran model in chapter 3,

$$\begin{aligned} c &= (a - b) / n & d &= a + b \\ c^\dagger &= \frac{n}{2} (a^\dagger - b^\dagger) & d^\dagger &= \frac{1}{2} (a^\dagger + b^\dagger) \end{aligned} \quad (4.5)$$

This is analogous to the transformation $z = 2x - 1$, of the Wright-Fisher/Moran Fokker-Planck equation mentioned in the last chapter, but carried out at every site. Once again we stress the necessity of using a canonical transformation so that we preserve the commutation relations, *i.e.* so that $[c, c^\dagger] = 1$ and $[d, d^\dagger] = 1$. For the on-site terms, this results in

$$\mathcal{L}_{\text{mut}} = -2\mu \sum_{\mathbf{r}} c_{\mathbf{r}}^\dagger c_{\mathbf{r}} \quad (4.6)$$

and

$$\mathcal{L}_{\text{drift}} = (2n)^{-1} \sum_{\mathbf{r}} \left(n^{-2} c_{\mathbf{r}}^{\dagger} c_{\mathbf{r}}^{\dagger} d_{\mathbf{r}} d_{\mathbf{r}} - c_{\mathbf{r}}^{\dagger} c_{\mathbf{r}}^{\dagger} c_{\mathbf{r}} c_{\mathbf{r}} \right) \quad (4.7)$$

In the ECSS Model, the conservation law at each site allows one to replace the $d_{\mathbf{r}}$ operators with the real number n , and the d^{\dagger} operators with 1. Including the exchange and diffusion parts, we now find respectively,

$$\begin{aligned} \mathcal{L}_{\text{ECSS}} &= m \sum_{\mathbf{r}} \sum_{\mathbf{e}} \left(c_{\mathbf{r}+\mathbf{e}}^{\dagger} c_{\mathbf{r}} + c_{\mathbf{r}-\mathbf{e}}^{\dagger} c_{\mathbf{r}} - 2c_{\mathbf{r}}^{\dagger} c_{\mathbf{r}} \right) \\ &\quad - 2\mu \sum_{\mathbf{r}} c_{\mathbf{r}}^{\dagger} c_{\mathbf{r}} \\ &\quad + (2n)^{-1} \sum_{\mathbf{r}} \left(c_{\mathbf{r}}^{\dagger} c_{\mathbf{r}}^{\dagger} - c_{\mathbf{r}}^{\dagger} c_{\mathbf{r}}^{\dagger} c_{\mathbf{r}} c_{\mathbf{r}} \right) \end{aligned} \quad (4.8)$$

and

$$\begin{aligned} \mathcal{L}_{\text{DCSS}} &= m \sum_{\mathbf{r}} \sum_{\mathbf{e}} \left(c_{\mathbf{r}+\mathbf{e}}^{\dagger} c_{\mathbf{r}} + c_{\mathbf{r}-\mathbf{e}}^{\dagger} c_{\mathbf{r}} - 2c_{\mathbf{r}}^{\dagger} c_{\mathbf{r}} \right) \\ &\quad + m \sum_{\mathbf{r}} \sum_{\mathbf{e}} \left(d_{\mathbf{r}+\mathbf{e}}^{\dagger} d_{\mathbf{r}} + d_{\mathbf{r}-\mathbf{e}}^{\dagger} d_{\mathbf{r}} - 2d_{\mathbf{r}}^{\dagger} d_{\mathbf{r}} \right) \\ &\quad - 2\mu \sum_{\mathbf{r}} c_{\mathbf{r}}^{\dagger} c_{\mathbf{r}} \\ &\quad + \frac{1}{2n} \sum_{\mathbf{r}} \left(n^{-2} c_{\mathbf{r}}^{\dagger} c_{\mathbf{r}}^{\dagger} d_{\mathbf{r}} d_{\mathbf{r}} - c_{\mathbf{r}}^{\dagger} c_{\mathbf{r}}^{\dagger} c_{\mathbf{r}} c_{\mathbf{r}} \right) \end{aligned} \quad (4.9)$$

If we now look at the Heisenberg equation for the expectation of the difference field, we find

$$\frac{d}{dt} \langle c_{\mathbf{r}} \rangle = -2\mu \langle c_{\mathbf{r}} \rangle, \quad (4.10)$$

which tells us that the critical point remains at $\mu = 0$.

4.2.1 Fourier transformation to wave-vector space.

We will assume periodic boundary conditions. For translationally invariant systems, further analysis is accomplished and greatly simplified by transforming to wave-vector space.

We use Fourier transforms defined by

$$\begin{aligned} c_{\mathbf{r}} &= N^{-1/2} \sum_{\mathbf{k}} c_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}}, & c_{\mathbf{r}}^{\dagger} &= N^{-1/2} \sum_{\mathbf{k}} c_{\mathbf{k}}^{\dagger} e^{-i\mathbf{k}\cdot\mathbf{r}} \\ d_{\mathbf{r}} &= N^{-1/2} \sum_{\mathbf{k}} d_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}}, & d_{\mathbf{r}}^{\dagger} &= N^{-1/2} \sum_{\mathbf{k}} d_{\mathbf{k}}^{\dagger} e^{-i\mathbf{k}\cdot\mathbf{r}} \end{aligned} \quad (4.11)$$

so that

$$\begin{aligned} c_{\mathbf{k}} &= N^{-1/2} \sum_{\mathbf{r}} c_{\mathbf{r}} e^{-i\mathbf{k}\cdot\mathbf{r}} & c_{\mathbf{k}}^{\dagger} &= N^{-1/2} \sum_{\mathbf{r}} c_{\mathbf{r}}^{\dagger} e^{i\mathbf{k}\cdot\mathbf{r}} \\ d_{\mathbf{k}} &= N^{-1/2} \sum_{\mathbf{r}} d_{\mathbf{r}} e^{-i\mathbf{k}\cdot\mathbf{r}} & d_{\mathbf{k}}^{\dagger} &= N^{-1/2} \sum_{\mathbf{r}} d_{\mathbf{r}}^{\dagger} e^{i\mathbf{k}\cdot\mathbf{r}} \end{aligned} \quad (4.12)$$

In the DCSS model, we have diffusion terms for both fields

$$\mathcal{L}_D = -m \sum_{\mathbf{k}} \omega_{\mathbf{k}} c_{\mathbf{k}}^{\dagger} c_{\mathbf{k}} - m \sum_{\mathbf{k}} \sum_{\delta} \omega_{\mathbf{k}} d_{\mathbf{k}}^{\dagger} d_{\mathbf{k}}, \quad (4.13)$$

where

$$\omega_{\mathbf{k}} = \sum_{\mathbf{e}} (2 - e^{i\mathbf{k}\cdot\mathbf{e}} - e^{-i\mathbf{k}\cdot\mathbf{e}}) \quad (4.14)$$

On hypercubic lattices in d dimensions,

$$\omega_{\mathbf{k}} = 4 \sum_{j=1}^d \sin^2 \frac{k_j l}{2\pi}, \quad (4.15)$$

where l is the lattice constant. For small kl , $\omega_{\mathbf{k}} \sim \mathbf{k}^2 l^2$. In the ECSS model, only the first term in Eq. (4.13) is present.

The mutation term is the same for both ECSS and DCSS models,

$$\mathcal{L}_{\text{mut}} = -2\mu \sum_{\mathbf{k}} c_{\mathbf{k}}^{\dagger} c_{\mathbf{k}}. \quad (4.16)$$

The selection/drift term for the ECSS model is

$$\mathcal{L}_{\text{drift}} = \frac{1}{2n} \sum_{\mathbf{k}} c_{\mathbf{k}}^{\dagger} c_{-\mathbf{k}}^{\dagger} - \frac{1}{2nN} \sum_{\mathbf{k}_1, \mathbf{k}_2, \mathbf{q}} c_{\mathbf{k}_1}^{\dagger} c_{\mathbf{k}_2}^{\dagger} c_{\mathbf{k}_1+\mathbf{q}} c_{\mathbf{k}_2-\mathbf{q}}, \quad (4.17)$$

and for the DCSS model

$$\mathcal{L}_{\text{drift}} = \frac{1}{2nN} \sum_{\mathbf{k}_1, \mathbf{k}_2, \mathbf{q}} \left(n^{-2} c_{\mathbf{k}_1}^\dagger c_{\mathbf{k}_2}^\dagger d_{\mathbf{k}_1+\mathbf{q}} d_{\mathbf{k}_2-\mathbf{q}} - c_{\mathbf{k}_1}^\dagger c_{\mathbf{k}_2}^\dagger c_{\mathbf{k}_1+\mathbf{q}} c_{\mathbf{k}_2-\mathbf{q}} \right). \quad (4.18)$$

4.2.2 Equation of motion for the two-point correlation function.

I now proceed to solve the equation of motion for the Fourier transform of the two-point correlation function of the difference field, which is defined by the expectation

$$\tilde{G}_{\phi\phi}(\mathbf{k}, t) \equiv \langle c_{\mathbf{k}} c_{-\mathbf{k}} \rangle = \langle S | c_{\mathbf{k}} c_{-\mathbf{k}} | \psi \rangle. \quad (4.19)$$

Here I use the Heisenberg representation, so that

$$\frac{d}{dt} (c_{\mathbf{k}} c_{-\mathbf{k}}) = [c_{\mathbf{k}} c_{-\mathbf{k}}, \mathcal{L}]$$

Taking expectations, one has

$$\frac{d}{dt} \tilde{G}_{\phi\phi}(\mathbf{k}, t) = \langle [c_{\mathbf{k}} c_{-\mathbf{k}}, \mathcal{L}] \rangle \quad (4.20)$$

The right side of Eq. (4.20) is found by direct evaluation of the commutator, using the canonical commutation relations. Examples of this procedure for creation and annihilation operators were given in chapter 2. A vast simplification occurs upon taking the expectation. This is due to the identity

$$\langle S | c_{\mathbf{k}}^\dagger | \psi \rangle = 0 \quad (4.21)$$

I do not present the somewhat tedious algebra here. The resulting equation of motion for the ECSS model is

$$\frac{d}{dt} \tilde{G}_{\phi\phi}(\mathbf{k}, t) = -2(m\omega_{\mathbf{k}} + 2\mu) \tilde{G}_{\phi\phi}(\mathbf{k}, t) + n^{-1} - \frac{1}{nN} \sum_{\mathbf{q}} \tilde{G}_{\phi\phi}(\mathbf{q}, t). \quad (4.22)$$

For the diffusively-coupled model, we have the system of equations

$$\begin{aligned} \frac{d}{dt} \tilde{G}_{\phi\phi}(\mathbf{k}, t) &= -2(m\omega_{\mathbf{k}} + 2\mu) \tilde{G}_{\phi\phi}(\mathbf{k}, t) \\ &+ \frac{1}{n^3 N} \sum_{\mathbf{q}} \tilde{H}_{\phi\phi}(\mathbf{q}, t) - \frac{1}{nN} \sum_{\mathbf{q}} \tilde{G}_{\phi\phi}(\mathbf{q}, t), \end{aligned} \quad (4.23)$$

$$\frac{d}{dt} \tilde{H}_{\phi\phi}(\mathbf{k}, t) = -2m\omega_{\mathbf{k}} \tilde{H}_{\phi\phi}(\mathbf{k}, t), \quad (4.24)$$

where I have denoted by $\tilde{H}_{\phi\phi}(\mathbf{k}, t)$, the Fourier transformed two-point correlation function of the total number field,

$$\tilde{H}_{\phi\phi}(\mathbf{k}, t) = \langle d_{\mathbf{k}} d_{-\mathbf{k}} \rangle = \langle S | d_{\mathbf{k}} d_{-\mathbf{k}} | \psi \rangle.$$

representing the Fourier transform of the total number field. For the case of a homogeneously distributed population, one has

$$\tilde{H}_{\phi\phi}(\mathbf{q}, t) = \begin{cases} Nn^2, & \mathbf{q} = \mathbf{0} \\ 0, & \mathbf{q} \neq \mathbf{0} \end{cases},$$

and therefore one recovers the same equation for $G_{\phi\phi}(\mathbf{k}, t)$ as in the exchange-coupled model. *For the rest of this thesis, we will assume a homogeneously distributed population and work only with the difference fields.* This means that as far as we're concerned, the DCSS and the ECSS are equivalent. It would be interesting, however, to use the system of equations, Eqs. (4.23), to examine situations where the population is initially localized over some small region.

We note that unlike many other stochastic population models, the equation of motion for the two-point correlation function does not couple to higher order correlation functions. This fact allows us to find exact results for this model.

4.2.3 The steady-state.

Using Eq. (4.22), we find the steady-state solution by trying the solution

$$\tilde{G}_{\phi\phi}(\mathbf{k}) = \frac{C}{m\omega_{\mathbf{k}} + 2\mu}, \quad (4.25)$$

where C is an undetermined constant. This results in an equation for C ,

$$\begin{aligned} C &= \frac{1}{n} - \frac{C}{nN} \sum_{\mathbf{q}} \frac{1}{m\omega_{\mathbf{q}} + 2\mu} \\ &= (1 - CI)/n, \end{aligned} \quad (4.26)$$

where

$$I = N^{-1} \sum_{\mathbf{q}} \frac{1}{m\omega_{\mathbf{q}} + 2\mu}.$$

with the solution for C ,

$$C = \frac{1}{n + I}$$

We recognize the two-point correlation function, in real space and evaluated at $\mathbf{r} = \mathbf{0}$, as the sum

$$G_{\phi\phi}(\mathbf{r}) = \frac{1}{N} \sum_{\mathbf{q}} \tilde{G}_{\phi\phi}(\mathbf{q}),$$

and using the solution for C ,

$$= \frac{I}{n + I}.$$

To get the correlation function in real space, we need to perform the inverse Fourier transform. We wish to study the spatial dependence of the correlation function in the continuum limit in which details such as lattice structure do not matter. In the continuum limit $\omega_{\mathbf{k}} = k^2 l^2$, so

$$m\omega_{\mathbf{q}} = mk^2 l^2 = Dk^2,$$

where we have obtained the diffusion constant $D = ml^2$. The two-point correlation function becomes

$$\tilde{G}_{\phi\phi}(\mathbf{k}) = \frac{C}{Dk^2 + 2\mu}, \quad (4.27)$$

which is C times a very familiar quantity in statistical physics: the two-point correlator of the Gaussian model. Furthermore, in the continuum limit the inverse Fourier transform becomes an integral

$$N^{-1} \sum_{\mathbf{k}} \rightarrow \frac{1}{(2\pi)^d} \int_0^\Lambda d^d \mathbf{k},$$

where $\Lambda \approx l^{-1}$ is a cut-off parameter. We define the correlation length

$$\xi \equiv \sqrt{\frac{D}{2\mu}}, \quad (4.28)$$

and the dimensionless constant

$$\kappa = n^{-1} I = \frac{1}{n(2\pi)^d} \int_0^\Lambda d^d \mathbf{k} \frac{1}{D\mathbf{k}^2 + 2\mu}, \quad (4.29)$$

and end up with the result

$$G_{\phi\phi}(\mathbf{r}) = \frac{\kappa}{1 + \kappa} \exp\left(-\frac{r}{\xi}\right) / |\mathbf{r}|^{\frac{d-1}{2}}. \quad (4.30)$$

The integral for κ is easily evaluated

$$\kappa = \begin{cases} (n^2 D \mu)^{-1/2}, & d = 1 \\ \frac{\ln(\Lambda \xi)}{2\pi D n}, & d = 2 \\ \frac{\Lambda}{2\pi^2 D n}, & d = 3 \end{cases}. \quad (4.31)$$

All along, we have been measuring time in generations. We can explicitly put back in the generation τ_g if needed,

$$\kappa = \begin{cases} (n^2 D \mu \tau_g)^{-1/2}, & d = 1 \\ \frac{\ln(\Lambda \xi)}{2\pi \tau_g D n}, & d = 2 \\ \frac{\Lambda}{2\pi^2 \tau_g D n}, & d = 3 \end{cases}. \quad (4.32)$$

I will refer to κ as the “strength of local differentiation”, for as κ becomes large, the amplitude of correlation function becomes one and the local populations are fixed to one

allele or the other. When κ is small, the local populations are polymorphic. As an alternative to κ , a more appropriate measure, would be the quantity $\frac{\kappa}{1+\kappa}$. In population genetics, this quantity referred to as F_{ST} , where F_{ST} stands for *fixation index*. If we go back to the lattice description and use our results for κ to calculate F_{ST} , we find

$$F_{ST} = \begin{cases} \frac{1}{1+n\sqrt{m\mu}}, & d = 1 \\ \frac{1}{1+4\pi mn/\ln(\frac{m}{2\mu})}, & d = 2 \\ \frac{1}{1+2\pi^2 mn}, & d = 3 \end{cases} \quad (4.33)$$

These results for F_{ST} should be compared to the steady-state value of F_{ST} in Wright's Island model [58],

$$F_{ST} = \frac{1}{1+4nm}. \quad (4.34)$$

The Island model of Wright also consists of colonies of size n , but in contrast to the Stepping-Stone model, each colony exchanges migrants with all the rest. We see that our results for F_{ST} for the Stepping-Stone model only qualitatively agree with the Island model for $d = 3$ ($4nm$ vs. $4\pi^2 nm$, as Wright uses the genetics convention of $2N$ for the population size). For $d \leq 2$, F_{ST} for the Stepping-Stone model depends explicitly on the mutation rate and is therefore qualitatively different than its cousin in the Island model.

Figure 4.1 shows several snapshots from computer simulations of the $d = 2$, Stepping-Stone model with $N = 200 \times 200$ sites, and illustrates the effects of varying ξ and κ . Simulations in the same row have the same correlation length, while those in the same column have the same value of F_{ST} . The method of simulation is described in appendix B.

The results up to this point, while perhaps expressed somewhat differently, are similar to those of Kimura and Weiss [1][2]. We now leave them behind.

4.3 Broken Symmetry or $d \leq 2$.

For any finite size, or for spatial dimension $d \leq 2$, I diverges as $\mu \rightarrow 0$, and hence

$$G_{\phi\phi}(\mathbf{0}) = \frac{I}{n + I} \Rightarrow 1$$

This implies fixation and the spontaneously broken symmetry likewise observed in the Wright-Fisher/Moran model. Another way to see this is to look at $\Phi_{\mathbf{k}}$, for $\mathbf{k} \rightarrow \mathbf{0}$, which is found to be

$$\tilde{G}_{\phi\phi}(\mathbf{k}) = \delta(\mathbf{k}).$$

Perhaps the simplest way to see this, however, is to look at the results for κ in the continuum limit. For $d \leq 2$, κ diverges as $\mu \rightarrow 0$, and the fixation index $F_{ST} \rightarrow 1$. This is not the case for $d > 2$ as κ is independent of μ . This result combined with the diverging correlation length ξ , means that no matter how large we make the system, it will eventually reach fixation.

4.4 Critical dynamics of the two-point correlation function.

4.4.1 The decay of heterozygosity.

To go beyond the steady-state solution and solve for the dynamics, I return to the equation of motion 4.22, which we re-express as

$$\frac{d}{dt} \tilde{G}_{\phi\phi}(\mathbf{k}, \mathbf{t}) = -2(D\omega_{\mathbf{k}} + 2\mu) \tilde{G}_{\phi\phi}(\mathbf{k}, \mathbf{t}) + n^{-1}h(t), \quad (4.35)$$

with

$$\begin{aligned} h(t) &= 1 - \Phi_{\mathbf{r}=\mathbf{0}} \\ &= 1 - F_{ST}(t). \end{aligned}$$

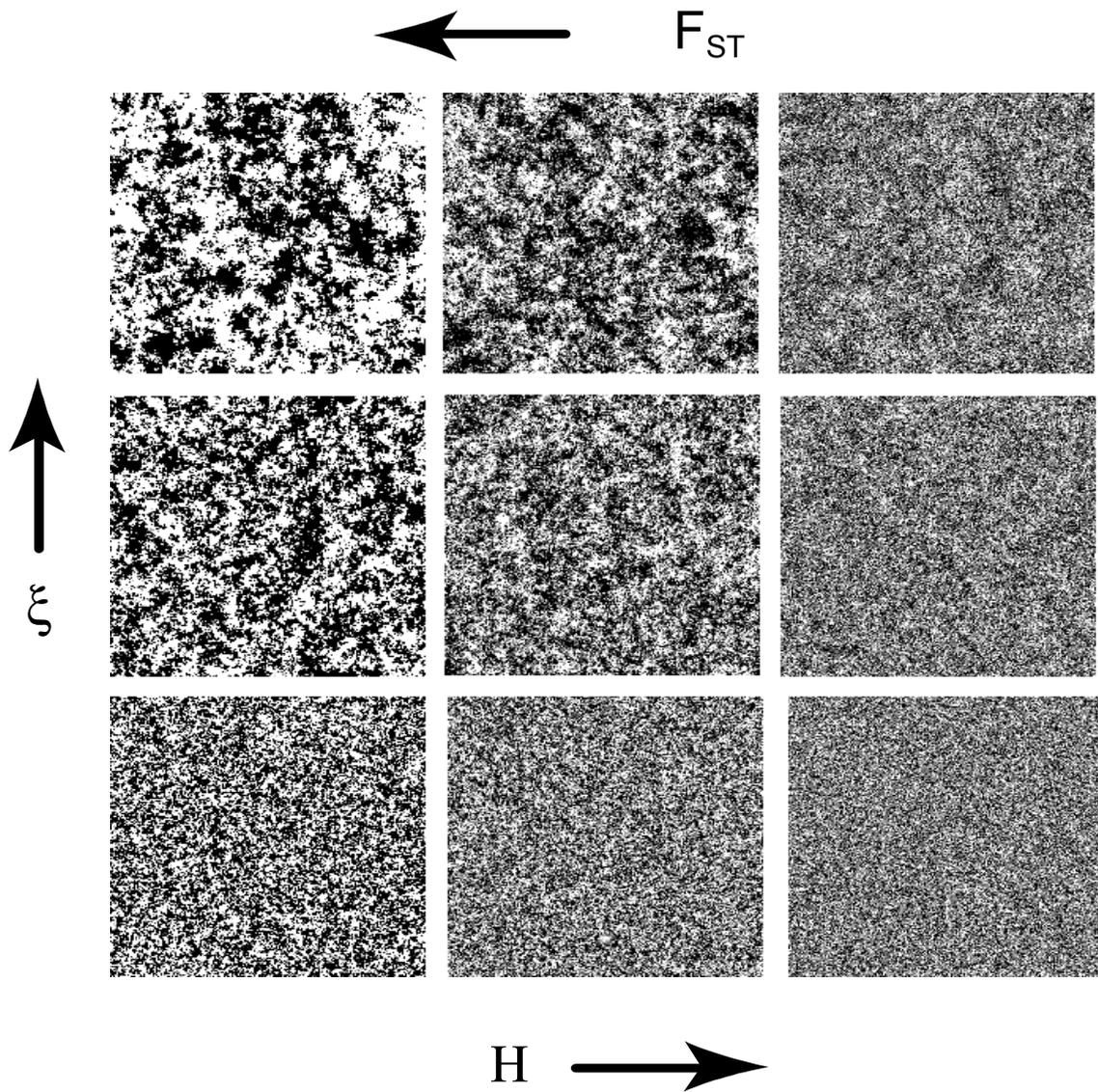


Figure 4.1: Snapshots of computer simulations of the $d = 2$ Stepping-Stone model on the 200×200 site square lattice. The mutation rate and migration rates were adjusted so that the simulations in each row have the same correlation length, while the simulations in the same column have the same value of F_{ST} .

The quantity $h(t)$ is known in the genetics literature as the *heterozygosity* or more properly as the local heterozygosity.

From above, we note that for $\mu = 0$, and $d \leq 2$, $h(t)$ decays asymptotically to 0. The integral solution to 4.35 is given by

$$\tilde{G}_{\phi\phi}(\mathbf{k}, t) = \frac{1}{n} \int_0^t dt' e^{-2m(t-t')\omega_{\mathbf{k}}} h(t').$$

Here we have used the initial condition $\tilde{G}_{\phi\phi}(\mathbf{k}, 0) = 0$. This means that we start out with $n_a = n_A = n/2$ at each site. Summing over both sides of Eq. (4.35) results in an equation for $h(t)$

$$h(t) = 1 - \frac{1}{nN} \sum_{\mathbf{k}} \int dt' e^{-2m(t-t')\omega_{\mathbf{k}}} h(t'), \quad (4.36)$$

Changing the order of summation and integration, we recognize the sum

$$\sum_{\mathbf{k}} e^{-2m(t-t')\omega_{\mathbf{k}}} = G_{2m}(\mathbf{0}, t - t') \quad (4.37)$$

as the diffusion Green's function for a random walk in d dimensions, with initial and final position at the origin, and diffusion rate $2m$ (Appendix A). We have

$$h(t) = 1 - \frac{1}{n} \int_0^t dt' G_{2m}(\mathbf{0}, t - t') h(t'), \quad (4.38)$$

with $G_{2m}(\mathbf{0}, t)$ for the hypercubic lattices given by

$$G_{2m}(\mathbf{0}, t) = e^{-4mdt} [I_0(4mt)]^d,$$

where $I_0(4mt)$ is a modified Bessel function of the first kind. Taking the Laplace transform of both sides results in the equation

$$\bar{h}(s) = s^{-1} \frac{1}{1 + \bar{G}_{2m}(\mathbf{0}, s)/n}. \quad (4.39)$$

Using the scale transformation properties of the Laplace transform,

$$\bar{G}_{bm}(\mathbf{0}, s) = b^{-1} \bar{G}_m(\mathbf{0}, bs),$$

and taking $b = 4dm$, we can write

$$\bar{G}_{2m}(\mathbf{0}, s) = (4dm)^{-1} \bar{G}_{\frac{1}{2d}}(\mathbf{0}, 4dms),$$

where $\bar{G}_{\frac{1}{2d}}$ is the Laplace transform of a random walk with rate $\frac{1}{2d}$. As shown in appendix A, the $s = 0$ limit of $\bar{G}_{\frac{1}{2d}}$ is fundamental and universal constant that is related to the return probability $Q(\mathbf{0})$ of a random walk through

$$Q(\mathbf{0}) = 1 - \frac{1}{\bar{G}_{\frac{1}{2d}}(\mathbf{0}, 0)}. \quad (4.40)$$

For $d \leq 2$, $\lim_{s \rightarrow 0} \bar{G}_{\frac{1}{2d}}(\mathbf{0}, s) = \infty$, so as noted above $G_{\phi\phi}(\mathbf{r} = \mathbf{0}) \rightarrow 1$ at large times. For $d > 2$, however, $G_{\phi\phi}(\mathbf{r} = \mathbf{0})$ approaches a constant. We can summarize the situation for arbitrary dimension by writing

$$\lim_{t \rightarrow \infty} G_{\phi\phi}(\mathbf{r} = \mathbf{0}) = \frac{1}{1 + dnm(1 - Q_d)}. \quad (4.41)$$

To find the dynamics of the approach of the two-point correlation function evaluated at $\mathbf{r} = \mathbf{0}$ to its limiting value, we take the Laplace transform of the Green's function for random walks which is explicitly calculated for the linear ($d = 1$), square ($d = 2$) and cubic ($d = 3$) lattices in appendix A. Asymptotically, the Laplace Transform for small s has the dependence

$$\bar{G}_{2m}(\mathbf{0}, s) \sim \begin{cases} (8ms)^{-1/2}, & d = 1 \\ \frac{\ln(8m/s)}{16m\pi}, & d = 2 \\ \frac{C}{12m} + B\sqrt{\frac{s}{12^3 m^3}}, & d = 3 \end{cases}. \quad (4.42)$$

where the constants C and B for the $d = 3$ result are found from a simple relation to the universal return probability of a random walk. Using this result gives us the behavior of $\bar{h}(s)$ for small s ,

$$\bar{h}(s) \sim \begin{cases} \sqrt{\frac{n^2 m}{2s}}, & d = 1 \\ \frac{4m\pi}{ns \ln(8m/s)}, & d = 2 \\ \frac{3mn}{3mn+C} s^{-1}, & d = 3 \end{cases}. \quad (4.43)$$

Standard ‘Tauberian’ theorems[59] for Laplace transforms relate the behaviour of $h(s)$ for small s to the behaviour of $h(t)$ for large t . One finds

$$h(t) \sim \begin{cases} \frac{2}{n\sqrt{mt}}, d = 1 \\ \frac{1}{\ln(mn^2t/4)}, d = 2 \\ \frac{3mn}{3mn+C}, d = 3 \end{cases} . \quad (4.44)$$

Furthermore, even though the situation is qualitatively different for $d = 3$, in that $h(t)$ approaches some non-zero constant, it is interesting to determine how $h(t)$ approaches this constant. From the result above, if the constant is subtracted, one finds that the small s behaviour of the Laplace transform of $g(t) = h(t) - \lim_{t \rightarrow \infty} h(t)$ is given by

$$\bar{g}(s) \sim \sqrt{\frac{mn^2}{4s}},$$

and hence

$$g(t) \sim \sqrt{\frac{4}{n^2mt}}.$$

The integral equation for $h(t)$, Eq. (4.36), is readily solved by numerical quadrature. This solution is presented in Figure (4.2), confirming our results for the asymptotic decay of $h(t)$.

Several remarks are in order: 1) the logarithmic dependence of this quantity for $d = 2$ is indicative of what is termed ‘‘upper critical dimension’’ in statistical physics; 2) Another interesting feature is the ‘‘crossover’’ region at early times; but most importantly, 3) even though strictly speaking, and as we have argued in the preceding section, fixation occurs for $d = 2$, the extremely slow logarithmic dependence makes one wonder whether or not, in any practical sense, the system ever gets there.

In summary we have found that the asymptotic value of $G_{\phi\phi}(\mathbf{0}, t)$, known as F_{ST} in the genetics literature, is related to the return probability of a random walk through

$$\lim_{t \rightarrow \infty} F_{ST}(t) = \frac{1}{1 + dnm(1 - Q_d)}$$

In the thermodynamic limit its approach to this value is $\sim t^{-1/2}$ for $d = 1$, $\sim 1/\ln t$ for $d = 2$ and $\sim t^{-1/2}$ for $d = 3$.

4.4.2 The infinite alleles model.

Sawyer[3] and Nagalaki[4] both investigated a related neutral model introduced by Malécot[55] that differs from the present model by having infinitely many alleles or types. They found asymptotics similar to those of Eqn. (4.44) for $d = 1$ and 2, but for a related quantity. Sawyer's starting point is an equation derived by Malécot[55] that describes the time dependence of the probability of identity by descent. This quantity is denoted $I(\mathbf{r}_1, \mathbf{r}_2; t)$ and gives the probability that two randomly selected individuals located at \mathbf{r}_1 and \mathbf{r}_2 are of the same type due to one or more common ancestors. For infinitely many alleles and assuming that at $t = 0$ all types in the population are unique, as in done in the paper by Sawyer, the probability of identity by descent (i.b.d) equals the probability of identity by type (i.b.t.). A similar equation was studied by Nagalaki for the probability of identity by type.

One can show that in a two-allele model, the probability of i.b.t. is related to the correlation function $G_{\phi\phi}(\mathbf{r}_1, \mathbf{r}_2)$. First we observe that when n is large, the discrete site variables become frequencies which we denote by the operators

$$f_{\mathbf{r}} = n^{-1}a_{\mathbf{r}}.$$

The probability of i.b.t. is given by the expectation

$$I(\mathbf{r}_1, \mathbf{r}_2) = \langle f_{\mathbf{r}_1} f_{\mathbf{r}_2} + (1 - f_{\mathbf{r}_1})(1 - f_{\mathbf{r}_2}) \rangle,$$

where the first (second) term is the contribution from selecting two a 's (A 's). The f -operators are related to the c -operators by the transformation

$$f_{\mathbf{r}} = \frac{1 - c_{\mathbf{r}}}{2},$$

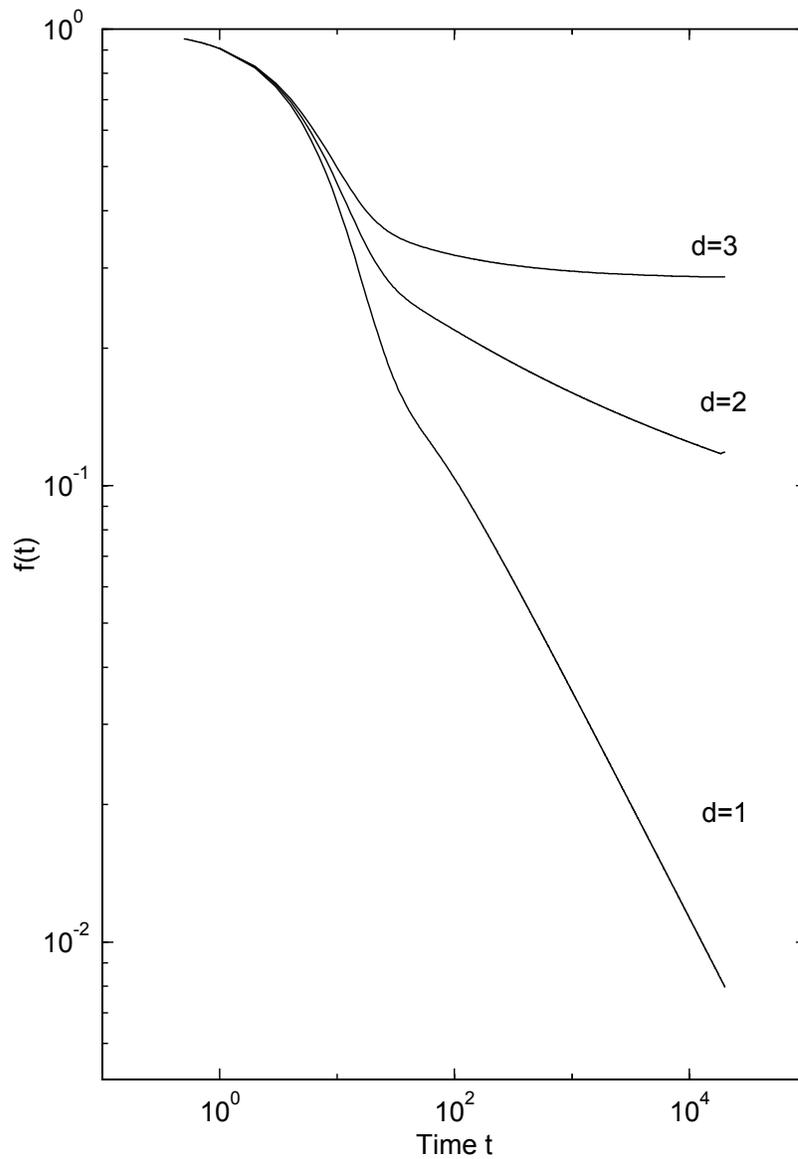


Figure 4.2: Numerical solution for the heterozygosity $h(t) = 1 - F_{ST}(t)$ showing the dynamics the fixation index (two-point correlation function). Note the slow logarithmic decay for $d = 2$. For $d = 1$, the asymptotic behaviour is clearly $t^{-1/2}$.

giving us

$$\begin{aligned} I(\mathbf{r}_1, \mathbf{r}_2) &= \frac{1}{2} + \frac{1}{2} \langle c_{\mathbf{r}_1} c_{\mathbf{r}_2} \rangle \\ &= \frac{1}{2} + \frac{1}{2} G_{\phi\phi}(\mathbf{r}_1, \mathbf{r}_2). \end{aligned} \quad (4.45)$$

Note that $I(\mathbf{r}_1, \mathbf{r}_2) \rightarrow 1$ iff $G_{\phi\phi}(\mathbf{r}_1, \mathbf{r}_2) \rightarrow 1$. Furthermore we see that the initial conditions used to solve Eqn. (4.35) which were $f_{\mathbf{r}} = 1/2$ for all \mathbf{r} , imply $G_{\phi\phi}(\mathbf{r}_1, \mathbf{r}_2; 0) = 0$ and $I(\mathbf{r}_1, \mathbf{r}_2; 0) = 1/2$, in agreement with Eqn. (4.45). Therefore, in the two-allele model studied in this thesis, $I(\mathbf{r}_1, \mathbf{r}_2; t)$ and $G_{\phi\phi}(\mathbf{r}_1, \mathbf{r}_2; t)$ have the same asymptotics. It is quite interesting that $I(\mathbf{r}_1, \mathbf{r}_2; t)$ in the infinite-allele model studied by Sawyer has the same asymptotics as for the two-allele model. One is tempted to conjecture that $I(\mathbf{r}_1, \mathbf{r}_2; t)$ has the same asymptotics in models with *arbitrary* number of alleles. An extension of the methods described in this thesis could likely answer this question.

Sawyer also found that $I(\mathbf{r}_1, \mathbf{r}_2; t)$ for any two individuals in any bounded set was asymptotically equal to unity iff the random walk performed by the individuals was recurrent. This result is also very similar to the results of this thesis for the asymptotic value of F_{ST} and the existence of broken symmetry. Sawyer's result, however, is obtained by a very different method than that described in the previous subsection and pertains only to infinite alleles model.

Further analysis of the infinite alleles model was carried out by Sawyer[60] who provided a limit theorem for patch sizes and found that the expectation of $N(t)$, the average number of individuals of the same type scales as $t^{1/2}$, $t/\ln t$ and t in $d = 1, 2$ and 3 , respectively.

4.5 Alternative Formulations of the Stepping-Stone Model.

In section 4.1 I formulated the Stepping-Stone model using the occupation number representation. Here I show that one can use either the response functional formalism or

the operator form of the Fokker-Planck equation to achieve an equivalent description. Moreover, the starting point for applying these techniques is the Langevin equation introduced by Kimura and Weiss [1][2] in their original formulation of the Stepping-Stone model. I also point out that all of these reformulations of the Stepping-Stone model lead to the same path-integral and its associated action. I use this path-integral exclusively in the section 4.6 and chapter 5.

To remind the reader, the Stepping-Stone model of Kimura and Weiss is spatial model where colonies (sub-populations) are arranged on a hypercubic lattice and are coupled by random migration of individuals. Apart from random migration between colonies, each sub-population evolves according to the Wright-Fisher model. We denote by $f_{\mathbf{r}}$, the frequency of the allele with a selective disadvantage, in the colony located at position \mathbf{r} . The Langevin equation reads

$$\frac{\partial f_{\mathbf{r}}(t)}{\partial t} = m \sum_{\mathbf{e}} (f_{\mathbf{r}+\mathbf{e}} + f_{\mathbf{r}-\mathbf{e}} - 2f_{\mathbf{r}}) - \mu(2f_{\mathbf{r}} - 1) - s\tau_{\mathbf{g}}^{-1}f_{\mathbf{r}}(1 - f_{\mathbf{r}}) \quad (4.46)$$

$$+ \sqrt{\frac{f_{\mathbf{r}}(1 - f_{\mathbf{r}})}{n\tau_{\mathbf{g}}}} \eta_{\mathbf{r}}(t), \quad (4.47)$$

where m is the migration rate, μ is the mutation rate, n is the total number at each site, and s is the coefficient of selection. The noise term is Gaussian and specified by the auto-correlation

$$\langle \eta_{\mathbf{r}}(t) \eta_{\mathbf{r}'}(t') \rangle = \delta_{\mathbf{r},\mathbf{r}'} \delta(t - t').$$

The amplitude of the noise in Eq. (4.46) is the result of Wright-Fisher binomial sampling at each site, and as was shown in section 3.2, is the same for Moran's dynamics. Finally, Eq. (4.46) is to be interpreted with Itô's prescription.

4.5.1 The Fokker-Planck equation and operators.

The Fokker-Planck equivalent to Eq. (4.46) is found in the standard way,

$$\begin{aligned} \frac{\partial W[f_{\mathbf{r}}]}{\partial t} &= \sum_{\mathbf{r}} \frac{\partial}{\partial f_{\mathbf{r}}} [m\Delta_{\mathbf{r}} + \mu(2f_{\mathbf{r}} - 1) + s\tau_{\mathbf{g}}^{-1}f_{\mathbf{r}}(1 - f_{\mathbf{r}})] W[f_{\mathbf{r}}] \\ &\quad + \frac{1}{2n\tau_{\mathbf{g}}} \sum_{\mathbf{r}} \frac{\partial^2}{\partial f_{\mathbf{r}}^2} f_{\mathbf{r}}(1 - f_{\mathbf{r}}) W[f_{\mathbf{r}}], \end{aligned} \quad (4.48)$$

where we have defined the lattice Laplacian $\Delta_{\mathbf{r}}$,

$$\Delta_{\mathbf{r}}f_{\mathbf{r}} = \sum_{\mathbf{e}} (f_{\mathbf{r}+\mathbf{e}} + f_{\mathbf{r}-\mathbf{e}} - 2f_{\mathbf{r}}).$$

For the operator description, we define the operator $\hat{f}_{\mathbf{r}}$ as multiplication by $f_{\mathbf{r}}$, and the operator $\hat{g}_{\mathbf{r}}$, as the partial derivative $\frac{\partial}{\partial f_{\mathbf{r}}}$. This gives us the Liouville operator

$$\begin{aligned} \mathcal{L} &= \sum_{\mathbf{r}} \hat{g}_{\mathbf{r}} [m\Delta_{\mathbf{r}}\hat{f}_{\mathbf{r}} + 2\mu\hat{f}_{\mathbf{r}} - 1] + s\tau_{\mathbf{g}}^{-1} \sum_{\mathbf{r}} \hat{g}_{\mathbf{r}}\hat{f}_{\mathbf{r}} (1 - \hat{f}_{\mathbf{r}}) \\ &\quad + \frac{1}{2n\tau_{\mathbf{g}}} \sum_{\mathbf{r}} \hat{g}_{\mathbf{r}}^2\hat{f}_{\mathbf{r}} (1 - \hat{f}_{\mathbf{r}}), \end{aligned} \quad (4.49)$$

Equation (4.49) can also be obtained from $\mathcal{L}_{\text{ECSS}}$ in Eq. (4.3) by performing a canonical transformation to eliminate the $b_{\mathbf{r}}$ and $b_{\mathbf{r}}^{\dagger}$ operators, followed by the system size expansion for Moran's model discussed in section 3.2.2, applied to each site. As discussed there and in chapter 2, this is accomplished by identifying appropriately scaled creation and annihilation operators with Fokker-Planck operators, which for the present discussion are the $\hat{f}_{\mathbf{r}}$ and $\hat{g}_{\mathbf{r}}$ operators.

4.5.2 The path-integral.

One can easily proceed from Fokker-Planck equation to a path-integral description as discussed in chapter 2. The path-integral can also describe the lattice model, but here

we take the continuum limit. Using the fields $f_{\mathbf{r}}$ and $\bar{f}_{\mathbf{r}}$, that correspond, respectively, to the operators $\hat{f}_{\mathbf{r}}$ and $\hat{g}_{\mathbf{r}}$, the action is given by

$$S[\bar{f}_{\mathbf{r}}, f_{\mathbf{r}}] = \int d^d \mathbf{r} dt [\bar{f}_{\mathbf{r}} \partial_t f_{\mathbf{r}} + \bar{f}_{\mathbf{r}} (D \nabla^2 + 2\mu f_{\mathbf{r}} - 1) + s\tau_g^{-1} \bar{f}_{\mathbf{r}} f_{\mathbf{r}} (1 - f_{\mathbf{r}})] \quad (4.50)$$

$$+ \frac{1}{2n\tau_g} \int d^d \mathbf{r} dt [\bar{f}_{\mathbf{r}}^2 f_{\mathbf{r}} (1 - f_{\mathbf{r}})],$$

where for simplicity we have dropped the explicit time-dependence of the fields. The continuum limit gives us the diffusion constant $D = ml^2$, where l is the lattice spacing.

I use the path-integral description in section 4.6 below and for the renormalization group analysis in chapter 5. I should also point out that this path-integral is completely equivalent to that obtained using the coherent-state representation of the creation and annihilation operators. This procedure is also discussed in chapter 2.

4.5.3 The response functional method.

We can use the response functional formalism to go directly from the Langevin Eq. (4.46) to the path integral, without making contact with operators. One follows the same procedure as the example in chapter 2.4.1. I will not show all the steps since terms in the action that come from the deterministic part of Eq. (4.46) are produced in exactly the same way as the example of chapter 2.4.1. I only highlight the difference from that example here, which is the form of the noise term. In our example of 2.4.1, the noise term was the constant, with variance Γ . Performing the functional integral over the noise field $\eta_{\mathbf{r}}(t)$ required the Gaussian integral

$$\int \exp\left(\eta_{\mathbf{r}} \bar{\phi}_{\mathbf{r}} - \frac{\eta_{\mathbf{r}}^2}{2\Gamma}\right) d\eta_{\mathbf{r}} \propto \exp\left(\frac{1}{2}\Gamma \bar{\phi}_{\mathbf{r}}^2\right), \quad (4.51)$$

and therefore generated the $\bar{\phi}_{\mathbf{r}}^2$ term in the Lagrangian density. For the Stepping-Stone model defined in Eq. (4.46),

$$\Gamma = \Gamma[f_{\mathbf{r}}] = \frac{f_{\mathbf{r}}(1 - f_{\mathbf{r}})}{n\tau_g},$$

and the integral over $\eta_{\mathbf{r}}$ becomes

$$\int \exp\left(\eta_{\mathbf{r}}\bar{f}_{\mathbf{r}} - \frac{\eta_{\mathbf{r}}^2}{2\Gamma[f_{\mathbf{r}}]}\right) d\eta_{\mathbf{r}} \propto \exp\left(\frac{1}{2}\bar{f}_{\mathbf{r}}^2\Gamma[f_{\mathbf{r}}]\right),$$

which adds the term

$$S'[\bar{f}_{\mathbf{r}}, f_{\mathbf{r}}] = \frac{1}{2n\tau_{\mathbf{g}}} \int d^d\mathbf{r}dt [\bar{f}_{\mathbf{r}}^2 f_{\mathbf{r}} (1 - f_{\mathbf{r}})]$$

to the action. The full action is identical to that given in Eq. (4.50).

4.5.4 The neutral regime.

For the neutral regime, $s = 0$. We can take either the Fokker-Planck Equation (4.48), or the Langevin Equation (4.46) and make the change of variables to the difference field,

$$\phi_{\mathbf{r}} = 2f_{\mathbf{f}} - 1.$$

This produces the Fokker-Planck equation,

$$\begin{aligned} \frac{\partial W[\phi_{\mathbf{r}}]}{\partial t} &= \sum_{\mathbf{r}} \frac{\partial}{\partial \phi_{\mathbf{r}}} \left(m \sum_{\mathbf{e}} (\phi_{\mathbf{r}+\mathbf{e}} + \phi_{\mathbf{r}-\mathbf{e}} - 2\phi_{\mathbf{r}}) + 2\mu\phi_{\mathbf{r}} \right) W[f_{\mathbf{r}}] \\ &\quad + \frac{1}{2n\tau_{\mathbf{g}}} \sum_{\mathbf{r}} \frac{\partial^2}{\partial \phi_{\mathbf{r}}^2} (1 - \phi_{\mathbf{r}}^2) W[\phi_{\mathbf{r}}]. \end{aligned}$$

For the operator description of the Fokker-Planck equation, we define multiplication by $\phi_{\mathbf{r}}$ as the operator $\hat{\phi}_{\mathbf{r}}$, and $\frac{\partial}{\partial \phi_{\mathbf{r}}}$, as the operator $\hat{\psi}_{\mathbf{r}}$. The Liouville operator is

$$\mathcal{L} = m \sum_{\mathbf{r}} \hat{\psi}_{\mathbf{r}} \left(\Delta_{\mathbf{r}} + 2\mu\hat{\phi}_{\mathbf{r}} \right) + \frac{1}{2n\tau_{\mathbf{g}}} \sum_{\mathbf{r}} \hat{\psi}_{\mathbf{r}}^2 (1 - \phi_{\mathbf{r}}^2), \quad (4.52)$$

and by now we should not be surprised that this operator is completely equivalent to $\mathcal{L}_{\text{ECSS}}$ in Eq. (4.8), if we identify the scaled creation-annihilation operators c and c^\dagger with $\hat{\phi}_{\mathbf{r}}$ and $\hat{\psi}_{\mathbf{r}}$.

4.6 The Broken-Symmetry State: Fate of a single mutant gene.

In this section, we will assume that the system has evolved into the broken symmetry state corresponding to fixation of a particular allele, along the lines of our discussion in section 4 of chapter 2. In that section, we found that the non-spatial model, in the thermodynamic limit, was equivalent to a simple branching process. Were we to write the non-spatial model using the creation-annihilation formalism, we would see that the terms of $\mathcal{O}(N^{-1})$ in the master equation, Eq. (3.25), that were safely ignored as $N \rightarrow \infty$, correspond to interactions between the otherwise independently branching particles. In the spatial model of this chapter, these terms are $\mathcal{O}(n^{-1})$, where n is the local density (number per site), and hence cannot be ignored in the thermodynamic limit. We are left with a type of interacting branching process, of which an exact solution appears very difficult. Fortunately, we will be able to find an exact result for the relaxation time τ . This quantity is the time scale over which long wavelength perturbations relax and is an important quantity to calculate since it pertains to any avalanche or cascade, induced by a single new mutant, that lasts long enough to spread throughout the population. This is relevant since we are ultimately interested in whether or not this cascade of new alleles significantly changes the composition of the population or indeed approaches fixation. In this section, we will find that τ diverges as $s \rightarrow 0$, and the exponent describing this divergence differs from its mean-field counterpart below $d = 2$.

As discussed in the preceding section, we can describe the model using several different formalisms. We will need the path-integral description for this section, however. In the previous section we described how one could accomplish this starting from the Stepping-Stone Langevin equation. We also pointed out the equivalence of this procedure with the same derivation of the path-integral from the occupation number representation. We remind the reader of this latter technique. The details can be found in the example of

section 2.4.2.

To get the path integral using the occupation number representation, one starts with the Liouville operator of Eq. (4.8) for the ECSS model. We then use canonical transformations of the kind discussed in section 3.2.2 for Moran's model to eliminate one set of operators. For this section, we will eliminate the operators corresponding to the fixed allele. Assuming a homogeneously distributed population, as was done for the neutral theory, produces identical results for the DCSS model, in that terms that couple the total field $d_{\mathbf{r}} = a_{\mathbf{r}} + b_{\mathbf{r}}$ and the $a_{\mathbf{r}}$ field we are interested in, do not affect the calculation of the relaxation time below. After Fourier transformation, we are left with the Liouville operator

$$\begin{aligned} \mathcal{L} = & - \sum_{\mathbf{k}} (m\omega_{\mathbf{k}} + s) a_{\mathbf{k}}^{\dagger} a_{\mathbf{k}} + sN^{-1} \sum_{\mathbf{k}, \mathbf{q}} a_{\mathbf{k}}^{\dagger} a_{\mathbf{k}-\mathbf{q}} a_{\mathbf{k}+\mathbf{q}} \\ & + \sum_{\mathbf{k}, \mathbf{q}} a_{\mathbf{k}+\mathbf{q}}^{\dagger} a_{\mathbf{k}-\mathbf{q}}^{\dagger} a_{\mathbf{k}} - (2nN)^{-1} \sum_{\mathbf{k}_1, \mathbf{k}_2, \mathbf{q}} a_{\mathbf{k}_1}^{\dagger} a_{\mathbf{k}_2}^{\dagger} a_{\mathbf{k}_1-\mathbf{q}} a_{\mathbf{k}_2+\mathbf{q}}, \end{aligned}$$

where time is in units of the generation time τ_g , (so $\tau_g = 1$). For a selectively advantageous allele, $s > 0$, and for a disadvantageous allele, $s < 0$. The a and a^{\dagger} operators here are scaled creation and annihilation operators and therefore represent the allele frequency. To facilitate the calculation, I take the continuum limit of \mathcal{L} , and replace the wave-vector sums by integrals. This will in no way change the results. Using the standard procedure in field theory discussed in section 2.4.2, I pass from the Liouville operator, *via* the coherent-state representation, to the path-integral.

To remind the reader, the path-integral is written as

$$Z = \int [d\bar{\phi}] [d\phi] \exp(-S[\bar{\phi}, \phi]),$$

where the “action” S reads,

$$S[\bar{\phi}, \phi] = \int dt \int d^d \mathbf{r} \left(\bar{\phi}_{\mathbf{r}}(t) \partial_t \phi(t)_{\mathbf{r}} - L[\bar{\phi}_{\mathbf{r}}(t), \phi_{\mathbf{r}}(t)] \right),$$

and where $L[\bar{\phi}_{\mathbf{r}}(t), \phi_{\mathbf{r}}(t)]$ obtained from $\mathcal{L}[a_{\mathbf{r}}^\dagger, a_{\mathbf{r}}]$ by replacing the $a_{\mathbf{r}}^\dagger$ and $a_{\mathbf{r}}$ operators in \mathcal{L} with the fields $\bar{\phi}_{\mathbf{r}}(t)$ and $\phi_{\mathbf{r}}(t)$, respectively. The quantity under the action integral, $\bar{\phi}_{\mathbf{r}}(t) \partial_t \phi(t)_{\mathbf{r}} - L[\bar{\phi}_{\mathbf{r}}(t), \phi_{\mathbf{r}}(t)]$, is called the Lagrangian density. As an alternative, we could stick with our Liouville operator, and use the formalism of “time-ordered” operators [44], but we would obtain the same results. Finally, I perform an additional Fourier transformation from time t to frequency ω . This produces the action

$$\begin{aligned} S[\bar{\phi}, \phi] &= \int \bar{d}\omega \int^\Lambda \bar{d}^d \mathbf{k} \quad (i\omega + D\mathbf{k}^2 + s) \bar{\phi}_{\mathbf{k},\omega} \phi_{\mathbf{k},\omega} & (4.53) \\ &+ s \int \bar{d}\omega \bar{d}\varepsilon \int^\Lambda \bar{d}^d \mathbf{k} \bar{d}^d \mathbf{q} \quad \bar{\phi}_{\mathbf{k},\omega} \phi_{\mathbf{k}-\mathbf{q},\omega-\varepsilon} \phi_{\mathbf{q},\varepsilon} \\ &+ (2n)^{-1} \int \bar{d}\omega \bar{d}\varepsilon \int^\Lambda \bar{d}^d \mathbf{k} \bar{d}^d \mathbf{q} \quad \bar{\phi}_{\mathbf{k},\omega} \bar{\phi}_{\mathbf{k}-\mathbf{q},\omega-\varepsilon} \phi_{\mathbf{q},\varepsilon} \\ &- (2n)^{-1} \int \bar{d}\omega_1 \bar{d}\omega_2 \bar{d}\varepsilon \bar{d}^d \mathbf{k} \bar{d}^d \mathbf{k}_2 \bar{d}^d \mathbf{q} \bar{\phi}_{\mathbf{k}_1,\omega_1} \bar{\phi}_{\mathbf{k}_2,\omega_2} \phi_{\mathbf{k}_1+\mathbf{q},\omega_1+\varepsilon} \phi_{\mathbf{k}_2-\mathbf{q},\omega_2-\varepsilon}. \end{aligned}$$

Here, and in the rest of this thesis, we’ll use $\bar{d} \equiv \frac{d}{2\pi}$, so $\bar{d}^d \mathbf{k} = \frac{d^d \mathbf{k}}{(2\pi)^d}$. A graphical representation of $S[\bar{\phi}, \phi]$ in terms of edges and vertices is given in Figure 4.3. Each vertex represents a term in the action. Lines with outgoing arrows represent $\bar{\phi}_{\mathbf{r}}$ (creation) fields, those with incoming lines represent $\phi_{\mathbf{r}}$ (annihilation) fields. The single edge (line) without a vertex represents the $\bar{\phi}_{\mathbf{r}} \phi_{\mathbf{r}}$ term. Except for this latter term, the label under each vertex gives the coupling constant that multiplies the respective term in the action. The factor $(i\omega + Dk^2 + s)^{-1}$ under the single edge is the two-point propagator of the “free theory”, which comes from the solution of the path-integral for the theory without the interaction terms. In the next section, I give a short description of how these diagrams assist us in the calculations that follow.

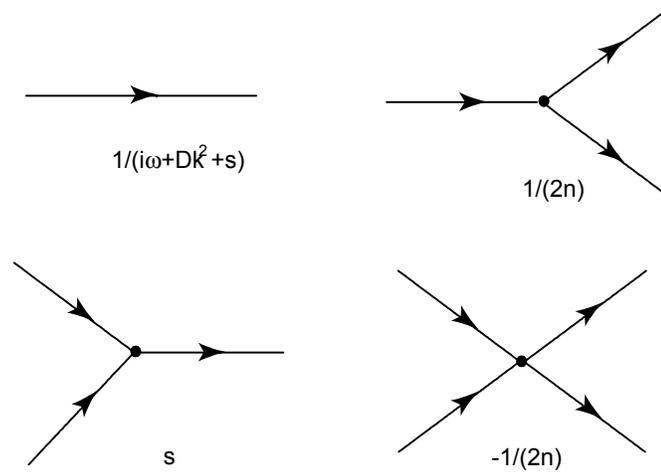


Figure 4.3: Vertices representing the action for propagation of new mutant alleles. The propagator of the non-interacting theory represented by the line in the upper left. The vertices representing the interactions are labeled with their respective coupling constants.

4.6.1 Diagrammatic Methods in Field Theory.

We are unable to solve the full path-integral for the action in Eq. (4.53) or alternatively Figure (4.3). At the end of section 2.4.1 of chapter 2, I mentioned that the procedure for dealing with such “interacting” systems was to break the action up into two parts: a Gaussian part S_0 , that is bilinear in the $\bar{\phi}$ and ϕ fields, and an interacting part S_I containing the higher order terms. A perturbation expansion of a given response function can then be carried out using the propagator from the action S_0 . I will not describe the technical details of how one generates such a perturbation expansion here. The reader is encouraged to consult one of the many texts on quantum field theory or the references cited in this thesis [38][61][37][33][46][61]. Here I give an account of the evaluation of such a perturbation expansion using diagrams generated from the vertices from Figure (4.3).

A response function is defined by the expectation

$$G_{n,m} = \langle \bar{\phi}^n \phi^m \rangle,$$

where for convenience, we have dropped the indices for wavevectors $\mathbf{k}_1, \dots, \mathbf{k}_{m+n}$ and frequencies $\omega_1, \dots, \omega_{m+n}$ (or alternative positions and times). The perturbation expansion for $G_{n,m}$ using the “bare” action S_0 will contain a countable infinite number of terms, each of which can be represented by a Feynman diagram constructed using the vertices and propagator line of Figure (4.3). A term in the perturbation expansion for $G_{n,m}$ is constructed from these vertices by connecting the outgoing lines of one vertex with the incoming lines of another vertex in such a way that the total number of incoming lines is equal to m and the total number of outgoing lines is equal to n . The “Feynman rules” are then applied to convert the diagram into an expression. For example, the diagram in Figure (4.4) comes from combining the vertices in the upper right and lower left of Figure (4.3). This diagram has one incoming line and one outgoing line and therefore represents a correction to the two-point propagator. The Feynman rules consist of assigning a

wavevector \mathbf{q} and frequency ω to each line which then brings a factor of $(i\omega + Dq^2 + s)^{-1}$, the propagator of the S_0 action, into the final expression for the perturbation term. Each vertex in the diagram contributes a factor equal to its coupling constant to this expression, *i.e.* the numbers under the vertices in Figure (4.3). Finally, all internal wave-vectors and frequencies are integrated and wave-vector and frequency are conserved at each vertex. Internal variables are those not associated with the m incoming or n outgoing lines; internal lines have both ends connected to a vertex. Conservation of wave-vector at each vertex means that the sum over the wave-vectors of the incoming lines at that vertex must equal the sum of the wave-vectors of the outgoing lines for that vertex.

As an example, let us compute the correction to the two-point propagator represented by the diagram in Figure (4.4). The diagram gives us the following integral,

$$I = -\frac{1}{4n^2} \frac{1}{(i\omega + Dk^2 + s)(i\omega + Dk^2 + s)} \quad (4.54)$$

$$\times \int d\varepsilon \int^{\Lambda} d^d \mathbf{q} \frac{1}{i(\omega - \varepsilon) + D(\mathbf{k} - \mathbf{q})^2 + s} \frac{1}{i(\omega + \varepsilon) + D(\mathbf{k} + \mathbf{q})^2 + s}.$$

The evaluation of this integral gives us the first-order correction to the full propagator, however, we will not use this result. Instead, we will use a similar perturbation expansion for Σ , the inverse of the relaxation time, in the next section.

4.6.2 Perturbation theory for the relaxation rate Σ .

At the beginning of section 4.2, we pointed out that none of the interaction terms in the Liouville operator produced corrections to the mutation rate μ . Unlike the situation with the mutation rate μ , however, the interaction terms in Eq. (4.53) do produce corrections or modifications to the two-point propagator. The situation is depicted using Feynman diagrams in Figure (4.5). We see that the contribution coming from diagrams with only

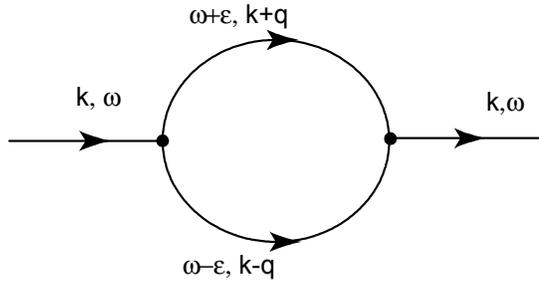


Figure 4.4: Feynman diagram representing the 1-loop correction to the two-point propagator

one “annihilation” vertex (the vertex in the bottom left of Figure (4.3)) are $\mathcal{O}(s)$, while the contributions from those with m “annihilation” vertices are $\mathcal{O}(s^m)$. For example, the bottom diagram in Figure (4.5) contributes a term $\mathcal{O}(s^2)$. Along the same lines, a perturbation expansion for Σ , the inverse of the relaxation time, is given in Figure (4.6 a). As we are interested in behavior of the relaxation time for the limit of $s \rightarrow 0$, we might be tempted to use the expansion in a) which uses the bare s in all of propagators of the loop integrals. This expansion leaves out diagrams of $\mathcal{O}(s^2)$ and higher and therefore is of questionable validity in the desired limit of $s \rightarrow 0$. We can remedy this however, by using the expansion in Figure (4.6 b), where we have substituted Σ back into the propagators to get a self-consistent equation, formally called a “Dyson” equation for Σ . This procedure has the effect of including all the diagrams we have left out. In the limit of $s \rightarrow 0$, the equation for Σ is exact. If the self-consistent equation can be solved, we’re in luck and can extract the asymptotic divergence of $\tau = \Sigma^{-1}$ with s .

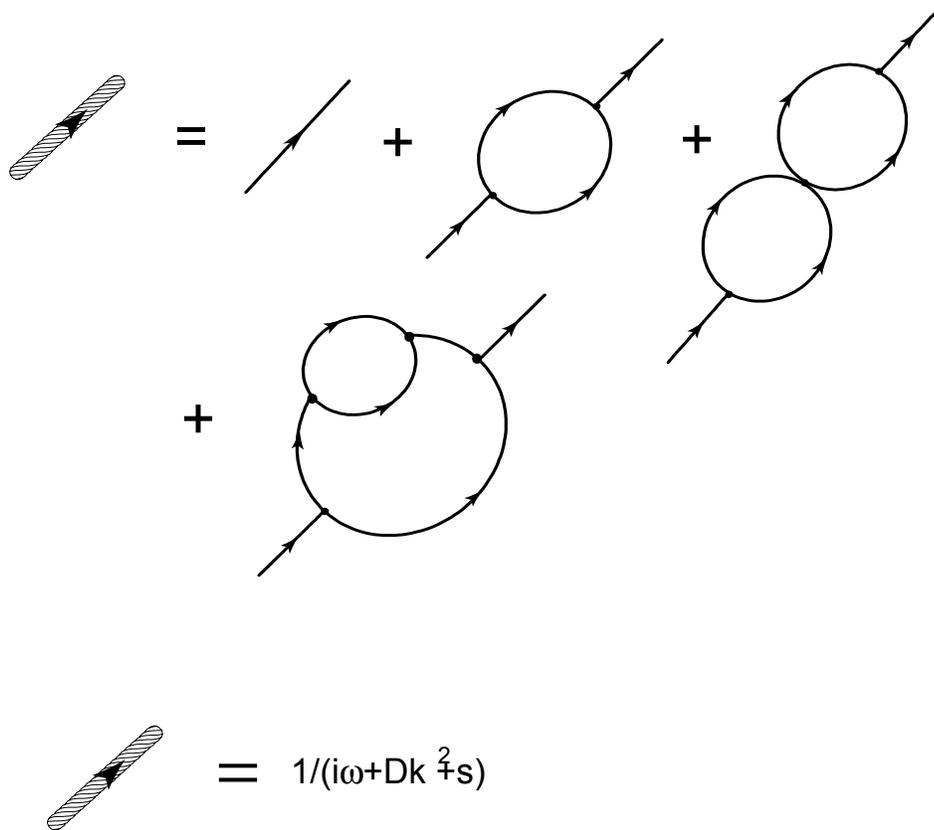


Figure 4.5: Perturbation expansion for the two-point propagator. The upper diagrams produce corrections of $\mathcal{O}(s)$ to the relaxation rate, while the lower diagrams contributes a correction of $\mathcal{O}(s^2)$.

a)

$$\Sigma = \bullet + \text{circle} + \text{two circles} + \dots$$

b)

$$\Sigma = \bullet + \text{hatched circle} + \text{two hatched circles} + \dots$$


 $= 1/(i\omega + Dk^2 + \Sigma)$

Figure 4.6: Perturbation expansion and Dyson equation for Σ . a) perturbation expansion to $\mathcal{O}(s)$ using the bare propagator. b) A self-consistent Dyson equation is derived using the full propagator.

4.6.3 The effective relaxation rate and “screening” of selection in low dimensions.

Using the diagrammatic expansion in Figure (4.6b), the Dyson equation for the inverse relaxation time $\Sigma = \tau^{-1}$, results in an alternating geometric series

$$\begin{aligned}\Sigma &= s(1 - I + I^2 - \dots) \\ &= \frac{s}{1 + I},\end{aligned}$$

with

$$I = \frac{1}{2n} \int_0^\infty \bar{d}\omega \int^\Lambda \bar{d}^d \mathbf{q} \frac{1}{\omega^2 + (Dq^2 + \Sigma)^2}.$$

Here again, we'll use $\bar{d} \equiv \frac{d}{2\pi}$, so $\bar{d}^d \mathbf{q} = \frac{d^d \mathbf{q}}{(2\pi)^d}$. The integral I is given by

$$I = \frac{1}{2n} \int^\Lambda \bar{d}^d \mathbf{q} \frac{1}{Dq^2 + \Sigma},$$

which we have seen before in Eq. (4.29), and evaluates to

$$I = \begin{cases} \frac{1}{4n\sqrt{2\Sigma D}}, & d = 1 \\ \frac{1}{8\pi n D} \ln\left(\frac{\Lambda^2 D}{\Sigma} + 1\right), & d = 2 \\ \frac{\Lambda}{4\pi^2 n D}, & d = 3 \end{cases}$$

We now proceed to solve the Dyson equation for $d = 1, 2$ and 3 .

For $d = 1$, we have

$$\Sigma = \frac{s}{1 + (32n^2 D \Sigma)^{-1/2}} \quad (4.55)$$

for $\Sigma \rightarrow 0$ this gives us

$$\Sigma = \sqrt{32n^2 s^2 \Sigma D}$$

Hence the relaxation time diverges as

$$\tau \sim \frac{1}{32Dn^2 s^2}$$

This is to be contrasted with the mean-field result in which $\tau \sim s^{-1}$.

For $d = 2$, the Dyson equation reads

$$\Sigma = \frac{s}{1 + \frac{1}{8\pi D n} \ln\left(\frac{\Lambda^2 D}{\Sigma}\right)}$$

The solution is

$$\Sigma = -8Ds\pi \frac{n}{W_{-1}\left(-8s\pi \frac{n}{\Lambda^2} e^{-8\pi D n}\right)},$$

where $W_k(x)$ is the k^{th} branch of Lambert's W-function, the inverse of

$$f(W) = W e^W$$

For small x , $W_{-1}(x) \simeq \ln x$, so

$$\Sigma = \frac{s}{\ln\left(\frac{\Lambda^2}{8s\pi n}\right) / 8Dn\pi + 1}$$

and we find that apart from logarithmic corrections, Σ is linear in s .

Next we consider $d = 3$. Here we simply have

$$\Sigma = \frac{s}{1 + \frac{\Lambda}{4\pi^2 n D}}.$$

Note that the integral I in this case is independent of Σ , so this result is simply what we would have gotten using the naive perturbation approach of Figure (4.6a). The complete solution of the Dyson equation for $d = 1, 2$ and 3 is plotted in Figure (4.7). For $d = 1$, the complete solution, as opposed to the result for small s is

$$\Sigma = s + \frac{2 - \sqrt{4 + 512sn^2D}}{128n^2D}$$

Note that for larger s , $\Sigma \sim s$, *i.e.* it approaches the mean-field behaviour. The crossover occurs roughly at $s^* \simeq (64n^2D)^{-1}$.

One can think of the relaxation time τ , or even the selection coefficient as being “renormalized” due to the spatial fluctuations. This is particularly acute for $d = 1$, as the present results demonstrate.

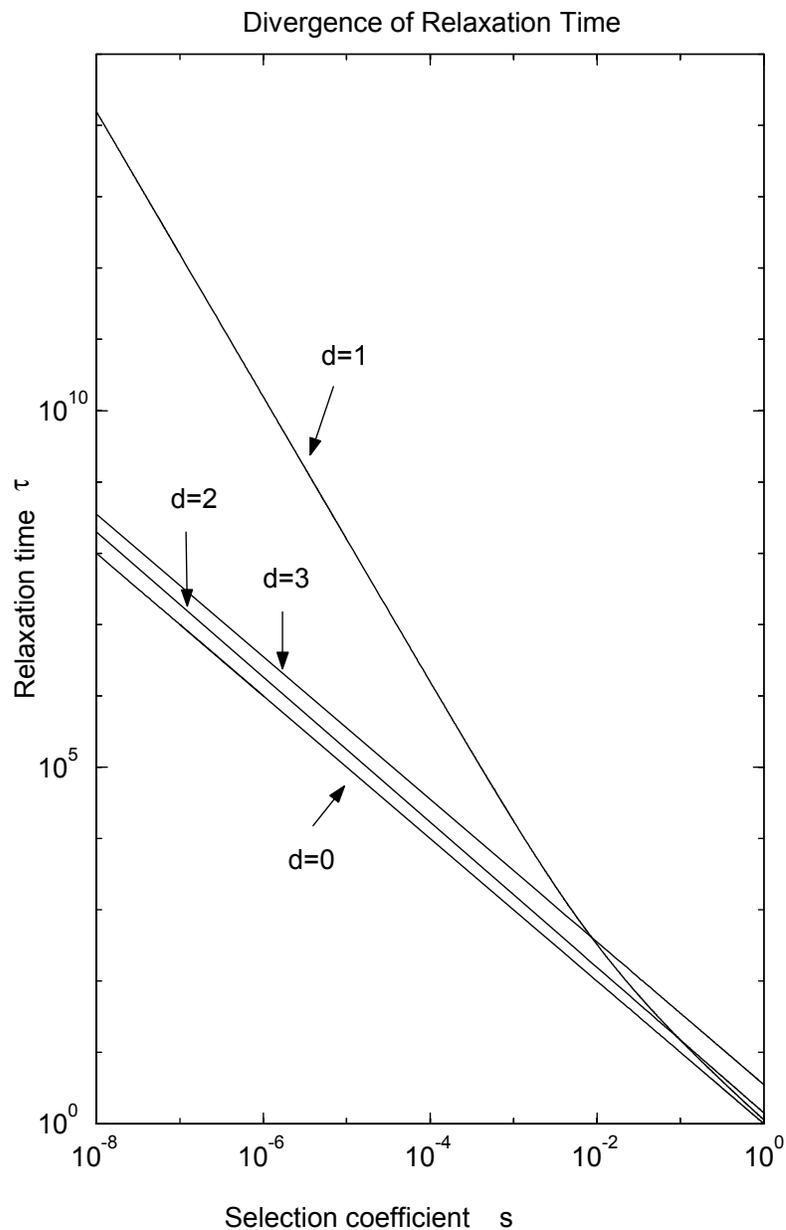


Figure 4.7: Divergence of the relaxation time for $d = 1, 2$ and 3 . For $d = 1$, the relaxation time diverges as s^{-2} , in contrast to s^{-1} in higher dimensions. We also see for $d = 2$, the slight logarithmic departure from s^{-1} .

4.6.4 Survival probabilities: scaling arguments.

In the previous section, we found the exponent ν that describes the divergence of the characteristic relaxation time as the selection coefficient $s \rightarrow 0$, *i.e.* $\tau \sim s^{-\nu}$. We would like to know the power law behaviour of other quantities such as the ultimate survival probability of a mutation-induced “avalanche” as a function of s or its time-dependent decay. In fact, a substantial effort was put forth by the author to calculate these quantities directly *via* several methods: generalization of Kimura’s approach [52] to the non-spatial model using the backward equation, closing of the moment equations for n -particle operators and direct summation of Feynman diagrams for small s . Unfortunately, none of these approaches were successful, which perhaps is a testament to the difficulty of understanding such non-conservative, non-equilibrium models—even the “stepping-stone” discussed in this thesis, which as we have shown, admits some exact results. Fortunately, we can adapt scaling arguments developed for directed percolation (DP) and other growth models to relate the exponents we know from our exact results to the otherwise unknown ones.

Here we use scaling arguments put forth by Dickman and Tretyakov [62] for the Domany-Kinzel cellular automaton (DKCA) [63]. The phase diagram of the DKCA possesses a line of critical points that separates an active phase from the absorbing vacuum state. The critical behaviour along this line is that of DP, except at one terminal point where the asymptotic behaviour is known exactly for $d = 1$ and is identical to that of the Voter model. Spreading of a critical process from a localized source is described by a power laws whose exponents are generally connected by two scaling relations [64]. One of these, known as “hyperscaling”, at first sight fails for the DKCA at the terminal critical point corresponding to the Voter model. Dickman and Tretyakov realized that this apparent failure of hyperscaling is due the fact that for the Voter model—and as

we have discussed in the above chapters of this thesis, the “stepping-stone” model—the order parameter is discontinuous at the transition. When this is taken into account, a different hyperscaling relation is obtained.

Spreading of a critical process, localized at the origin is described by a set of exponents defined by

$$P(t) \propto t^{-\delta}, \quad (4.56)$$

$$n(t) \propto t^\eta, \quad (4.57)$$

and

$$\mathbf{R}^2(t) \propto t^z, \quad (4.58)$$

where the survival probability $P(t)$ is simply $1 - P_0(t)$, the probability of being in the 0-particle state w.r.t. new alleles and the power law behaviour of Eq. (4.56) is valid at long times. In Eq. (4.57), $n(t)$ is the mean population size of new alleles and $\mathbf{R}^2(t)$ in Eq. (4.58), the mean-squared displacement of the new particles. The ultimate survival probability is characterized by the exponent β ,

$$P_\infty = \lim_{t \rightarrow \infty} P(t) \propto s^\beta. \quad (4.59)$$

The scaling hypothesis [64] for spreading from an initial perturbation (in this case, a single new mutant at $\mathbf{r} = \mathbf{0}$, $t = 0$) postulates the existence of two scaling functions G and ϕ defined by

$$\rho(\mathbf{r}, t) \sim t^{\eta-dz/2} G(\mathbf{r}^2/t^z, st^{1/\nu}), \quad (4.60)$$

and

$$P_S(t) \sim t^{-\delta} \phi(t/\tau) = t^{-\delta} \phi(ts^\nu).$$

Here $\rho(\mathbf{r}, t)$ is the local order-parameter density, *i.e.* the density of the new alleles. The form of Eq. (4.60) follows by noting that for $s = 0$, integrating this equation over all space produces Eq. (4.57).

The existence of the limit in Eq. (4.59) implies

$$\phi(x) \sim x^{\beta/\nu} = t^{\beta/\nu} s^\beta$$

and hence, in order to cancel the t -dependence in Eq. (4.59),

$$\beta = \delta\nu. \quad (4.61)$$

In the case of a selectively advantageous new mutation that reaches fixation, the local density, for fixed \mathbf{r} , must approach the stationary density ρ_0 which for our model is simply 1. Since $\rho(\mathbf{r}, t)$ is the average over *all* trials, for small but finite s we have,

$$\lim_{t \rightarrow \infty} \rho(\mathbf{r}, t) \sim s^\beta \rho_0. \quad (4.62)$$

It follows that $G(\mathbf{0}, y) \sim y^\beta$ for large y . Then in order to cancel the t -dependence in Eq. (4.60), we must have $\beta/\nu = \frac{dz}{2} - \eta$, or using the first scaling relation above

$$\delta + \eta = \frac{dz}{2}. \quad (4.63)$$

This ‘‘hyperscaling’’ relation depend crucially upon the validity of our assumption that even in the limit that $s \rightarrow 0$, the order parameter is discontinuous, or in other words, that the symmetry is spontaneously broken. As we have shown, this is true only for $d \leq 2$. This situation is not unlike equilibrium critical phenomena, where hyperscaling is not valid above the upper critical dimension.

In the previous section, we found $\nu = 1$ for $d \geq 2$ and $\nu = 2$ for $d = 1$. For the Stepping-Stone model, this is the only difficult exponent to determine. The other exponents, η and z are relatively easy to get. First, we note that when $s = 0$, the two-point propagator coming from the action of Eq. (4.53), or Figure (4.3), is simply that of a random walk. Therefore, $z = 1$. The next observation to make, is that, also for $s = 0$, the expectation of the number of particles is easily shown to be conserved.

(Formally, one can show this by going back to the Liouville operator and evaluating this expectation using the commutation relations). This tells us that $\eta = 0$. We can use the scaling relations above to get the rest. In particular, for $d = 1$

$$\delta = \frac{1}{2},$$

but β is the same as in higher dimensions

$$\beta = 1.$$

For higher dimensions, δ takes its “mean-field” value of 1, in the sense that this is what we derived for the Wright-Fisher/Moran model in section 2.4. Figure (4.1) shows simulation results for the time-dependent survival probability in the $d = 1$ Stepping-Stone model. These results agree with the theoretical calculation of $\delta = \frac{1}{2}$.

We should remark that the scaling relations above are definitely correct for $d = 1, 3$, but in $d = 2$ one expects logarithmic corrections.

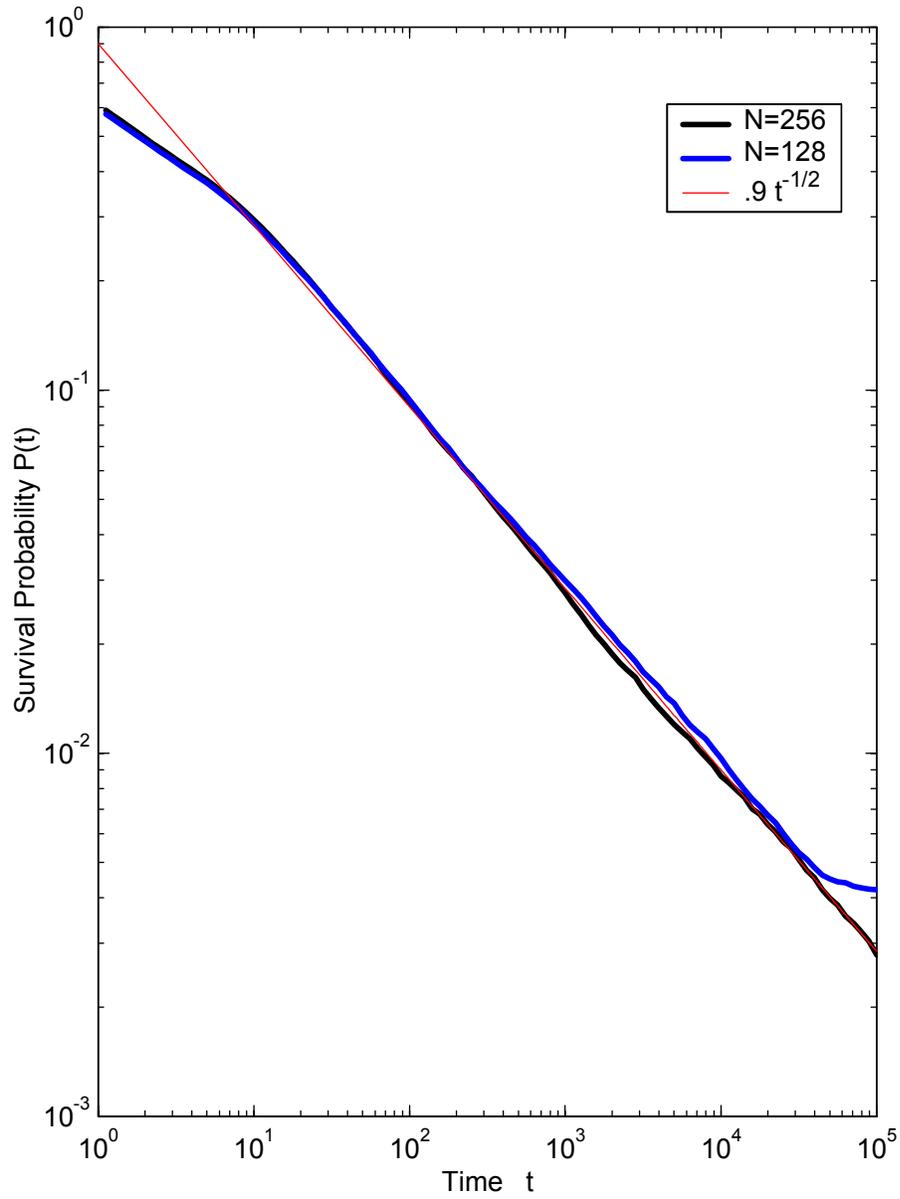


Figure 4.1: Simulation results for the time-dependent survival probability for $d = 1$. Shown are simulations for the $N = 128$ and 256 lattices which clearly show the $t^{-1/2}$ dependence of the survival probability.

Chapter 5

Renormalization Group Analysis

5.1 Introduction.

A very significant advance in our understanding of critical phenomena occurred with the advent of Renormalization Group (RG) . The basic ideas of the RG were introduced in a series of papers by Ken Wilson and co-workers in the early 1970's [65]. These ideas have since taken hold in practically every avenue of theoretic physics and have changed the way we understand and look at problems in many-body systems. Prior to the inception of the RG picture, phenomenological assumption about the form of thermodynamic functions near critical points lead to scaling laws, similar to those described at the end of the previous chapter. RG theory not only justifies the scaling laws, it also provides a framework to calculate the exponents in a systematic way, and for difficult problems that defy exact results. RG techniques fall roughly into two classes. The first goes by the name ' k -space' or 'wave-vector shell' Renormalization Group, and the second 'real-space renormalization'. In this thesis, we will use the wave-vector shell Renormalization Group to analyze the Stepping-Stone model, and recover many of the results presented earlier, but from a very different point of view.

The basic idea of the RG is fairly simple. One wishes to know how the model in question behaves at large scales of both space and time. One constructs a series of transformations on the model, First, by integrating out the small-scale, fast-time dynamics. This is followed by a re-scaling of space and time. The parameters characterizing the

model are thus modified and repeating this procedure generates a flow in the parameter space of the model. The ultimate goal is to find fixed points of this transformation which can be due to 1) the fact that the fluctuations have all averaged out at large scales, or even more interesting, 2) the fluctuations have no characteristic scale and persist from the micro-scale to the macro-scale. The latter possibility is the hallmark of critical behaviour. Fixed points corresponding to these critical points are unstable with respect to one, or perhaps multiple parameters of the model and it is the linearized eigenvalues of the RG flow that determine the critical exponents.

5.2 A Simple Example of Renormalization: The biased random walk.

In this section we will provide a simple illustration of Wilson's wave-vector shell Renormalization Group.

The time-evolution operator for a random walk with diffusion constant D and bias α in the x_1 -direction is given by

$$\mathbf{U} [\bar{\phi}, \phi] = e^{-S[\bar{\phi}, \phi]}, \quad (5.1)$$

with action

$$S [\bar{\phi}, \phi] = \int dt \int d^d \mathbf{r} \bar{\phi}_{\mathbf{r}} (\partial_t - D \nabla^2) \phi_{\mathbf{r}} - \alpha \left(\frac{\partial \bar{\phi}_{\mathbf{r}}}{\partial x_1} \right) \phi_{\mathbf{r}}.$$

In wave-vector space, one has

$$S [\bar{\phi}, \phi] = \int dt \int d^d \mathbf{k} \bar{\phi}_{\mathbf{k}} (\partial_t + D \mathbf{k}^2) \phi_{\mathbf{k}} + i k_1 \alpha \bar{\phi}_{\mathbf{k}} \phi_{\mathbf{k}}.$$

(Recall that $\bar{d} \equiv \frac{d}{2\pi}$, so $\bar{d}^d \mathbf{k} = \frac{d^d \mathbf{k}}{(2\pi)^d}$).

Step 1: Separate the \mathbf{k} integral into “fast” and “slow” modes and integrate out modes near the cutoff.

This resulting action is

$$S[\bar{\phi}, \phi] = S_{<}[\bar{\phi}, \phi] + S_{>}[\bar{\phi}, \phi],$$

with

$$S_{<}[\bar{\phi}, \phi] = \int dt \int_0^{\Lambda/b} d^d \mathbf{k} \bar{\phi}_{\mathbf{k}}(\partial_t + m\mathbf{k}^2)\phi_{\mathbf{k}} + ik_1 \alpha \bar{\phi}_{\mathbf{k}}\phi_{\mathbf{k}}, \quad (5.2)$$

and

$$S_{>}[\bar{\phi}, \phi] = \int dt \int_{\Lambda/b}^{\Lambda} d^d \mathbf{k} \bar{\phi}_{\mathbf{k}}(\partial_t + m\mathbf{k}^2)\phi_{\mathbf{k}} + ik_1 \alpha \bar{\phi}_{\mathbf{k}}\phi_{\mathbf{k}}.$$

Here $b \gtrsim 1$ is our factor describing our wave-vector shell just below the cutoff. Now we integrate out the “fast” modes. This is easy since the action, written in wave-vector space, does not mix “fast” and “slow” modes and is of gaussian form. The integration simply adds a multiplicative constant to the exponential in Eq. (5.1), and the resulting action of the “slow” fields is simply what we had before in Eq. (5.2).

Step 2: Re-scale momenta and time.

As the objective is to end up with an action that looks the same as before, in particular, we want to restore the cutoff on wave-vector to its original value, so we let

$$k_i \rightarrow b^{-1}k_i,$$

and re-scale time by some as yet undetermined factor b^z , *i.e.*

$$t \rightarrow b^z t.$$

The resulting action for the “slow” fields becomes

$$S_{<}[\bar{\phi}, \phi] = \int dt \int_0^{\Lambda} d^d \mathbf{k} b^{-d} \bar{\phi}_{\mathbf{k}}(\partial_t + b^{z-2} D\mathbf{k}^2)\phi_{\mathbf{k}} + ib^{z-d-1} \alpha k_1 \bar{\phi}_{\mathbf{k}}\phi_{\mathbf{k}}.$$

Step 3: Re-scale Fields to find fixed-points of the transformation.

To accomplish this we let

$$\begin{aligned}\bar{\phi}_{\mathbf{k}} &\rightarrow b^p \bar{\phi}_{\mathbf{k}} \\ \phi_{\mathbf{k}} &\rightarrow b^q \phi_{\mathbf{k}},\end{aligned}$$

and now look for values of the exponents z , p and q that produce a fixed point. It is important at this stage to point out the necessity of choosing independent re-scaling factors for the $\bar{\phi}_{\mathbf{k}}$ and $\phi_{\mathbf{k}}$ fields. This is due to the fact that within both the creation/annihilation formalism, and the response function or Fokker-Planck based field theory, most processes, especially those that describe an interaction between diffusing particles (unlike the ones considered in the present section, but see below) contribute two terms to the action that have different numbers of creation operators, or equivalently $\bar{\phi}$ fields.

First we consider the term under the action integral with the time derivative,

$$dt \bar{d}^d k \partial_t \bar{\phi}_{\mathbf{k}} \phi_{\mathbf{k}} \rightarrow b^{q+p-d} dt \bar{d}^d k \partial_t \bar{\phi}_{\mathbf{k}} \phi_{\mathbf{k}}.$$

We need to keep this fixed so

$$p + q = d.$$

Next we consider the diffusive and drift terms, which as a result of the above choice give us new constants *via* the recursion relations

$$D' = b^{z-2} D \tag{5.3}$$

and

$$\alpha' = b^{z-1} \alpha. \tag{5.4}$$

From the recursion relations above, we find two different fixed point corresponding to different choices of z . The choice $z = 2$ produces the ‘random walk’ fixed point, however, the drift term grows like b at each iteration of the renormalization. In other words, this

fixed point is unstable to drift. On the other hand, the choice $z = 1$ gives us the ‘Drift’ fixed point, for which the diffusion constant D flows to zero. This is naturally what we expect for drift: at large length and time scales, the effective motion is rectilinear. Using the terminology of the renormalization group, the drift term is *relevant*, with respect to the random walk fixed point.

Albeit simple, this example illustrates some of the basic notions of the renormalization group and provides a foundation upon which to generalize the procedure to more complicated systems. With further analysis of the above RG flow, one could extract the various exponents that characterize the random walk with and without drift, however in a sense, the results are somewhat trivial given that these are exactly known, *i.e.* the propagators are easily derived due to the non-interacting character of the system.

5.3 Perturbations from the Random Walk Fixed Point.

It will be necessary for us to understand more complicated perturbations from the random walk fixed point than considered above. In general, we cannot easily make the separation into “slow” and “fast” fields as above—the resulting action will contain terms that mix “slow” and “fast” modes. It will, however, be very instructive to consider the RG flow for small perturbations from the random walk fixed point. In this manner we can study the stability of the random walk fixed point. In the process, we will discover yet another non-interacting fixed point of central importance in population-like systems.

5.3.1 n -particle to m -particle processes.

We start with a random walk action as discussed above, but omitting the drift term. Suppose we add small perturbation corresponding to the stochastic process of going from n -particles to m -particles. This contributes 2 terms to the action

$$\delta S_{m,n}^{(2)} = u_{m,n} \int dt \int^\Lambda \bar{d}^d \mathbf{k}_1 \dots \bar{d}^d \mathbf{k}_{m+n} \bar{\phi}_{\mathbf{k}_1} \bar{\phi}_{\mathbf{k}_2} \dots \bar{\phi}_{\mathbf{k}_m} \phi_{\mathbf{k}_{m+1}} \phi_{\mathbf{k}_{m+2}} \dots \phi_{\mathbf{k}_{m+n}} \delta \left(\sum_i^{m+n} \mathbf{k}_i \right),$$

and

$$\delta S_{m,n}^{(2)} = -u_{m,n} \int dt \int^\Lambda \bar{d}^d \mathbf{k}_1 \dots \bar{d}^d \mathbf{k}_{2n} \bar{\phi}_{\mathbf{k}_1} \bar{\phi}_{\mathbf{k}_2} \dots \bar{\phi}_{\mathbf{k}_n} \phi_{\mathbf{k}_{n+1}} \phi_{\mathbf{k}_{n+2}} \dots \phi_{\mathbf{k}_{2n}} \delta \left(\sum_i^{2n} \mathbf{k}_i \right).$$

Following the three steps outlined in the example of the preceding section, and requiring that the term with the time-derivative be invariant under re-scaling, we have for the field re-scaling factors

$$p + q = d.$$

The random walk fixed point is accessed by choosing

$$z = 2.$$

For arbitrary p and q , the two terms will scale differently. Accordingly, we will denote the two new coupling constants $u_{m,n}^{(1)}$ and $u_{m,n}^{(2)}$. The resulting recursion relations describing the growth (or decay) of the perturbations are

$$\begin{aligned} u_{m,n}^{(1)} &= b^{2-(m+n-1)d+mp+nq} u_{m,n} \\ u_{m,n}^{(2)} &= b^{2-(2n-1)d+n(p+q)} u_{m,n} \end{aligned}$$

We can make both coupling constants behave in the same manner if we choose

$$\begin{aligned} p &= d, \\ q &= 0. \end{aligned}$$

The result being $u_{m,n}^{(1)} = u_{m,n}^{(2)} = u'_{m,n}$, and

$$u'_{m,n} = b^{2-(n-1)d} u_{m,n} \quad (5.5)$$

So the relevance of a perturbation from the random walk fixed point depends only on the number of incoming ϕ fields.

5.3.2 Relevance of independent decay and branching processes.

The simplest process to consider are those which have only one annihilation operator or ϕ field, for example random decay (or death), and branching (or birth). Having only one power of ϕ in the action means that these processes occur independently of other particles. These processes are always relevant from random walk fixed point, as can be seen by putting $n = 1$ in Eq. (5.5), which tells us that these terms grow as b^2 . One could take $z = 0$ in the above RG procedure and find

$$D' = b^{-2}D$$

and

$$u'_{m,1} = u_{m,1},$$

i.e. the diffusion constant D flows to zero. This type of fixed point is non-critical, which is easy to see since it is stable to any perturbation (except, of course other terms with $n = 1$, which are marginal).

5.3.3 Behavior of coagulation, or annihilation processes.

Next we consider processes such as coagulation ($A+A \rightarrow A$) or annihilation ($A+A \rightarrow 0$). For example, the former would have a term in the Lagrangian

$$u (\bar{\phi}_{\mathbf{r}} - \bar{\phi}_{\mathbf{r}}^2) \phi_{\mathbf{r}}^2. \quad (5.6)$$

and would produce a RG flow $u' = b^{2-d}u$, so these processes are irrelevant with respect to the random walk fixed point for $d > 2$.

5.3.4 A paradox.

Higher order process (*i.e.* $n \geq 3$) would all be irrelevant at $d \geq 2$ and so one might be tempted to conclude that the upper critical dimension $d_c = 2$ for these models. The

processes $A + A \rightarrow A$ and $A + A \rightarrow 0$ have been quite well studied in the literature where both exact results, RG arguments and simulations clearly establish that d_c is in fact equal to 2.

Perhaps the most pervasive critical behaviour known in non-equilibrium statistical mechanics is that of *directed percolation* (DP), also known as Reggeon field theory in particle physics. DP is quite easily formulated using the operator/path-integral techniques discussed in this thesis. Its population model analog contains, in addition to births and deaths, density-dependent mortality—or effectively decay, branching and a coagulation term, all of which have been discussed above. So shall we conclude on the basis of the above argument that $d_c = 2$ for DP also? Unfortunately this is quite wrong. DP has been quite well-studied in the literature and is found to have an upper critical dimension $d_c = 4$. This fact is well-established by the breakdown of perturbation theory for $d < 4$ and non-so-transparent field-theoretic RG results.

It seems unfortunate that for DP, $d_c = 4$ does not come out of the simple wave-vector RG argument above. Is there something wrong with the theory or our argument? The resolution of this ‘paradox’ is that our argument about the stability of the random walk fixed point is in fact correct. What we have neglected, however, is the possibility of another critical fixed point. It is with respect to this fixed point that $d_c = 4$ for directed percolation.

5.4 The Critical Branching Fixed Point.

In the preceding section we found that perturbations described by a single annihilation operator or ϕ field were classified as ‘relevant’ with respect to the random walk fixed point. Independent branching is described by such a term and therefore the random walk fixed point is always (in arbitrary dimension) unstable to such a branching process.

We should expect from our study of branching processes in appendix C to find another fixed point corresponding to a branching random walk where the branching rate is equal to the decay rate. The qualitative difference between a critical branching random walk and the simple random walk is that the latter conserves the number of particles. In terms of the spreading exponents introduced in the last section of chapter 3, the simple random walk has a survival probability identically equal to 1, so both the exponent β in Eq. (4.59) and the exponent δ in Eq. (4.56) are identically equal to 0. For the critical branching process, however, we know from the exact results in appendix C (see also section 2.4.2) that $\delta = 1$ and $\beta = 1$. The mean number of particles in both models is conserved so both have $\eta = 0$. Of course, in both cases these exponents are consistent with the scaling laws in Eq. (4.61) and Eq. (4.63).

We will now explicitly find the RG fixed point corresponding to the critical branching process and then consider perturbations as was done for the simple random walk above. In addition to the random walk term, the action for the branching random walk has both a decay term and a branching term, *i.e.*

$$S = S_{RW} + S_B + S_D,$$

with

$$S_{RW} = \int dt \int d^d \mathbf{r} \bar{\phi}_{\mathbf{r}} (\partial_t - D \nabla^2) \phi_{\mathbf{r}},$$

$$S_B = \gamma \int dt \int d^d \mathbf{r} (\bar{\phi}_{\mathbf{r}}^2 \phi_{\mathbf{r}} - \bar{\phi}_{\mathbf{r}} \phi_{\mathbf{r}}),$$

and

$$S_D = \lambda \int dt \int d^d \mathbf{r} (\phi_{\mathbf{r}} - \bar{\phi}_{\mathbf{r}} \phi_{\mathbf{r}}).$$

and the system is critical when $\gamma = \lambda$. It will be convenient to shift the $\bar{\phi}_{\mathbf{r}}$ field by 1, by defining the field

$$\tilde{\phi}_{\mathbf{r}} \equiv \bar{\phi}_{\mathbf{r}} - 1.$$

Grassberger [43] has discussed this shift in the context of the creation-annihilation operator formalism, showing that it corresponds to an alternative, but equally valid choice of the scalar product. The resulting action becomes

$$S_{RW} = \int dt \int d^d \mathbf{r} \tilde{\phi}_{\mathbf{r}} (\partial_t - D \nabla^2) \phi_{\mathbf{r}},$$

$$S_B + S_D = \int dt \int d^d \mathbf{r} \gamma \tilde{\phi}_{\mathbf{r}}^2 \phi_{\mathbf{r}} - (\gamma - \lambda) \bar{\phi}_{\mathbf{r}} \phi_{\mathbf{r}},$$

At the critical point, therefore one finds only the $\gamma \tilde{\phi}_{\mathbf{r}}^2 \phi_{\mathbf{r}}$ term.

Perform the first two steps of the renormalization procedure, we find the recursion relations

$$D' = b^{2-z} D,$$

and

$$\gamma' = b^{2-2d+2p+q} \gamma.$$

Once again we must have $p + q = d$. We can keep both the diffusion constant D and the branch rate γ fixed by the choices

$$q = 2,$$

$$p = d - 2,$$

$$z = 2.$$

Thus we have found the fixed point corresponding to the critical branching process.

5.4.1 Perturbations from the Critical Branching fixed point.

We now consider perturbations from the Critical Branching fixed point. Following our standard procedure, we look at a perturbation due to a n -particle to m -particle process, and find the coupling constant behaves under our renormalization procedure as

$$u'_{m,n} = b^{2-(m+n-1)d+m(d-2)+2n} u_{m,n}. \quad (5.7)$$

Let us first consider the coagulation, or density-dependent mortality term, whose coupling constant we denote by $u_{1,2}$. Recall that in the last section, this term was found relevant for $d < 2$ with respect to the random walk fixed point, and thereby provided us with something of a paradox given the well-known upper critical dimension of DP, $d_c = 4$. Using the above recursion relation, we see that

$$u'_{1,2} = b^{4-d} u_{1,2},$$

confirming the upper critical dimension for DP.

In the next section, when we renormalize the stepping-stone model in the broken symmetry state, the term corresponding to $u_{1,2}$ will vanish at the critical point, thus will not play a role in the critical behaviour. It will be necessary, however, to consider the process with $m = 2$, $n = 2$. One finds

$$u'_{2,2} = b^{2-d} u_{2,2}.$$

Hence, models that fall into this category will have an upper critical dimension, $d_c = 2$, which agrees with the exact results presented in the last chapter.

5.5 The Gaussian Fixed Point.

Before we proceed to renormalize the stepping-stone model, there is one more “easy” fixed point to consider. We shall need it when we consider renormalization of the stepping stone model for the description of the steady-state critical behaviour obtained as the mutation rate goes to zero. It is also important to consider since it represents a convenient way to incorporate dynamics into equilibrium statistical mechanics and has been thoroughly studied in the context of *dynamical critical phenomena*. [30][31][32]

In equilibrium statistical mechanics one has a free-energy H , which for the present context can be considered a functional of some field $\phi_{\mathbf{r}}$. The simplest way to add dynamics to the static picture of equilibrium statistical mechanics is to postulate a Langevin equation for $\phi_{\mathbf{r}}$

$$\frac{d\phi_{\mathbf{r}}}{dt} = \frac{\delta H[\phi_{\mathbf{r}}]}{\delta\phi_{\mathbf{r}}} + \eta(t), \quad (5.8)$$

where $\eta(t)$ is a Gaussian noise term with auto-correlation

$$\langle \eta(t) \eta(t') \rangle = 2\gamma \delta(t - t').$$

Eq. (5.8) is also known as the time-dependent Landau-Ginzburg (TDLG) equation. The action S , or dynamical functional corresponding to Eq. (5.8) is found by looking at the equivalent Fokker-Planck equation, or using the response function method, and is given by

$$S = \int dt \int d^d\mathbf{r} \bar{\phi}_{\mathbf{r}} \frac{\delta H[\phi_{\mathbf{r}}]}{\delta\phi_{\mathbf{r}}} - \gamma \bar{\phi}_{\mathbf{r}}^2. \quad (5.9)$$

If we consider H to represent the Gaussian model, we have

$$H = \frac{D}{2} (\nabla\phi_{\mathbf{r}})^2 + r_0\phi_{\mathbf{r}}^2, \quad (5.10)$$

and the dynamical functional (*e.g.* from the response-function formalism) associated with the TDLG equation above is

$$S = \int dt \int d^d\mathbf{r} \bar{\phi}_{\mathbf{r}} (\partial_t - D\nabla^2)\phi_{\mathbf{r}} + r_0\phi_{\mathbf{r}}^2 - \gamma\bar{\phi}_{\mathbf{r}}^2. \quad (5.11)$$

Here r_0 is proportional to the reduced temperature, and hence the critical point occurs at $r_0 = 0$. We note that the action in Eq. (5.11) can be viewed as a random walk action with a “source” term. We now have to deal with this “source” term, $\bar{\phi}_{\mathbf{r}}^2$. Setting $n = 0$ in Eq. (5.5) shows that it is relevant with respect to the Random Walk fixed point in all dimensions.

Once again, one needs to find the appropriate field re-scaling exponents p and q to produce the RG fixed point. One simply repeats the above arguments for the $\gamma\bar{\phi}_{\mathbf{r}}^2$ term.

One finds the results

$$p = \frac{d-2}{2},$$

$$q = \frac{d+2}{2}.$$

5.5.1 Perturbations from the Gaussian Fixed Point.

Consider now the effect of our standard perturbation $\propto \bar{\phi}_{\mathbf{r}}^m \phi_{\mathbf{r}}^n$ on the RG flow near the Gaussian fixed point. Using the above values for p and q , we find

$$u'_{m,n} = b^{2-d(\frac{m+n}{2}-1)+n-m} u_{m,n}$$

To make further contact with equilibrium statistical mechanics, we consider the RG flow arising from a $\phi_{\mathbf{r}}^{n+1}$ perturbation in H . The corresponding term in the dynamical functional, would be $\propto \bar{\phi}_{\mathbf{r}} \phi_{\mathbf{r}}^n$. It is important to note here that within the context of equilibrium statistical mechanics, except for the $\bar{\phi}_{\mathbf{r}}^2$ term that we have kept invariant under the RG transformation, only terms with one power of $\bar{\phi}_{\mathbf{r}}$ will appear in the dynamical functional. This is evident from Eq. (5.9). The form of the noise term, $\propto \bar{\phi}_{\mathbf{r}}^2$, is the result of a constant noise source (*i.e.* the noise is independent of $\phi_{\mathbf{r}}$). The term coming from the free-energy H is given by $\bar{\phi}_{\mathbf{r}} \frac{\delta H[\phi_{\mathbf{r}}]}{\delta \phi_{\mathbf{r}}}$ and only contains the explicit $\bar{\phi}_{\mathbf{r}}$ since $H[\phi_{\mathbf{r}}]$ does not depend on any $\bar{\phi}_{\mathbf{r}}$. The coupling constant for a $\bar{\phi}_{\mathbf{r}} \phi_{\mathbf{r}}^n$ term in S we denote by u_n , flows as

$$u'_n = b^{1+n-\frac{1}{2}d(n-1)} u_n$$

from the Gaussian fixed point. In particular, a $\phi_{\mathbf{r}}^4$ term in H , which corresponds to a $\bar{\phi}_{\mathbf{r}} \phi_{\mathbf{r}}^3$ term in the action, flows as

$$u'_3 = b^{4-d} u_3.$$

from the Gaussian fixed point. In a similar manner, a $\phi_{\mathbf{r}}^6$ term in H , produces the flow

$$u'_5 = b^{\frac{3-d}{2}} u_5.$$

Both of these results are identical to the standard, non-dynamical RG theory of equilibrium statistical mechanics. (see, for example [37]).

Finally, in the following section, we will need to consider a term in the Lagrangian proportional to $\bar{\phi}_{\mathbf{r}}^2 \phi_{\mathbf{r}}^2$. As mentioned above, this kind of term would never occur in the equilibrium theory. Using the above results, this term scales as

$$u'_{2,2} = b^{2-d} u_{2,2},$$

further confirming $d_c = 2$, for the stepping stone model.

5.6 Summary of non-interacting fixed points.

In the preceding sections of this chapter, we went to great lengths to elaborate the various non-interacting fixed points described by the creation/annihilation formalism or alternatively the response-function/Fokker-Planck formalism. Our reasons for doing this are twofold. First, this distinction has not always been appreciated or even made in the literature. This is perhaps because all of these non-interacting models can be called “gaussian” in a sense—the propagator is of gaussian form. Yet, as we have discussed, each one describes different physics. Although they share some critical exponents (specifically, those describing the two-point correlation function), they differ in other respects, for example, the survival exponents and conservation of particles. More importantly, however, the distinctions we have elucidated above demonstrate both the flexibility and breath of these operator/path-integral formalisms to describe the diverse world of stochastic processes.

Let us briefly summarize these fixed points and their stability within the RG picture. First, we have the Random Walk fixed point. A prototypical model would be a random walk (or collection of random walkers) with spontaneous decay. When the decay rate vanishes, the number of particles (in contrast to the mean number of particles) is conserved. The Random Walk fixed point is always (in arbitrary dimension) unstable to a “branching term”. When the branching rate equals the decay rate, the relevant fixed point becomes what we have named the Critical Branching Process fixed point. Only the mean number of particles is conserved, and non-zero survival exponents are found. Next we found that the Gaussian fixed point of the TDLG extension of the equilibrium Gaussian model could be viewed as the inclusion of a “source” term to the Random Walk Lagrangian. This term also is relevant in arbitrary dimension, with respect to the Random Walk fixed point, giving way to the Gaussian fixed point. Indeed, we have uncovered a very interesting equivalence: fluctuations in the TDLG extension of dynamics to equilibrium statistical field theory can be described as a source of spontaneously decaying random walkers. In the Landau-Ginzburg model of equilibrium phase transitions, one has a $\phi_{\mathbf{r}}^4$ term in H and hence the walkers would also be interacting. In general, using the random walker picture, the absolute number of walkers is not conserved. (We should note, however, that it is possible to construct a free-energy in Eq. (5.8) that does conserve the absolute number, by making the source term $\propto k^2$. This model is called “model B” in the literature on dynamical critical phenomena [30][31]. It is quite straight-forward to see that a term like this grows like b^d from the Random Walk fixed point.)

Finally, we have evaluated the stability of these fixed points to perturbations due to various interactions between particles. Some of these terms, like those found in directed percolation or ϕ^4 -theory are relevant for $d < 4$, however, these terms are not present at the critical points of the Stepping-Stone model, nor do they get generated in the renormalization procedure. We have discussed the terms that *are* present, finding them

relevant for $d < 2$. In the next section, we will go beyond the perturbation analysis in this section and explicitly find the non-trivial fixed points for $d < 2$ and show how they give rise to “non-classical” critical exponents.

5.7 Renormalization of the Stepping-Stone Model.

Using the wave-vector shell RG introduced in the previous sections, we now proceed with full renormalization of the Stepping-Stone Model. First we must distinguish between the two different regimes of the model, already discussed in chapter 3. The first describes the steady-state of the “neutral theory”, which occurs when selection is absent and becomes critical as the mutation rate vanishes. Using the results of the last few sections, mere inspection of the Lagrangian shows that the critical behavior for $d \geq 2$ is controlled by the Gaussian fixed point. On the other hand, for the description of the broken-symmetry state we have a Lagrangian describing single new alleles, with vanishingly small selective difference, propagating in a “sea” of the fixed allele. From the Lagrangian of the broken-symmetry state, we see that the critical behaviour for $d \geq 2$ is controlled by the Critical Branching Process fixed point. We also note that $d = 2$ is marginal for both these situations, therefore one expects logarithmic corrections for $d = 2$ systems.

5.7.1 The Neutral Regime.

Here we follow the Wilson-like procedure [65][38] introduced in section 4.2. Because of the similar diagrammatic expansions, the results of our renormalization group analysis are somewhat similar to those of the diffusive annihilation ($A + A \rightarrow 0$) and diffusive coagulation ($A + A \rightarrow A$) models.[66][67] First, we employ the same procedure discussed at the beginning of section 3.5 to get an action S from the Liouville operator for the ECSS model given in Eq. (4.8). One again, the procedure is to introduce the fields

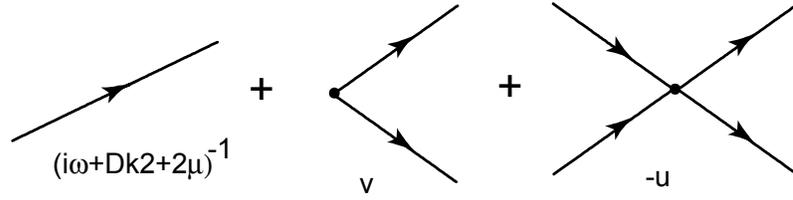


Figure 5.2: Vertices representing the Lagrangian of the Neutral regime

$\bar{\phi}_{\mathbf{r}}(t)$ and $\phi_{\mathbf{r}}(t)$ and write S as

$$S[\bar{\phi}, \phi] = \int dt \int d^d \mathbf{r} \bar{\phi}_{\mathbf{r}}(t) \partial_t \phi(t)_{\mathbf{r}} - L[\bar{\phi}_{\mathbf{r}}(t), \phi_{\mathbf{r}}(t)],$$

where $L[\bar{\phi}_{\mathbf{r}}(t), \phi_{\mathbf{r}}(t)]$ obtained from $\mathcal{L}[a_{\mathbf{r}}^\dagger, a_{\mathbf{r}}]$ by replacing the $a_{\mathbf{r}}^\dagger$ and $a_{\mathbf{r}}$ operators of \mathcal{L} with the fields $\bar{\phi}_{\mathbf{r}}(t)$ and $\phi_{\mathbf{r}}(t)$, respectively. The resulting action reads,

$$\begin{aligned} S[\bar{\phi}, \phi] &= \int \bar{d}\omega \int^\Lambda \bar{d}^d \mathbf{k} \quad (i\omega + D\mathbf{k}^2 + 2\mu) \bar{\phi}_{\mathbf{k},\omega} \phi_{\mathbf{k},\omega} \\ &\quad + v \int \bar{d}\omega \int^\Lambda \bar{d}^d \mathbf{k} \quad \bar{\phi}_{\mathbf{k},\omega} \bar{\phi}_{-\mathbf{k},-\omega} \\ &\quad - u \int \bar{d}\omega_1 \bar{d}\omega_2 \bar{d}\varepsilon \int^\Lambda \bar{d}^d \mathbf{k}_1 \bar{d}^d \mathbf{k}_2 \bar{d}^d \mathbf{q} \bar{\phi}_{\mathbf{k}_1,\omega_1} \bar{\phi}_{\mathbf{k}_2,\omega_2} \phi_{\mathbf{k}_1+\mathbf{q},\omega_1+\varepsilon} \phi_{\mathbf{k}_2-\mathbf{q},\omega_2-\varepsilon}, \end{aligned} \quad (5.12)$$

which can be represent graphically as the edges and vertices in Figure (5.2). In the previous chapter, we had

$$u = v = (2n)^{-1}.$$

In the bare model, they are identical, and our choice of time measured in generations made them equal to 1. For the present purposes, we need to distinguish between them to allow for the possibility of distinct RG flows. The separation of fields into “fast” and “slow” modes results in a term that mixes the two and hence the integration over the “fast” modes produces additional contributions to the renormalized coupling constants. We must use our knowledge of field theory to evaluate these contributions—the net

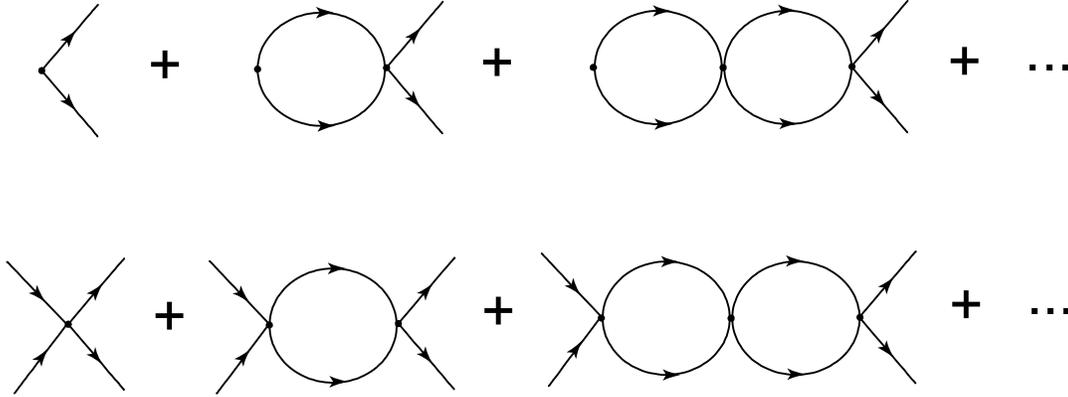


Figure 5.3: Bubble diagrams contributing to the renormalization of v (above) and u (below).

result is precisely what we would arrive at from a standard perturbation expansion, with the important exception that the wave-vector \mathbf{q} in the “loop” integrals is restricted to $\Lambda/b \leq |\mathbf{q}| \leq \Lambda$.

The first important observation to make is that applying steps 1 and 2 of our RG procedure to the Lagrangian in Eq. (5.12) or Figure (5.2) does not “renormalize” the mutation rate μ . (Obviously, this is also the case for the standard perturbation expansion.) So its behavior under the full RG transformation will be identical to its flow from the Gaussian fixed point and grow as b^2 . Hence, we can forget about it for the rest of the discussion. The second observation is that the RG procedure only can only generate additional terms $\propto \bar{\phi}^m \phi^n$, with $n \geq 2$ and $m \geq 3$. These terms are easily shown to be irrelevant. This leaves us with only the two original terms and these have similar renormalizations, which we can perform to all orders by summing the “parquet” diagrams in Figure (5.3), we have already seen in the previous chapter.

For the evaluation of the “bubble” integral, we can take the external wave-vector to be

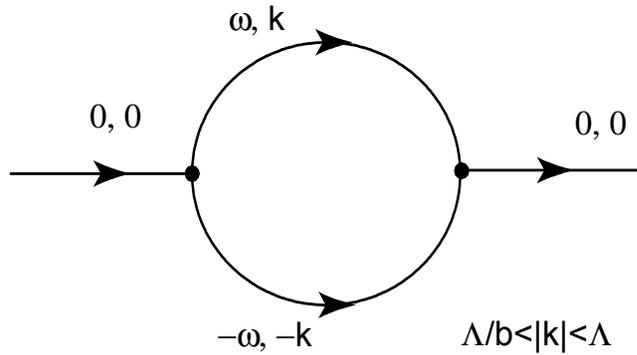


Figure 5.4: Diagram representing the 1-loop integral that contributes to the differential RG recursion relation.

zero, see Figure (5.4). This is because any k -dependence generated in the renormalization of u is easily show to be irrelevant. Furthermore, we only need to know the *differential* recursion relation for u and v to find the RG fixed point and its associated eigenvalue. We will show shortly that in the limit $b = e^l \rightarrow 1$, diagrams with n loops proportional to l^n , therefore only the one-loop diagrams contribute to the differential recursion relation.

The net result of our analysis, is the recursion relations

$$\begin{aligned} u' &= b^{2-d} u \left(1 - \frac{u}{D} I_d \right), \\ v' &= v \left(1 - \frac{u}{D} I_d \right) \end{aligned}$$

where

$$\begin{aligned} I_d &= \int_{-\infty}^{\infty} d\omega \int_{\Lambda/b}^{\Lambda} d^d \mathbf{k} \frac{1}{\omega^2 + k^4} \\ &= \int_{\Lambda/b}^{\Lambda} d^d \mathbf{k} \frac{1}{k^2}. \end{aligned}$$

We have

$$I_d = \begin{cases} \frac{1}{\pi} \frac{b-1}{\Lambda}, d = 1 \\ \frac{1}{\pi D} \ln b, d = 2 \\ \frac{1}{\pi D} \Lambda \frac{b-1}{b}, d = 3 \end{cases}$$

We now define l through $e^l \equiv b$. In the limit $l \rightarrow 0$, $b \simeq 1 + l$ and we have

$$I_d = \begin{cases} \frac{1}{\pi} \frac{l}{\Lambda}, d = 1 \\ \frac{l}{\pi D}, d = 2 \\ \frac{1}{\pi D} \Lambda l, d = 3 \end{cases} .$$

As mentioned earlier, this tells us the diagrams with n loops are proportional to l^n .

The differential recursion relations for $d = 1$ read

$$\begin{aligned} \frac{du}{dl} &= u - \frac{u^2}{D\pi\Lambda}, \\ \frac{dv}{dl} &= -\frac{uv}{D\pi\Lambda}. \end{aligned}$$

These equations have the non-trivial fixed point

$$\begin{aligned} u^* &= D\pi\Lambda, \\ v^* &= 0. \end{aligned}$$

For $d = 2$, the recursion relations are

$$\begin{aligned} \frac{du}{dl} &= -\frac{u^2}{\pi D}, \\ \frac{dv}{dl} &= -\frac{uv}{D\pi}. \end{aligned}$$

These solution to these equations is

$$\begin{aligned} u(l) &= u_0 \left(1 + \frac{u_0}{D\pi} l\right)^{-1}, \\ v(l) &= v_0 \left(1 + \frac{u_0}{D\pi} l\right)^{-1}. \end{aligned}$$

Now the only fixed point is at $u^* = v^* = 0$.

For $d = 3$, the differential recursion relations are, once again, qualitatively different,

$$\begin{aligned}\frac{du}{dl} &= -u - \frac{u^2}{\pi D}, \\ \frac{dv}{dl} &= -\frac{uv}{D\pi}.\end{aligned}$$

The somewhat surprising result is a line of fixed points with $u^* = 0$ and v dependent on the initial values.

The RG flow for all of the above is plotted in Figure (5.5). At first sight the result for $d = 3$ might seem somewhat strange but a moment's reflection tells us that this is in complete harmony with the exact results from the previous chapter. For $d = 2$ and 3, we see that the u -term representing the interaction is irrelevant. This term is responsible for the anomalous decay of the two-point correlation function in $d = 1$. For $d = 1$ and 2, we see that the v -term, which loosely speaking is the 'source term' for the fluctuations, flows to zero at large length and time scales. This reflects an absence of fluctuations in the steady-state and is precisely what happens in the absorbing state of fixation that occurs in $d = 1$ and 2, where as we have argued the symmetry is spontaneously broken. For $d = 3$, v remains finite and the effective description a large scales is that of the Gaussian model.

5.7.2 The Broken-Symmetry State.

We now follow a similar procedure and apply the RG to the action that describes propagation of single, new mutant alleles. We re-write the action from Eq. (4.53) here, making the parameters of the model explicit by introducing the constant u as was done in the previous sub-section,

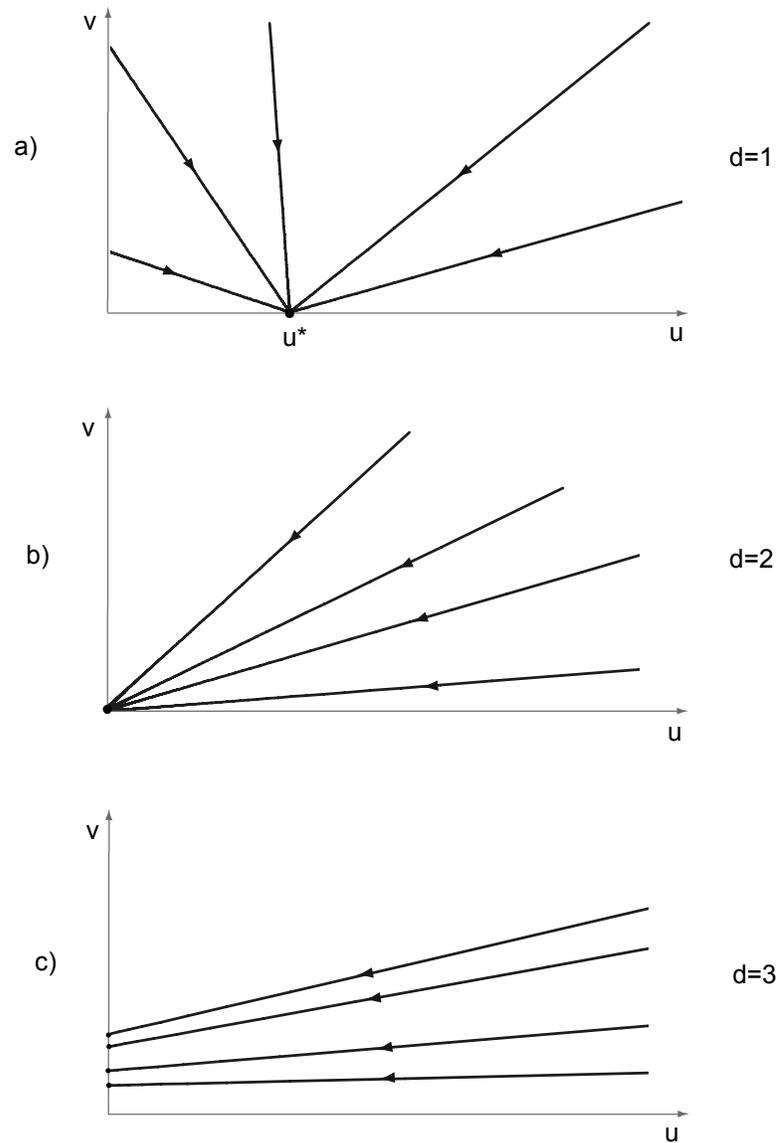


Figure 5.5: Renormalization group flow for the neutral theory. Top: The flow for $d = 1$ shows the non-trivial fixed point at $u^* = D\pi\Lambda$. Top and Centre: v flows to zero for $d = 1$ and 2 and indicates that the absorbing state is eventually reached. Centre and Bottom: u flows to zero but v reaches a line of fixed points that correspond to the Gaussian model.

$$\begin{aligned}
S[\bar{\phi}, \phi] &= \int \bar{d}\omega \int^\Lambda \bar{d}^d \mathbf{k} \quad (i\omega + D\mathbf{k}^2 + su) \bar{\phi}_{\mathbf{k},\omega} \phi_{\mathbf{k},\omega} \\
&\quad + su \int \bar{d}\omega \bar{d}\varepsilon \int^\Lambda \bar{d}^d \mathbf{k} \bar{d}^d \mathbf{q} \quad \bar{\phi}_{\mathbf{k},\omega} \phi_{\mathbf{k}-\mathbf{q},\omega-\varepsilon} \phi_{\mathbf{q},\varepsilon} \\
&\quad + \frac{u}{2n} \int \bar{d}\omega \bar{d}\varepsilon \int^\Lambda \bar{d}^d \mathbf{k} \bar{d}^d \mathbf{q} \quad \bar{\phi}_{\mathbf{k},\omega} \bar{\phi}_{\mathbf{k}-\mathbf{q},\omega-\varepsilon} \phi_{\mathbf{q},\varepsilon} \\
&\quad - \frac{u}{2n} \int \bar{d}\omega_1 \bar{d}\omega_2 \bar{d}\varepsilon \bar{d}^d \mathbf{k} \bar{d}^d \mathbf{k}_2 \bar{d}^d \mathbf{q} \bar{\phi}_{\mathbf{k}_1,\omega_1} \bar{\phi}_{\mathbf{k}_2,\omega_2} \phi_{\mathbf{k}_1+\mathbf{q},\omega_1+\varepsilon} \phi_{\mathbf{k}_2-\mathbf{q},\omega_2-\varepsilon}
\end{aligned}$$

Once again, we can graphically represent this action using the vertices and edges of Figure (4.3). Using the same arguments as above, we find the recursion relations

$$\begin{aligned}
s' &= b^2 s \left(1 - \frac{u}{D} I_d\right), \\
u' &= b^{2-d} u \left(1 - \frac{u}{D} I_d\right),
\end{aligned}$$

with I_d the same as above.

For $d = 1$, we have the differential recursion relations

$$\begin{aligned}
\frac{ds}{dl} &= 2s - \frac{su}{\pi D \Lambda}, \\
\frac{du}{dl} &= u - \frac{u^2}{\pi D \Lambda}.
\end{aligned}$$

These have the non-trivial fixed point $s^* = 0$, $u^* = D\pi\Lambda$.

For $d = 2$, the RG flow is

$$\begin{aligned}
\frac{ds'}{dl} &= 2s - \frac{us}{\pi D}, \\
\frac{du'}{dl} &= -\frac{u^2}{\pi D},
\end{aligned}$$

and the only fixed point is at $s^* = 0$, $u^* = 0$.

For $d = 3$, we find

$$\frac{ds'}{dl} = 2s - \frac{1}{2} \frac{su\Lambda}{\pi D},$$

$$\frac{du'}{dl} = -u - \frac{1}{2} \frac{u^2 \Lambda}{\pi D},$$

and like the previous result, the only fixed point is at $s^* = 0$, $u^* = 0$.

The RG flow for this regime of the Stepping-Stone model, and for $d = 1$ and 2 is shown in Figure (??). The flow for $d = 3$ is qualitatively similar to that for $d = 2$. These results also correspond nicely to what we found in the previous chapter, namely, that the behavior of the model is qualitatively different for $d = 1$.

5.8 Determining critical exponents from RG eigenvalues.

In the previous section, we deduced the RG flow and found non-trivial fixed point for $d = 1$. Here we show how this result can be used to find ν , the relaxation time exponent calculated in the previous chapter.

Recall that the relaxation time τ is a divergent function of s ,

$$\tau = \tau(s) \sim s^{-\nu}. \quad (5.13)$$

Under a renormalization transformation parametrized by b , time and hence τ gets re-scaled by b^z . On the other hand, the RG analysis above also tells us how s gets re-scaled. This allows us to write

$$\tau(s) = b^z \tau(sb^{y_s}), \quad (5.14)$$

where y_s is the eigenvalue coming from the linearized RG flow near the non-trivial fixed point u^* . (At the Critical Branching fixed point, $y_s = 2$). We are free to choose b , and making the choice $b = s^{-1/y_s}$ results in the relation

$$\tau(s) = b^{-z/y_s} \tau(1), \quad (5.15)$$

and so we see that

$$\nu = \frac{z}{y_s}. \quad (5.16)$$

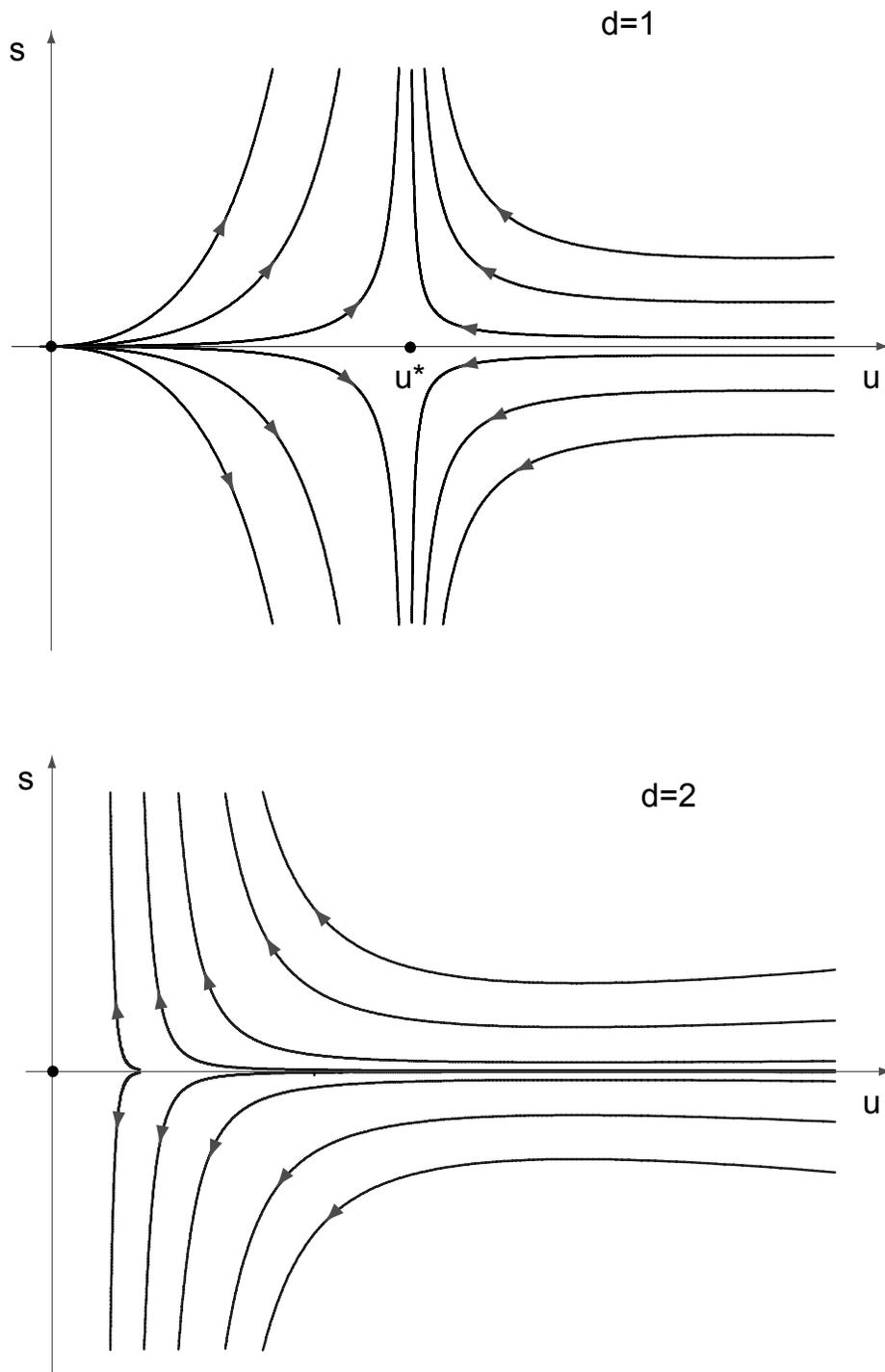


Figure 5.6: RG flow for the broken symmetry state. Top: For $d = 1$, the linearized RG flow around u^* determines the critical exponent $\nu = 2$.

We have related an eigenvalue of the RG to a critical exponent. The RG we performed in the previous section had $z = 2$. For the non-trivial fixed point, y_s is determined from the eigenvalues of the matrix

$$M = \left[\begin{array}{cc} \frac{\partial s'(u,s)}{\partial s} & \frac{\partial s'(u,s)}{\partial u} \\ \frac{\partial u'(u,s)}{\partial s} & \frac{\partial u'(u,s)}{\partial u} \end{array} \right]_{s=s^*, u=u^*}, \quad (5.17)$$

where $s'(u, s) = ds/dl$, $u'(u, s) = du/dl$. One finds

$$M = \left[\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array} \right], \quad (5.18)$$

therefore $y_s = 1$, and hence

$$\nu = 2, \quad (5.19)$$

in complete agreement with the calculation of chapter 4.

Chapter 6

Summary of Results and Conclusions

6.1 The Wright-Fisher and Moran Models.

In chapter 3 we introduced the Wright-Fisher Model, which for large population size is described by a Fokker-Planck equation with a heterogeneous diffusion term. We also introduced Moran's model, formulated using a master equation, and used a system size expansion to show that for large population size it obeys the same Fokker-Planck equation. For all intensive purposes, therefore, the Moran model is identical to the Wright-Fisher model. By examining the Fokker-Planck equation from the point of view of the effective potential, we showed that both the steady-state and dynamical behaviour were qualitatively different, depending on the value of $2N\mu$, where N is the population size and μ is the mutation rate. We also showed that for $2N\mu \ll 1$, the mean-field picture that results from ignoring the fluctuations becomes irrelevant. This observation is the basis of Kimura's Neutral Theory[8]. When there is no selective advantage of one allele relative to the other, and when the mutation rate vanishes, the eventual fixation of the population to one of the two equivalent absorbing states was shown to be an example of what is termed in statistical physics, a "spontaneously broken symmetry". Broken symmetry in the Wright-Fisher model was also shown to be associated with a kind of critical behaviour, that of the branching process. In contrast to the familiar examples in statistical physics, however, the critical behaviour in the Wright-Fisher model is accompanied by a discontinuous jump of the order parameter. Although we presented

no new results for the Wright-Fisher-Moran model, a review of the main features of this model from perhaps a slightly different point of view than is discussed in the population genetics literature was necessary in preparation for our analysis of the Stepping-Stone model.

6.2 The Stepping-Stone Model.

The main results of this thesis were exact calculations for steady-state and dynamical properties of the Stepping-Stone model. We examined two regimes of the model: the *neutral regime* and the *broken symmetry state*.

6.2.1 The neutral regime.

In the neutral regime there is no selective advantage of one allele over the other, and for non-zero mutation rate, a steady state is achieved that is characterized by the correlation length and the fixation index F_{ST} . The fixation index is the two-point correlation function evaluated at vanishing separation of the field that relates the difference between frequencies of the two alleles. In other words, F_{ST} is the variance of the frequency difference averaged over the sub-populations. When the mutation rate becomes vanishingly small, the correlation length diverges for all spatial dimensions d , but $F_{ST} \rightarrow 1$ only for $d \leq 2$. This indicates a broken symmetry of the type found in the Wright-Fisher model. The asymptotics of F_{ST} , including its time-dependence, were explicitly shown to be related to the return probability of a random walk and are also indicative of critical behaviour. However, for $d > 2$, the critical behaviour is not associated with symmetry breaking.

The above results for Stepping-Stone model are reminiscent of the Voter model, where “consensus” is reached for infinite lattices with $d \leq 2$ and is also a consequence of the

eventual return of the random walk[68]. In contrast to the Stepping-Stone model, F_{ST} is identically 1 in the Voter model since the site variable consists of a “spin”, $\sigma = \pm 1$. The fixation index F_{ST} is actually measured in natural populations (see ch.5 of [69] and references therein). Therefore the Voter model is of limited use for population genetics. Exact results have been given for the dynamics of reactive interfaces in the Voter model that are in rough agreement with our results, in that logarithmic behaviour is observed in $d = 2$ [70][71]. My results are also similar to those found by Sawyer[3] and Nagalaki[4] for the infinite alleles versions of the Stepping-Stone model and were discussed in section 4.4.2.

Most measurements of F_{ST} for natural populations are interpreted using Wright’s formula from his “Island model”, a model with infinite range migration[58]. The results I have found for the steady-state value of F_{ST} for the Stepping-Stone model suggest that this may not be appropriate for large population size and $d \leq 2$. I should point out that natural populations with $d > 2$ are quite unrealistic. The results of Eq. (4.33) are presented here using the genetics convention for a population size of $2N$,

$$F_{ST} = \begin{cases} \frac{1}{1+2n\sqrt{m\mu}}, & d = 1 \\ \frac{1}{1+8\pi mn/\ln(\frac{m}{2\mu})}, & d = 2 \\ \frac{1}{1+4\pi^2 nm}, & d = 3 \end{cases} \quad (6.1)$$

and are to be compared with the result from the Island model,

$$F_{ST} = \frac{1}{1 + 4nm}. \quad (6.2)$$

in Wright’s Island model. Only for $d = 3$ do the results qualitatively agree. In the limit of vanishing mutation rate μ , the results for $d \leq 2$ are quite different. One should also be careful in applying these results to $d = 2$ populations for very small μ , due to the extremely slow logarithmic relaxation of F_{ST} near the critical point. A final

point concerning F_{ST} and long-range migration is also in order. Long-range migration, for example Levý flights, will push the critical dimension for symmetry breaking lower than $d = 2$. (The infinite range migration of the Island model is an extreme limit where symmetry is not broken even for $d = 1$). Therefore one should expect long-range migration to have a very strong effect on F_{ST} for two-dimensional populations. In other words, species with long-range migration should exhibit greater polymorphism than those with finite-range migration.

6.2.2 The broken symmetry state.

I have calculated the relaxation time τ for long-wavelength perturbations due to rare, new mutant alleles with *selective disadvantage* s . The relaxation time τ diverges with s ,

$$\tau \sim s^{-\nu}. \quad (6.3)$$

The exponent ν characterizing the divergence of τ with s is identical to the non-spatial result $\nu = 1$ for $d \geq 2$. For $d = 1$, however $\nu = 2$. For small selective advantage, τ in the Stepping-Stone model is higher than that of the non-spatial Wright-Fisher model, the result of a “screening” of selective advantage due to the spatial structure of the population. The shift of τ is finite for $d = 3$, logarithmically dependent on s for $d = 2$, and very different for $d = 1$ as indicated by exponent $\nu = 2$ in Eq. (6.3).

Other critical exponents describing the spread of new alleles at the critical point $s = 0$ were easily obtained and found to be equal to their random walk counterparts. Scaling arguments were used to find the asymptotic behaviour of the survival probability $P_S(t)$ of the cascade of new neutral alleles. For $d = 1$ the non-classical value of ν leads to the result

$$P_S(t) \sim t^{-1/2},$$

while for $d \geq 2$, the asymptotic behaviour of $P_S(t)$ is that of a critical branching process,

$$P_S(t) \sim t^{-1}.$$

Although the relaxation time and survival probability are characterized by non-classical exponents for $d = 1$, the ultimate survival probability $P_\infty = \lim_{t \rightarrow \infty} P_S(t)$, or fixation probability for alleles with a *selective advantage* s' grows as

$$P_\infty \sim (s')^\beta, \tag{6.4}$$

with $\beta = 1$ for *all* dimensions. This “classical” value of β describes the fixation probability in the Wright-Fisher model as well as the ultimate survival probability for supercritical branching processes. One can summarize these results: Selectively disadvantageous new alleles tend to last longer and selectively advantageous alleles tend to take longer to reach fixation, and much more for $d = 1$ populations, but the fixation probability scales the same in all dimensions as well as in the non-spatial model. Unfortunately, I was unable to calculate the much desired prefactor in Eq. (6.4). To remind the reader, in the Wright-Fisher model $P_\infty \simeq 2s$ for small s and large population size[54][52]. This would be a nice factor to know as the fixation probability is a key ingredient to Kimura’s neutral theory.

6.3 Renormalization Group For Reaction-Diffusion and Related Stochastic Processes.

In preparation for the renormalization group analysis of the Stepping-Stone model, a good deal of effort was spent examining the general structure of the functional integral description of stochastic processes. The complexity of this elegant description, from the point of view of the renormalization group, manifested itself in the existence of several non-interacting fixed points. This fact has not been appreciated or even noticed in

the literature concerning the application of renormalization group methods to spatially-distributed stochastic models. The need for this distinction between non-interacting fixed points of the RG may be due to the specific wave-vector shell renormalization used by the author and therefore may not be necessary for others type of renormalization group approaches, e.g. real-space RG[72], field-theoretic RG[33], operator-product expansion RG[32] or “dynamic RG”[73]. On the other hand, understanding the structure of the general theory aids the application of RG methods to not only the Stepping-Stone model, but other more complicated population and reaction-diffusion models that defy exact results.

6.3.1 Renormalization to all orders.

Here I reiterate my argument of chapter 5 that the renormalization group analysis for the Stepping-Stone model is valid to all orders of perturbation theory and therefore represents a novel, exact application of the renormalization group . This is contrasted with the usual situation in which the RG is an expansion in $\varepsilon = d_c - d$. The reason for this is the existence of a small parameter, s that cuts off a large number of diagrams at the critical point $s = 0$. These diagrams are generated from the $s\bar{\phi}\phi^2$ term represented by the vertex in the bottom left of Figure (4.3). For example, in the model of directed percolation this vertex is non-zero at the critical point and hence the differential RG flow obtained by including only one loop diagrams is only valid for small $\varepsilon = d_c - d$.

6.3.2 Modern Theory of Stochastic Processes.

In addition to finding exact results for the Stepping-Stone model, I have decribed the synthesis of several theoretical formalisms of stochastic processes. These ranged from the classical Fokker-Planck and Langevin equations to the more modern description using operators and path integrals. Using Moran’s model and the Stepping-Stone model, I

demonstrated the equivalence of the several of these approaches. Hopefully this exercise will further clarify the extension of these formalisms to more complicated models.

Bibliography

- [1] M. Kimura and G. H. Weiss, *Genetics* **49**, 561 (1964).
- [2] G. H. Weiss and M. Kimura, *J. Appl. Prob.* **2**, 129 (1965).
- [3] S. Sawyer, *Ann. Prob.* **4**, 699 (1976).
- [4] T. Nagylaki, *Theoretical Population Biology* **10**, 70 (1976).
- [5] M. Kimura, *Cold Spring Harbor Symposium* **20**, 33 (1955).
- [6] S. Wright, *Genetics* **16**, 97 (1931).
- [7] R. A. Fisher, *Proc. Roy. Soc. Edin.* **42**, 321 (1922).
- [8] M. Kimura, *The Neutral Theory of Molecular Evolution*, Cambridge University Press, 1983.
- [9] M. Kimura, *Scientific American* **241**, 94 (1979).
- [10] J. F. Crow, *Science* **253**, 973 (1991).
- [11] J. A. Coyne, N. H. Barton, and M. Turelli, *Evolution* **51**, 634 (1997).
- [12] S. Rouhani and N. Barton, *Theor. Pop. Biol.* **31**, 465 (1987).
- [13] M. Kimura, *Ann. Math. Stat.* **28**, 882 (1957).
- [14] A. D. Fokker, *Ann. Physik* **43**, 810 (1914).
- [15] M. Planck, *Sitzber. Preuß Acad. Wiss.* , 324 (1917).
- [16] H. Risken, *The Fokker-Planck Equation, 2nd ed.*, Springer, 1989.
- [17] R. L. Stratonovich, *Topics in the Theory of Random Noise*, Gordon and Breach, New York, 1963.
- [18] N. S. Goel and N. Richter-Dyn, *Stochastic Models in Biology*, Academic, New York, 1974.
- [19] N. G. van Kampen, *Adv. Chem. Phys.* **34**, 245 (1976).

- [20] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, North-Holland, 1981.
- [21] L. Reichl, *A Modern Course in Statistical Physics*, Edward Arnold Ltd., 1981.
- [22] W. Horsthemke and L. Brenig, *Zeitschrift für Physik B* **27**, 341 (1977).
- [23] W. Horsthemke, M. Malke-Mansour, and L. Brenig, *Zeitschrift für Physik B* **28**, 135 (1977).
- [24] T. G. Kurtz, *J. Appl. Prob.* **8**, 344 (1971).
- [25] T. G. Kurtz, *J. Chem. Phys.* **57**, 2976 (1972).
- [26] P. R. Langevin, *C. R. Acad. Sci. (Paris)* **146**, 530 (1908).
- [27] N. G. van Kampen, *J. Stat. Phys.* **24**, 175 (1981).
- [28] K. Itô, *Proc. Imp. Acad.* **20**, 519 (1944).
- [29] R. L. Stratonovich, *Conditional Markov Process and Their Application to the Theory of Optimal Control*, Elsevier, New York, 1968.
- [30] S. K. Ma, *Modern Theory of Critical Phenomena*, Addison-Wesley, 1976.
- [31] P. C. Hohenberg and B. I. Halperin, *Rev. Mod. Phys.* **49**, 435 (1976).
- [32] J. Cardy, *Scaling and Renormalization in Statistical Physics*, Cambridge University Press, 1996.
- [33] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena, 2nd ed.*, Oxford University Press, 1993.
- [34] P. C. Martin, E. D. Siggia, and H. H. Rose, *Phys. Rev. A* **8**, 423 (1973).
- [35] H. K. Janssen, *Z. Physik B* **23**, 377 (1976).
- [36] R. Bausch, H. K. Janssen, and H. Wagner, *Z. Physik B* **24**, 113 (1976).
- [37] J. J. Binney, N. J. Dowrick, A. J. Fisher, and M. E. J. Newman, *The Theory of Critical Phenomena*, Oxford University Press, 1992.
- [38] N. Goldenfeld, *Lectures on Phase Transitions and the Renormalization Group*, Addison-Wesley, 1992.
- [39] R. D. Mattuck, *A guide to Feynman diagrams in the many-body problem, 2nd ed.*, McGraw-Hill, 1976.

- [40] M. Doi, J. Phys. A **9**, 1465 (1976).
- [41] M. Doi, J. Phys. A **9**, 1479 (1976).
- [42] Y. B. Zel'dovich and A. Ovchinnikov, Sov. Phys. JETP **47**, 829 (1978).
- [43] P. Grassberger and A. de la Torre, Fortschr. Phys. **28**, 547 (1980).
- [44] D. C. Mattis and M. L. Glasser, RMP **70**, 979 (1998).
- [45] A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics*, Prentice-Hall Inc., 1963.
- [46] J. W. Negele and H. Orland, *Quantum Many-Particle Systems*, Addison-Wesley, 1978.
- [47] D. L. Hartl and A. G. Clark, *Principles of Population Genetics*, Sinauer Associates Inc., 1997.
- [48] R. A. Fisher, *The Genetical Theory of Natural Selection*, Oxford Clarendon Press, 1930.
- [49] S. Wright, Ecology **26**, 415 (1945).
- [50] P. A. P. Moran, Proc. Camb. Phil. Soc. **54**, 60 (1958).
- [51] M. Kimura, Proc. Natl. Acad. Sci. USA **41**, 144 (1955).
- [52] M. Kimura, Genetics **28**, 713 (1962).
- [53] J. D. Murray, *Mathematical Biology*, Springer-Verlag, N.Y., 1993.
- [54] J. B. S. Haldane, Proc. Cambridge Phil. Soc. **23**, 838 (1927).
- [55] G. Malecot, Proc. Fifth Berkeley Symp. Math. Statist. Prob. **4**, 317 (1967).
- [56] S. Itatsu, Nagoy Math. J. **114**, 143 (1989).
- [57] T. Shiga and K. Uchiyama, Probab. Th. Rel. Fields **73**, 87 (1986).
- [58] S. Wright, Genetics **28**, 114 (1943).
- [59] W. Feller, *An Introduction to Probability Theory and Its Applications*, Wiley, New York, 1966.
- [60] S. Sawyer, J. Appl. Prob. **16**, 482 (1979).
- [61] M. L. Bellac, *Quantum and Statistical Field Theory*, Oxford University Press, 1991.

- [62] R. Dickman and A. Y. Tretyakov, *Phys. Rev. E* **52**, 3218 (1995).
- [63] E. Domany and W. Kinzel, *Phys. Rev. Lett.* **53**, 447 (1984).
- [64] P. Grassberger and M. Scheunert, *Ann. Phys. (NY)* **122**, 373 (1979).
- [65] K. G. Wilson and J. Kogut, *Phys. Rep.* **12C**, 75 (1974).
- [66] L. Peliti, *J. Phys. A* **19**, L365 (1986).
- [67] M. Droz and L. Sasvari, *Phys. Rev. E.* **48**, 2343 (1993).
- [68] T. M. Liggett, *Interacting Particle Systems*, Springer-Verlag, NY, 1985.
- [69] J. B. Mitton, *Selection in Natural Populations*, Oxford University Press, 1997.
- [70] L. Frachebourg and P. L. Krapivsky, *Phys. Rev. E.* **53**, R3009 (1996).
- [71] E. Ben-Naim, L. Frachebourg, and P. L. Krapivsky, *Phys. Rev. E.* **53**, 3078 (1996).
- [72] T. W. Burkhardt and J. M. J. van Leeuwen, *Real-Space Renormalization*, Springer-Verlag, 1982.
- [73] A.-L. Barabási and H. E. Stanley, *Fractal Concepts in Surface Growth*, Cambridge University Press, 1995.
- [74] G. S. Joyce, *J. Phys. A.* **5**, L65 (1972).
- [75] G. N. Milshtein, *Theory Prob. Appl.* **23**, 396 (1978).
- [76] J. Bentley, *Comm. ACM* **28**, 245 (1985).

Appendix A

Random Walks on the Linear, Square and Cubic Lattices.

A.1 Random Walk Green's Functions.

In this appendix, we calculate the Green's function and return probability for continuous time random walks. Most treatments of random walks in the mathematical physics literature deal with discrete time random walks, and even though these have similar asymptotics, there are some subtle differences to continuous time walks. We need the properties of continuous random walks for the analysis of the Stepping-Stone model in chapter 4.

Written in terms of creation-annihilation operators, a random walk on a lattice is described by the Liouville operator

$$\mathcal{L} = \alpha \sum_{\mathbf{r}} \sum_{\mathbf{e}} a_{\mathbf{r}+\mathbf{e}}^\dagger a_{\mathbf{r}} + a_{\mathbf{r}-\mathbf{e}}^\dagger a_{\mathbf{r}} - 2a_{\mathbf{r}}^\dagger a_{\mathbf{r}}. \quad (\text{A.1})$$

For hypercubic lattices, $\mathbf{r} = \sum_i^d n_i \mathbf{e}_i$, and the lattice vector \mathbf{e}_i is simply l times the unit vector in the i th direction. Here α is the migration rate or jump rate. We identify the α in Eq. (A.1) with D/l^2 , so that in the continuum limit D becomes the diffusion constant.

We will consider a lattice with a total number of sites $N = \prod_i^d N_i$ and assume periodic boundary conditions. We will also take $N_1 = N_2 = \dots = L$.

Fourier transforming to wave-vector space using

$$a_{\mathbf{r}} = N^{-1/2} \sum_{\mathbf{k}} a_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}}, \quad a_{\mathbf{r}}^\dagger = N^{-1/2} \sum_{\mathbf{k}} c_{\mathbf{k}}^\dagger e^{-i\mathbf{k}\cdot\mathbf{r}}, \quad (\text{A.2})$$

in which, $\mathbf{k} = \sum_i^d \frac{m_i \mathbf{b}_i}{N_i}$, with $|\mathbf{b}_i| = \frac{2\pi}{l}$, one finds

$$\mathcal{L} = -\alpha \sum_{\mathbf{k}} \omega_{\mathbf{k}} a_{\mathbf{k}}^{\dagger} a_{\mathbf{k}}, \quad (\text{A.3})$$

where

$$\omega_{\mathbf{k}} = \sum_{\mathbf{e}} (2 - e^{i\mathbf{k}\cdot\mathbf{e}} - e^{-i\mathbf{k}\cdot\mathbf{e}}). \quad (\text{A.4})$$

On hypercubic lattices in d dimensions,

$$\omega_{\mathbf{k}} = 4 \sum_{j=1}^d \sin^2 \frac{k_j l}{2}, \quad (\text{A.5})$$

For small kl , $\omega_{\mathbf{k}} \simeq \mathbf{k}^2 l^2$, so in the continuum limit

$$\mathcal{L}/D = -\frac{1}{(2\pi)^d} \int k^2 a_{\mathbf{k}}^{\dagger} a_{\mathbf{k}} d\mathbf{k}, \quad (\text{A.6})$$

The Heisenberg equations don't couple higher order operators,

$$\frac{da_{\mathbf{k}}}{dt} = [a_{\mathbf{k}}, \mathcal{L}] = -\alpha \omega_{\mathbf{k}} a_{\mathbf{k}} \quad (\text{A.7})$$

$$\frac{da_{\mathbf{k}}^{\dagger}}{dt} = [a_{\mathbf{k}}^{\dagger}, \mathcal{L}] = \alpha \omega_{\mathbf{k}} a_{\mathbf{k}}^{\dagger} \quad (\text{A.8})$$

So

$$a_{\mathbf{k}}(t) = a_{\mathbf{k}}(0) e^{-\alpha \omega_{\mathbf{k}} t}, \quad (\text{A.9})$$

$$a_{\mathbf{k}}^{\dagger}(t) = a_{\mathbf{k}}^{\dagger}(0) e^{\alpha \omega_{\mathbf{k}} t}.$$

Thus the propagator, in wave-vector space is

$$\begin{aligned} G_{\alpha}(\mathbf{k}, t_f - t_i) &= \left\langle S \left| a_{\mathbf{k}}(t_f) a_{\mathbf{k}}^{\dagger}(t_i) \right| 0 \right\rangle, \\ &= \left\langle S \left| a_{\mathbf{k}}(0) a_{\mathbf{k}}^{\dagger}(0) e^{-\alpha \omega_{\mathbf{k}}(t_f - t_i)} \right| 0 \right\rangle, \\ &= e^{-\alpha \omega_{\mathbf{k}}(t_f - t_i)} \left\langle S \left| a_{\mathbf{k}}(0) a_{\mathbf{k}}^{\dagger}(0) \right| 0 \right\rangle, \\ &= e^{-\alpha \omega_{\mathbf{k}}(t_f - t_i)} \end{aligned} \quad (\text{A.10})$$

Now we Fourier transform back to get

$$G_\alpha(\mathbf{r}, t) = \frac{1}{N} \sum_{\mathbf{k}} e^{-\alpha\omega_{\mathbf{k}}t + i\mathbf{k}\cdot\mathbf{r}}. \quad (\text{A.11})$$

Written in terms of indices,

$$\begin{aligned} G_\alpha(\{n\}, t) &= \frac{1}{N} \prod_{j=1}^d \sum_{m_j=0}^L \exp(-\alpha\omega_{k_j}t) \exp\left(2\pi\frac{m_j n_j}{L}\right), \\ &= \frac{1}{N} \prod_{j=1}^d \sum_{m_j=0}^L \exp\left(-4\alpha t \sin^2 \frac{\pi m_j}{L}\right) \exp\left(i2\pi\frac{m_j n_j}{L}\right), \\ &= \frac{e^{-2\alpha t}}{N} \prod_{j=1}^d \sum_{m_j=0}^L \exp\left(2\alpha t \cos \frac{2\pi m_j}{L}\right) \exp\left(i2\pi\frac{m_j n_j}{L}\right). \end{aligned} \quad (\text{A.12})$$

The random walk on the infinite lattice corresponds to the limit as $L \rightarrow \infty$. The multiple sum becomes a multiple integral by introduction of the angles $\theta_j = \frac{2\pi m_j}{L}$, with $d\theta_j = \frac{2\pi}{L}$.

Then

$$\begin{aligned} G_\alpha(\{n\}, t) &= \frac{1}{(2\pi)^d} \prod_{j=1}^d \int_0^{2\pi} d\theta_j \exp(-4\alpha t \sin^2 \theta/2) \exp(in_j\theta) d\theta_j, \\ &= \frac{e^{-2\alpha t}}{(2\pi)^d} \prod_{j=1}^d \int_0^{2\pi} d\theta_j \exp(2\alpha t \cos \theta) \exp(in_j\theta) d\theta_j. \end{aligned} \quad (\text{A.13})$$

Recalling the integral representation of the modified Bessel function of the first kind

$$I_n(x) = \frac{1}{2\pi} \int_0^{2\pi} e^{in\theta} e^{x \cos \theta} d\theta, \quad (\text{A.14})$$

one has

$$G_\alpha(\{n\}, t) = e^{-2\alpha t} \prod_{j=1}^d I_{n_j}(2\alpha t), \quad (\text{A.15})$$

$$I_n(x) = \frac{1}{2\pi} \int_0^{2\pi} e^{in\theta} e^{x \cos \theta} d\theta \quad (\text{A.16})$$

First we note that at any time $G_\alpha(\mathbf{r}, t)$ is normalized to 1. This follows from the identities

$$e^z = I_0(z) + 2I_1(z) + 2I_2(z) + \dots, \quad (\text{A.17})$$

and

$$I_n(z) = I_{-n}(z). \quad (\text{A.18})$$

We also see for large \mathbf{r} and t , the anticipated result for the diffusion propagator

$$G_D(\mathbf{r}, t) \simeq (4\pi t D)^{-d/2} \exp(-\mathbf{r}^2/4tD), \quad (\text{A.19})$$

which follows from the asymptotic properties of the modified Bessel function.

A.2 Relation of the first-passage time to the Green's function.

Using the Green's function of a continuous-time random walk, $G_\alpha(\mathbf{r}, t)$ gives the probability of finding a random walker at the lattice site \mathbf{r} at time t , given that the walker started at the origin at $t = 0$. Since this is not necessarily the first visit to \mathbf{r} , we define the *first-passage time distribution* $F_\alpha(\mathbf{r}, t) dt$, as the probability that the walker reaches \mathbf{r} for the first time in the interval $[t, t + dt]$.

The probability of eventually reaching the origin, or the *return probability*, is obtained by integrating the first-passage time,

$$Q(\mathbf{0}) = \int_0^\infty F_\alpha(\mathbf{0}, t) dt. \quad (\text{A.20})$$

We need to relate $F_\alpha(\mathbf{r}, t)$ to the Green's function $G_\alpha(\mathbf{r}, t)$. To do this we note that if the walker is a point \mathbf{r} at time t , it must have been there for the first time at some $t' < t$.

Therefore, one has the convolution

$$G_\alpha(\mathbf{r}, t) = \int_0^t G_\alpha(\mathbf{0}, t - t') F_\alpha(\mathbf{r}, t') dt', \quad (\text{A.21})$$

which is valid for $\mathbf{r} \neq \mathbf{0}$. For $\mathbf{r} = \mathbf{0}$, we must add a term for the probability $P(t)$ that the walker has never left $\mathbf{r} = \mathbf{0}$, *i.e.*

$$G_\alpha(\mathbf{r}, t) = P(t) \delta_{\mathbf{r}} + \int_0^t G_\alpha(\mathbf{0}, t - t') F_\alpha(\mathbf{r}, t') dt', \quad (\text{A.22})$$

Taking the Laplace Transform of Eq. (A.22) for $\mathbf{r} = \mathbf{0}$ results in

$$\tilde{F}_\alpha(\mathbf{0}, s) = 1 - \frac{\tilde{P}(s)}{\tilde{G}_\alpha(\mathbf{0}, s)} \quad (\text{A.23})$$

And since $Q(\mathbf{0}) = \tilde{F}_\alpha(\mathbf{0}, s)$,

$$Q(\mathbf{0}) = 1 - \frac{\tilde{P}(0)}{\tilde{G}_\alpha(\mathbf{0}, 0)}. \quad (\text{A.24})$$

For the Markovian random walks considered in the previous section, the time interval between successive jumps is exponentially distributed with the characteristic time αz , where z is the number of nearest neighbors and equal to $2d$ for hypercubic lattices considered here. The probability of never leaving the origin is therefore

$$P(t) = e^{-2\alpha dt}, \quad (\text{A.25})$$

and its Laplace transform

$$\tilde{P}(s) = \frac{1}{s + 2\alpha d} \quad (\text{A.26})$$

The result for $Q(\mathbf{0})$ is then

$$Q(\mathbf{0}) = 1 - \frac{1}{2\alpha d \tilde{G}_\alpha(\mathbf{0}, 0)} \quad (\text{A.27})$$

Under a change of time scale, or equivalently, $\alpha \rightarrow b\alpha$, the Laplace transform behaves as

$$\tilde{G}_{b\alpha}(\mathbf{0}, s) = b^{-1} \tilde{G}_\alpha(\mathbf{0}, s/b), \quad (\text{A.28})$$

Taking $b^{-1} = 2\alpha d$, we can write for the return probability

$$Q(\mathbf{0}) = 1 - \frac{1}{\tilde{G}_{1/2d}(\mathbf{0}, 0)} \quad (\text{A.29})$$

We see that the propagator for a random walk with rate $(2d)^{-1}$ is quite special, so in what follows we simply drop the subscript and refer to it and its Laplace transform as G and \tilde{G} . The fundamental quantity $\tilde{G}(\mathbf{0}, 0)$, which is of course independent of the jump rate, gives us the return probability through Eq. (A.29).

A.3 The Critical Dimension of the Random Walk.

The return probability is unity only if the integral of the Green's function $G(\mathbf{0}, t)$ diverges. The result (first proved by Polya for a random walk in discrete time) is that for $d \leq 2$, $Q(\mathbf{0}) = 1$, but $Q(\mathbf{0}) = \text{const}$ for $d > 2$. We will recover this result for our continuous-time random walks by explicitly calculating $\tilde{G}(\mathbf{0}, 0)$ below. So we see that $d = 2$ is special for random walks.

A.4 Asymptotic Behaviour of the Random Walk Green's Function $G_\alpha(\mathbf{0}, t)$.

It is not difficult to explicitly calculate $\tilde{G}_\alpha(\mathbf{0}, s)$ in $d = 1$ and $d = 2$, for the linear and square lattice, respectively. The case of $d = 3$ is a much harder integral but has been done by Joyce. Nevertheless, we first show how one can extract the asymptotic dependence of $\tilde{G}_\alpha(\mathbf{0}, s)$, for $s \rightarrow 0$, in arbitrary dimension, without explicitly calculating the return probability (for $d > 2$). This is accomplished by looking at the dependence of $G_\alpha(0, t)$ for large t . It is independent of the particular choice of lattice, *e.g.* triangular, hexagonal, f.c.c., and depends only on dimension d through

$$G_\alpha(\mathbf{0}, t) \sim (\alpha t)^{-d/2}. \quad (\text{A.30})$$

We also observe that for small t , $G_\alpha(0, t)$ is bounded, so any divergence from the integral over t , comes from the large t part of the integral. So $\tilde{G}_\alpha(\mathbf{0}, 0)$ diverges in $d = 1$ as $\lim_{T \rightarrow \infty} T^{1/2}$, and in $d = 2$ as $\lim_{T \rightarrow \infty} \ln T$, but approaches a constant C for $d > 2$.

The asymptotic behaviour of the Laplace transform of $G_\alpha(\mathbf{0}, t)$ is obtained from so-called 'Abelian' theorems and only depends on behaviour for large t given above. One

finds

$$\tilde{G}_\alpha(\mathbf{0}, s) \sim \begin{cases} (4\alpha s)^{-1/2}, d = 1 \\ \frac{\ln(4\alpha/s)}{8\alpha\pi}, d = 2 \\ \frac{C}{6\alpha} + \frac{B}{6\alpha}\sqrt{\frac{s}{6\alpha}}, d = 3 \end{cases}. \quad (\text{A.31})$$

The constant C in A.31 is of course related to the return probability by

$$C = \frac{1}{1 - Q(\mathbf{0})}. \quad (\text{A.32})$$

A.5 The Laplace Transform of $G_\alpha(\mathbf{0}, t)$.

Above we defined $G(\mathbf{r}, t)$ as the propagator of a random walk with rate $(2d)^{-1}$ and found that its time integral $\tilde{G}(\mathbf{0}, 0)$ was intimately related to the return probability of the random walk. Here we explicitly calculate the Laplace transform $\tilde{G}(\mathbf{0}, s)$ for hypercubic lattices in $d = 1, 2$ and 3 . A change of scale will then give us

$$\tilde{G}_\alpha(\mathbf{0}, s) = \frac{1}{2d\alpha} \tilde{G}\left(\mathbf{0}, \frac{s}{2d\alpha}\right),$$

From above,

$$G(\mathbf{0}, t) = e^{-t} [I_0(t/d)]^d$$

First one uses the integral representation of the modified Bessel function

$$I_0(z) = \frac{1}{2\pi} \int_0^{2\pi} d\theta \ e^{z \cos \theta}, \quad (\text{A.33})$$

to obtain

$$\begin{aligned} \tilde{G}(\mathbf{0}, s) &= \int_0^\infty e^{-st} e^{-t} I_0(t/d) dt \\ &= \frac{1}{(2\pi)^d} \int_0^\infty dt \int_0^{2\pi} d^d \theta \exp\left(-t + \frac{t}{d} \sum_{i=1}^d \cos \theta_i - st\right) \end{aligned} \quad (\text{A.34})$$

Integrating over t results in the multiple integral

$$\tilde{G}(\mathbf{0}, s) = \frac{1}{(2\pi)^d} \int_0^{2\pi} d^d \theta \frac{1}{s + 1 - d^{-1} \sum_{i=1}^d \cos \theta_i} \quad (\text{A.35})$$

A.5.1 The Linear Lattice.

For $d = 1$, we have

$$\begin{aligned}\tilde{G}(\mathbf{0}, s) &= \frac{1}{2\pi} \int_0^{2\pi} \frac{1}{s + 1 - \cos \theta} d\theta \\ &= \frac{1}{\sqrt{s(s+2)}}\end{aligned}\tag{A.36}$$

The corresponding quantity for the random walk with rate α is

$$\tilde{G}_\alpha(\mathbf{0}, s) = \frac{1}{\sqrt{s(s+4\alpha)}}\tag{A.37}$$

We see that for small s ,

$$\tilde{G}_\alpha(\mathbf{0}, s) \sim (4\alpha s)^{-1/2}.\tag{A.38}$$

For large s ,

$$\tilde{G}_\alpha(\mathbf{0}, s) \sim s^{-1},\tag{A.39}$$

reflecting the fact that $\tilde{G}_\alpha(\mathbf{0}, s) \rightarrow 1$ with slope 0 as $t \rightarrow 0$.

A.5.2 The Square Lattice.

For $d = 2$,

$$\begin{aligned}\tilde{G}(\mathbf{0}, s) &= \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} \frac{1}{s + 1 - \frac{1}{2}(\cos \theta_1 - \cos \theta_2)} d\theta_1 d\theta_2 \\ &= \frac{1}{2\pi} \int_0^{2\pi} \frac{1}{\sqrt{(2s + 2 - \cos \theta_2)^2 - 1}} d\theta_2 \\ &= \frac{1}{\pi(s+1)} K\left(\frac{1}{s+1}\right),\end{aligned}\tag{A.40}$$

where $K(z)$ is the Complete Elliptic Integral of the first kind, so

$$\tilde{G}_\alpha(\mathbf{0}, s) = \frac{1}{\pi(s+4\alpha)} K\left(\frac{4\alpha}{s+4\alpha}\right),\tag{A.41}$$

For small s

$$\tilde{G}(\mathbf{0}, s) \sim \ln(2\sqrt{2})/\pi + \ln(s^{-1})/2\pi \quad (\text{A.42})$$

and hence

$$\tilde{G}_\alpha(\mathbf{0}, s) \sim (4\alpha\pi)^{-1} \left(\ln(2\sqrt{2}) + \frac{1}{2} \ln(4\alpha/s) \right) \quad (\text{A.43})$$

A.5.3 The Cubic Lattice.

For $d = 3$, we have one of the three so-called ‘Watson integrals’

$$\tilde{G}(\mathbf{0}, s) = \frac{1}{8\pi^3} \int_0^{2\pi} \int_0^{2\pi} \int_0^{2\pi} \frac{1}{s + 1 - \frac{1}{3}(\cos\theta_1 - \cos\theta_2 - \cos\theta_3)} d\theta_1 d\theta_2 d\theta_3 \quad (\text{A.44})$$

The other two ‘Watson integrals’ give the corresponding random walk propagator for the body centered and face centered cubic lattices. These integrals originally were derived to analyze the discrete time random walk, but as we have found, also pertain to the continuous time random walk discussed in this thesis. A closed form expressions for this integral has been given by Joyce [74],

$$\tilde{G}(\mathbf{0}, s) = \frac{4}{\pi^2} \frac{1}{s + 1} \frac{\sqrt{1 - \frac{3}{4}x_1}}{1 - x_1} K(k_+) K(k_-), \quad (\text{A.45})$$

with

$$k_\pm^2 = \frac{1}{2} \pm \frac{1}{4}x_2\sqrt{4 - x_2} - \frac{1}{4}(2 - x_2)\sqrt{1 - x_2} \quad (\text{A.46})$$

and

$$x_1 = \frac{(s + 1)^2 + \frac{1}{3} - \frac{1}{3}\sqrt{s(3s + 2)(3s + 4)(s + 2)}}{2(s + 1)^2}, \quad (\text{A.47})$$

$$x_2 = \frac{x_1}{x_1 - 1}$$

For small s ,

$$\tilde{G}(\mathbf{0}, s) \sim C + B\sqrt{s},$$

where

$$\begin{aligned} C &= \frac{6\sqrt{2}}{\pi^2} K\left(\sqrt{\frac{1}{2} - \frac{1}{2}\sqrt{6} - \sqrt{3}}\right) K\left(\sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{6} - \sqrt{3}}\right) \\ &= 1.516386059 \end{aligned} \quad (\text{A.48})$$

and

$$B = \frac{3}{2}C = 2.2745790885 \quad (\text{A.49})$$

Then for $d = 3$ and small s , we have

$$\tilde{G}_\alpha(\mathbf{0}, s) \sim \frac{C}{6\alpha} + B\sqrt{\frac{s}{6^3\alpha^3}}, \quad (\text{A.50})$$

Appendix B

Computer Simulations

This appendix provides additional details on the computer simulations presented in chapter 3 and 4. These simulations were written using C++ and compiled for the Windows, Unix and Linux operating systems.

B.1 Stochastic Differential Equations.

Stochastic differential equations (s.d.e.'s) are easily simulated on the computer using techniques similar to those developed for integration of ordinary differential equations (o.d.e.'s). The only complications involve questions of accuracy and efficiency. The simulations in Figure (3.5) were produced using a method of second-order accuracy described by Milshstein.[75].

B.2 Continuous-time Reaction Diffusion Processes.

For simulations of the Stepping-Stone model, a C++ library for the simulation of continuous time reaction-diffusion processes was developed by the author. This library consists of classes representing various fundamental reaction processes (for example $A + B \rightarrow A + A$, $A + B \rightarrow B + B$, $A \rightarrow A + A$ and $B \rightarrow 0$), in addition to random walk processes ($A_{\mathbf{r}} \rightarrow A_{\mathbf{r}+\mathbf{e}}$), and exchange processes ($A_{\mathbf{r}} + B_{\mathbf{r}+\mathbf{e}} \rightarrow A_{\mathbf{r}+\mathbf{e}} + B_{\mathbf{r}}$). Not all of these processes were necessary for the Stepping-Stone model, however, the library was designed in such a way that programs for the simulation of other lattice reaction-diffusion processes

could easily be implemented without rewriting the core code. We will not present all of the details of the class library in this appendix. Instead, we will point out the important differences of our simulations for continuous time reaction-diffusion models with those commonly used in statistical physics. The main differences are two-fold.

First of all, most simulations in the literature use a spin-like variable at each site which has at most several possible values. The theoretical formulation of birth-death type processes we use is the occupation number formalism, which describes classical “Bosons”. By this we mean that there can be an arbitrary number at each site. Therefore we need simulations where the site variable is an integer, and more importantly, not fixed to a particular value. Two-species models have two integers for site variables. Furthermore, the fundamental processes divide naturally into two classes: those that are purely on-site—the reaction processes, and those that couple sites—diffusion and exchange.

The second major difference involves time. In the standard Monte-Carlo method, time is discrete and measured in steps. At each step, one tries to change the configuration according to some allowed “rules”. The new configuration is either accepted or discarded according to probabilities that depend on the difference of the new configuration with the old configuration. For example, when simulating equilibrium systems like the Ising model, at each step a site is selected and the spin is flipped. If the new configuration has a lower energy ($\Delta E < 0$), it is accepted. If it has a higher energy ($\Delta E > 0$), it is accepted with a probability $e^{-\Delta E/k_b T}$. Once again, we strive to make the simulation fit the theoretical description, which is for continuous time Markov processes where the “event” times are exponentially distributed with a characteristic time for each event that depends on the local configuration. Using the Monte-Carlo method, events cannot happen simultaneously. Furthermore, the Monte-Carlo method is inefficient for simulations near a critical point where ‘critical slowing down’ makes the rejection rate extremely high. In reaction-diffusion models, the critical point is often associated with

an absorbing state, so using the Monte-Carlo method is wasteful because “dead” sites are still picked at random and rejected. For these reasons, a technique was developed for the simulation of continuous time processes.

The main idea for the simulation engine, was to keep track of the next event for each site. The simulation is first initialized by iterating through the lattice, site by site, and picking an allowed event from set of the fundamental processes that specify the model. The occurrence time of the event is also recorded and the times of the events from all the sites are sorted. After this is completed, the simulation begins by executing the earliest timed event. This event changes the configuration of its associated site, in addition to connected sites, depending on the nature of the event. A new event is generated for this site as well as those coupled to it. The times are again sorted and the procedure is iterated.

The reordering of times, especially for the whole lattice seems somewhat inefficient: the time cost in sorting an unsorted collection of N items grows as $O(N^2)$. The solution is to use a “heap”, or “priority queue” to maintain the items such that the earliest time is on top. A heap is a special type of binary tree that maintains a set of elements under the operations of inserting new elements and extracting the smallest element[76]. The worst case initial sort using a heap is $O(N \log N)$, but more importantly, the operation of removing the smallest, inserting a new element and reordering the tree requires a time cost of $O(\log N)$. This provides an efficient way to implement the continuous time simulations. For this purpose, a class known as a “dynamic priority que” was developed as the core component of the simulation engine.

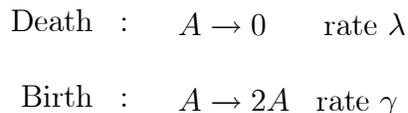
Appendix C

The Branching Process

The branching process is perhaps the simplest stochastic birth-death process, however, it is necessary to understand its properties, especially its critical behaviour. This reason is that many critical points in stochastic population models behave as a critical branching processes in the thermodynamic or large system size limit. A even more compelling motivation is that the critical points of spatial versions of these models above their upper critical dimension exhibit the same critical exponents as the simple branching process. Thus, the RG fixed point corresponding to the critical branching process plays the same role in these transitions as does the Gaussian fixed point in equilibrium statistical mechanics.

C.1 The model.

Consider a system of ‘ A ’-particles or individuals subject to the Poisson-distributed random events



The state of zero individuals is an absorbing state, so the system hits this state it stays there.

The corresponding master equation is

$$\frac{\partial P_n(t)}{\partial t} = \gamma(n-1)P_{n-1}(t) + \lambda(n+1)P_{n+1}(t) - (\lambda + \gamma)nP_n(t) \quad (\text{C.1})$$

Suppose we start with one individual at $t = 0$. We'd like to know something about the sizes and durations of the resulting cascades. The interesting quantities therefore are

1. *The survival probability.*

$$P_S(t) = 1 - P_0(t),$$

or the number of cascades (or “avalanches”) lasting longer than time t , and in particular, its asymptotic value

$$P_\infty = \lim_{t \rightarrow \infty} P_S(t)$$

2. *Cumulative size distribution $n(s)$* , the number of cascades larger than size s , or the probability that a cascade involves at least s individuals.

One finds critical behaviour when $\Delta \equiv \gamma - \lambda = 0$, *i.e.* all of these quantities behave as power laws

$$P_S(t) \sim t^{-\delta}$$

$$P_\infty \sim \Delta^\theta$$

$$n(s) \sim s^{-\alpha}$$

C.2 Survival probability.

The survival probability can be found by using the generating function

$$F(z, t) = \sum_{n=0}^{\infty} P_n(t) z^n \quad (\text{C.2})$$

From the generating function, we can find the moments by taking derivative w.r.t z and then setting $z = 1$. For example

$$\langle n \rangle = \left. \frac{\partial F(z)}{\partial z} \right|_{z=1}$$

and

$$\langle n^2 \rangle - \langle n \rangle = \left. \frac{\partial^2 F(z)}{\partial z^2} \right|_{z=1}$$

often the moments

$$\phi_m = \langle n(n-1)(n-2) \cdots (n-m+1) \rangle = \left. \frac{\partial^m F(z)}{\partial z^m} \right|_{z=1} \quad (\text{C.3})$$

are referred to as ‘factorial’ moments.

From the factorial moments, one can always get back the probabilities

$$P_n = \frac{1}{n!} \sum_{m=n}^{\infty} \frac{(-1)^k}{m!} \phi_m \quad (\text{C.4})$$

We can use Eq. (C.1) to obtain an equation for the generating function for this birth-death process. By multiplying by z^n and summing over n , one finds

$$\frac{\partial F(z, t)}{\partial t} = (z-1)(\gamma z - \lambda) \frac{\partial F(z, t)}{\partial z} \quad (\text{C.5})$$

This can be solved using the *method of characteristics*.

C.2.1 The method of characteristics.

Let $[z(\tau), t(\tau)]$ be some curve in the (z, t) plane. The change in F along this curve is

$$\frac{dF}{d\tau} = \frac{\partial F}{\partial t} \frac{dt}{d\tau} + \frac{\partial F}{\partial z} \frac{dz}{d\tau} \quad (\text{C.6})$$

We want to choose a curve such that

$$\left(\frac{dz}{d\tau} \right) / \left(\frac{dt}{d\tau} \right) = -(z-1)(\gamma z - \lambda)$$

Eq. (C.5) implies $F(z, t)$ is constant along this *characteristic*. Integrating

$$\frac{dz}{dt} = -(z-1)(\gamma z - \lambda)$$

we find

$$C = \frac{z-1}{\gamma z - \lambda} e^{(\gamma-\lambda)t}$$

where C is some constant of the curve $z(t)$. Then $F(z, t)$ must be some function of C . If we consider initial conditions such that at $t = 0$, the number of individuals was exactly m , or $P_n(0) = \delta_{n,m}$, we have

$$F(z, 0) = z^m$$

One finds the solution

$$F(z, t) = \left[\frac{e^{(\lambda-\beta)t}(\gamma z - \lambda) + \lambda(1-z)}{e^{(\lambda-\beta)t}(\gamma z - \lambda) + \gamma(1-z)} \right]^m \quad (\text{C.7})$$

This complicated looking expression tells us everything we need to know about this birth death process. Expanding in a power series, we can find $P_n(t)$. The survival probability is

$$P_S(t) = \frac{\lambda - \gamma}{e^{(\lambda-\gamma)t}\lambda - \gamma} \quad (\text{C.8})$$

For $\gamma < \lambda$,

$$P_\infty = 0$$

the population becomes extinct. For $\lambda < \gamma$

$$P_\infty = 1 - \frac{\lambda}{\gamma} \quad (\text{C.9})$$

$$= \frac{\Delta}{\gamma} \quad (\text{C.10})$$

For the critical case, $\lambda = \gamma$, one finds

$$P_S(t) = \frac{1}{1 + \gamma t} \quad (\text{C.11})$$

The result for the critical point can also be found by solving the (factorial) moment equations. Using Eq. (C.3), one can show that

$$\frac{d\phi_m(t)}{dt} = \frac{\gamma k!}{(k-2)!} n_{m-1}$$

and that for 1-initial particle boundary conditions, these equations can be recursively integrated to find

$$\phi_m(t) = m! (\beta t)^{m-1},$$

which is then substituted into Eq. (C.4) to obtain Eq. (C.11).

C.2.2 The backward equation.

There's actually an easier way to get the duration results. For this method, we need conditional probabilities.

$$P_{n/m}(t_2, t_1) \equiv \text{prob. of finding } n \text{ at time } t_2 \\ \text{given } m \text{ at time } t_1.$$

The forward equation reads,

$$\frac{dP_{n/m}(t_2, t_1)}{dt_2} = \sum_{n'} w_{n,n'} P_{n'/m}(t_2, t_1) - w_{n',n} P_{n/m}(t_2, t_1),$$

or with $t = t_2 - t_1$,

$$\frac{dP_{n/m}(t)}{dt} = \sum_{n'} w_{n,n'} P_{n'/m}(t) - w_{n',n} P_{n/m}(t).$$

The backward equation is,

$$\frac{dP_{n/m}(t_2, t_1)}{dt_1} = - \sum_{m'} w_{m',m} P_{n/m'}(t_2, t_1) + w_{m',m} P_{n/m}(t_2, t_1),$$

or with $t = t_2 - t_1$,

$$\frac{dP_{n/m}(t)}{dt} = \sum_{m'} w_{m',m} P_{n/m'}(t) - w_{m',m} P_{n/m}(t).$$

Note that we can obtain the master equation by multiplying the both sides of the forward equation by P_m and summing over m , so that one only need worry about state probabilities at time t_2 and not conditional probabilities. For this simple branching process, the backward equation for 1-initial individual boundary conditions is

$$\frac{dP_{n,1}(t)}{dt} = -(\lambda + \gamma) P_{n/1} + \gamma P_{n/2} + \lambda P_{n/0}$$

Using the generating functions

$$F_1(z) = \sum_{n=0}^{\infty} P_{n/1}(t) z^n$$

$$F_2(z) = \sum_{n=0}^{\infty} P_{n/2}(t) z^n$$

the backward equation becomes

$$\frac{\partial F_1(z, t)}{\partial t} = -(\lambda + \gamma) F_1 + \gamma F_2(z, t) + \lambda,$$

but since

$$F_m(z, t) = [F_1(z, t)]^m$$

$$\frac{\partial F_1(z, t)}{\partial t} = -(\lambda + \gamma) F_1 + \gamma F_1^2(z, t) + \lambda,$$

subject to the initial condition

$$F_1(z, 0) = z$$

hence

$$F_1(z, t) = \frac{e^{(\lambda-\gamma)t}(\gamma z - \lambda) + \lambda(1-z)}{e^{(\lambda-\gamma)t}(\gamma z - \lambda) + \gamma(1-z)}$$

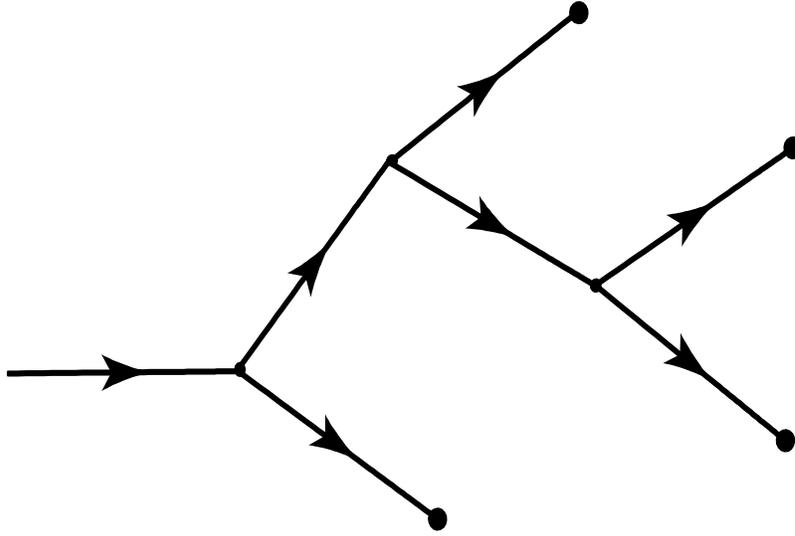


Figure C.1: A typical tree from a branching process

C.3 Size distribution.

In order to find the size distribution, we look at a generic tree

The probability of a tree with n leaves is

$$\begin{aligned}
 Q_n &= \left(\begin{array}{c} \# \text{ of trees} \\ \text{with } n \text{ leaves} \end{array} \right) \times [\text{Prob}(\text{death})]^{\# \text{ leaves}} \times [\text{Prob}(\text{branch})]^{\# \text{ vertices}} \\
 &= (\text{Cayley's formula}) \times \left(\frac{\lambda}{\lambda + \gamma} \right)^n \times \left(\frac{\gamma}{\lambda + \gamma} \right)^{n-1} \\
 &= \frac{1}{2n-1} \binom{2n-1}{n} \left[\frac{\lambda\gamma}{(\lambda + \gamma)^2} \right]^n \left(\frac{\lambda + \gamma}{\gamma} \right)
 \end{aligned}$$

Using Stirling's formula

$$Q_n \simeq \frac{1}{2\sqrt{\pi}} n^{-3/2} \left[\frac{4\lambda\gamma}{(\lambda + \gamma)^2} \right]^n \left(\frac{\lambda + \gamma}{2\gamma} \right)$$

or

$$Q_n \simeq \frac{1}{2\sqrt{\pi}} n^{-3/2} e^{-n/\zeta} \tag{C.12}$$

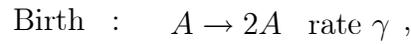
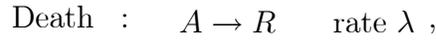
with the characteristic size

$$\zeta = -\frac{1}{\ln\left(\frac{4\lambda\gamma}{(\lambda+\gamma)^2}\right)} \quad (\text{C.13})$$

So the cumulative size distribution

$$n(s) \sim s^{-1/2}$$

We briefly mention another way to calculate the size distribution by using a slightly modified model. Suppose we write a master equation for the process corresponding to



where R is a ‘removed’ individual. Using a two species generating function

$$F(y, z) = \sum_{n,l} P_{n,l} y^n z^l ,$$

one can use either the backward or forward equation and the relevant method above to solve for the $t \rightarrow \infty$, R probabilities, defined by $Q_l = \sum_n P_{n,l}$.