

The Statistical Dynamics of Epochal Evolution

De Statistische Dynamica van Gefaseerde Evolutie
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de Rector Magnificus, Prof. Dr. H.O. Voorma,
ingevolge het besluit van het College voor Promoties
in het openbaar te verdedigen op
maandag 27 september 1999 des middags te 4.15 uur

door

Erik Jan van Nimwegen

geboren op 5 november 1970
te Amsterdam.

Promotores:

Prof. P. Hogeweg
Faculteit Biologie
Universiteit Utrecht.

Prof. J. P. Crutchfield
Santa Fe Institute

Promotie Commissie:

Prof. B. Derrida
Laboratoire de Physique Statistique
Département de Physique
Ecole Normale Supérieure, Paris.

Prof. P. Schuster
Institut für Theoretische Chemie
Universität Wien.

Prof. J. Shapiro
Computer Science Department
University of Manchester.

The studies described in this thesis were performed at the Santa Fe Institute in Santa Fe, New Mexico, U.S.A. and at the department of theoretical biology and bioinformatics at Utrecht University. The investigations were financially supported by the Santa Fe Institute and by the Priority Program Nonlinear Systems of the Netherlands Organization for Scientific Research (NWO).

ISBN 90-393-2162-0

Keine Annahme scheint mir natürlicher, als daß dem Assoziieren oder Denken kein Prozeß in Gehirn zugeordnet ist; so zwar, daß es also unmöglich wäre, aus Gehirnprozessen Denkprozesse abzulesen. Ich meine das so: Wenn ich rede oder schreibe, so geht, nehme ich an, ein meinem gesprochenen oder geschriebenen Gedanken zugeordnetes System von Impulsen von meinem Gehirn aus. Aber warum sollte das *System* sich weiter in zentraler Richtung fortsetzen? Warum soll nicht sozusagen diese Ordnung aus dem Chaos entspringen? Der Fall wäre ähnlich dem—daß sich gewisse Pflanzenarten durch Samen vermehren so daß ein Same immer dieselbe Pflanzenart erzeugt, von der er erzeugt wurde,— daß aber *nichts* in dem Samen der Pflanze, die aus ihm wird, entspricht; so daß es unmöglich ist, aus den Eigenschaften oder der Struktur des Samens auf die der Pflanze, die aus ihm wird, zu schließen,—daß man dies nur aus seiner *Geschichte* tun kann. So könnte also aus etwas ganz Amorphem ein Organismus sozusagen ursachelos werden; und es ist kein Grund, warum sich dies nicht mit unserem Gedanken, also mit unserem Reden oder Schreiben etc. wirklich so verhalten sollte.

No supposition seems to me more natural than that there is no process in the brain correlated with associating or with thinking; so that it would be impossible to read off thought-processes from brain-processes. I mean this: if I talk or write there is, I assume, a system of impulses going out from my brain and correlated with my spoken or written thoughts. But why should the *system* continue further in the direction of the center? Why should this order not proceed, so to speak, out of chaos? The case would be like the following—certain kinds of plants multiply by seed, so that a seed always produces a plant of the same kind as that from which it was produced—but *nothing* in the seed corresponds to the plant which comes from it; so that it is impossible to infer the properties or structure of the plant from those of the seed that comes out of it—this can only be done from the *history* of the seed. So an organism might come into being even out of something quite amorphous, as it were causelessly; and there is no reason why this should not really hold for our thoughts, and hence for our talking, and writing.

Ludwig Wittgenstein, Zettel [145]

Contents

1	Introduction	1
1.1	Formalizing the Darwinian Paradigm	1
1.2	Simple Evolutionary Systems	4
1.3	Epochal Evolutionary Dynamics	5
1.4	Outline of the Thesis	8
2	Statistical Mechanics and Thermodynamics	13
2.1	Macroscopic Laws	14
2.1.1	Elasticity: Hooke's Law	15
2.1.2	Electric current: Ohm's Law	15
2.1.3	Diffusion: Fick's Law	16
2.1.4	Ideal Gas: Equation of State	16
2.2	Equilibrium Statistical Mechanics	17
2.2.1	The Maximum Entropy Method	17
2.2.2	Is that all?	21
2.3	Phase Transitions	22
2.4	Non-equilibrium Statistical Mechanics	24
2.4.1	Memoryless Approximation	27
2.5	The Second Law	28
2.5.1	Definitions	29
2.5.2	Information Theoretic Viewpoint	30
2.5.3	Implications	33
3	Macroscopic Evolutionary Dynamics	35
3.1	Microscopic Description	35
3.2	Evolutionary Macrostates	38
3.2.1	Neutrality and the Macroscopic State Spaces	40
3.3	Infinite-Population Dynamics	41
3.3.1	Memoryless Approximation	42
3.3.2	The Generation Operator	43
3.4	Finite Population Dynamics	46
3.5	Metastability and Phase Space Unfolding	48
3.5.1	Unfolding Dimensions	50
3.5.2	Unfolding and Phase Transitions	52
4	Statistical Dynamics of the Royal Road Genetic Algorithm	55
4.1	Epochal Evolution	56
4.1.1	Search and Evolution	56
4.1.2	Organization of the Analysis	58
4.2	A Simple Genetic Algorithm on the Royal Road Fitness Function	59
4.2.1	The Fitness Function	59

4.2.2	The Genetic Algorithm	60
4.3	Observed Behavior of the Royal Road GA	60
4.4	The Royal Road GA's State Space	65
4.5	Infinite Population Dynamics on Fitness Distribution Space	68
4.5.1	Block Alignment Dynamics	69
4.5.2	The Mutation Operator M	70
4.5.3	The Selection Operator S	71
4.5.4	The Generation Operator G	72
4.5.5	Properties of the Generation Operator G	75
4.5.6	GA Dynamics as a Flow in Fitness Distribution Space	76
4.6	Finite Population Dynamics	78
4.6.1	Fitness Epochs	79
4.6.2	Predicted Epoch Levels	80
4.6.3	Epochal Dynamics as the Evolutionary Unfolding of the State Space	84
4.6.4	Eigensystem of the Restricted Generation Operators	85
4.6.5	Crossover	88
4.6.6	Stable and Unstable Manifolds	89
4.6.7	Innovation Durations	92
4.6.8	Fitness Fluctuations	94
4.6.9	Destabilizing Fluctuations and the Error Threshold	99
4.7	Epoch Durations	102
4.7.1	Creation of a Higher Fitness String	102
4.7.2	Takeover of the Population by a Higher Fitness String	103
4.8	Discussion	107
4.8.1	Low Mutation Rate Results	107
4.8.2	Metastability, Unfolding, and Landscapes	109
4.8.3	Future Work	111
5	Optimizing Epochal Evolutionary Search:	
	Population-Size Independent Theory	113
5.1	Engineering Evolutionary Search	113
5.2	Landscape Architecture	114
5.3	The Royal Staircase Fitness Function	118
5.4	The Genetic Algorithm	119
5.5	Observed Population Dynamics	120
5.6	Statistical Dynamics of Evolutionary Search	123
5.6.1	Generation Operator	125
5.6.2	Finite-Population Dynamics	126
5.6.3	Epochal Dynamics	127
5.6.4	Selection Operator	128
5.6.5	Mutation Operator	129
5.7	Quasispecies Distributions and Epoch Fitness Levels	130
5.8	Quasispecies Genealogies and Crossover's Role	132
5.9	Mutation Rate Optimization	134
5.10	Fitness Function Evaluations	137

5.11	Theory versus Experiment	138
5.12	Conclusion and Future Analyses	141
6	Optimizing Epochal Evolutionary Search:	
	Population-Size Dependent Theory	145
6.1	Designing Evolutionary Search	145
6.2	Royal Staircase Fitness Functions	148
6.3	The Genetic Algorithm	150
6.4	Observed Population Dynamics	150
6.5	Statistical Dynamics of Evolutionary Search	153
	6.5.1 Macrostate Space	154
	6.5.2 The Evolutionary Dynamic	154
	6.5.3 Finite-Population Sampling	156
	6.5.4 Epochal Dynamics	157
6.6	Quasispecies Distributions and Epoch Fitness Levels	159
6.7	Mutation Rate Optimization	160
6.8	Epoch Destabilization: Population-Size Dependence	164
6.9	Theory versus Experiment	168
6.10	Search-Effort Surface and Generalized Error Thresholds	170
6.11	Conclusions	173
	6.11.1 Genetic Algorithms versus Hill Climbers	174
	6.11.2 Coarse Graining the Fitness Function	175
7	Neutral Evolution of Mutational Robustness	179
7.1	Introduction	179
7.2	Modeling Neutrality	180
	7.2.1 Infinite-Population Solution	181
	7.2.2 Blind and Myopic Random Neutral Walks	183
7.3	Mutational Robustness	184
7.4	Finite-Population Effects	185
7.5	RNA Evolution on Structurally Neutral Networks	185
7.6	Conclusions	188
8	Metastable Evolutionary Dynamics: Crossing Fitness Barriers or Escaping via Neutral Paths?	191
8.1	Introduction	191
	8.1.1 Evolutionary Pathways and Metastability	193
	8.1.2 Overview	194
8.2	Evolutionary Dynamics	195
8.3	Crossing a Single Barrier	196
	8.3.1 Metastable Quasispecies	197
	8.3.2 Valley Lineages	199
	8.3.3 Crossing the Fitness Barrier	202
	8.3.4 Additional Time in Valley Bushes	203
	8.3.5 Theory versus Simulation	204
	8.3.6 Scaling of the Barrier Crossing Time	207

8.4	The Entropy Barrier Regime	210
8.4.1	Error Thresholds	211
8.4.2	The “Landscape” Regime	212
8.4.3	Time Scales in the Entropic Regime	213
8.4.4	Anomalous Scaling	215
8.5	Traversing Complex Fitness Functions	217
8.5.1	The Royal Staircase with Ditches	218
8.5.2	Evolutionary Dynamics	219
8.5.3	Observed Population Dynamics	220
8.5.4	Epoch Quasispecies and The Statistical Dynamics Approach	221
8.5.5	Crossing the Fitness Barrier	223
8.5.6	Theoretical and Experimental Epoch Times	225
8.6	Conclusions	228
9	Conclusions	231
9.1	Summary of Results	232
9.2	Integrative Unfolding View	235
9.3	Neutrality in the Sea of Microscopic Variables	238
A	Information theory	241
A.1	Axiomatic Definition of Entropy	241
A.2	Average Number of Yes/No Questions	243
A.3	Mutual Information	243
B	The Method of Lagrange Multipliers	245
B.1	Derivation	245
B.2	Example	246
C	Uniqueness of the Asymptotic Fitness Distribution	249
D	Finite Population Dynamics Convergence in the Infinite Population Limit	251
E	Generation Operator and Quasispecies for the Royal Staircase	255
E.1	Selection Operator	255
E.2	Mutation Operator	255
E.3	Epoch Fitnesses and Quasispecies	256
F	Epoch Quasispecies for the Royal Staircase with Ditches	259
G	Publications	263
H	Curriculum Vitae	265
I	Samenvatting	267
	Bibliography	279

1

Introduction

1.1 Formalizing the Darwinian Paradigm

The Darwinian theory of evolution [26] of reproducing entities with heritable variation that effects their reproductive success is enormously flexible. In essence it is a paradigm for an explanative theory that can be applied to *any* system which contains objects

- that reproduce with variation.
- that have heritable characters.
- whose reproductive success is affected by their characters.

This flexibility and general applicability is the paradigm's strength, on the one hand, while on the other hand it leaves unspecified *how* exactly the theory of evolution should be mathematically formalized.

This is in contrast to the typical situation found in physics where theories are generally formulated in mathematical terms from the start. The analysis of the system under study is generally conceptually clear. The dynamics is rigorously defined at some microscopic level and one imagines that the (higher level) behavior of the system is in principle derivable from these microscopic equations of motion when combined with particular boundary conditions. One in general focuses on the task of deriving the behavior of interest from these underlying microscopic laws. This task is in itself highly nontrivial and may not always be accomplished—such as in chaotic dynamical systems, where closed-form solutions of the equations of motion are often not available. In other words, one explores, to the best of one's abilities, the consequences of rigorously defined microscopic equations of motion.

From this particular viewpoint, the Darwinian theory of evolution doesn't seem so much a *theory* of evolution but instead more of a *paradigm* for a theory of the evolution of life on earth. That is, with respect to the theory of evolution, the question is still of a conceptual nature: What is the "right" way to formalize the theory in a mathematical framework?

For example, which objects in the evolving system should one take as the fundamental reproducing entities when implementing the Darwinian paradigm? Should one

take species, individuals, genomes, genes, or other less obvious constructions, such as collections of interacting individuals? Indeed, precisely this question has been the source of heated debates among evolution theorists; [29] for instance argues that genes are *the* units at which natural selection acts. The main problem, in the opinion of the present author, is that there is no such *right* or *fundamental* level: the Darwinian paradigm is capable of describing evolutionary dynamics on many different levels without any of these providing a fundamental level of description by itself.

Surely, different genes are competing with each other within the genomes of organisms and in that sense these genes undergo Darwinian evolution. However, this evolutionary dynamics cannot be rigorously defined without taking into account interactions with the evolution on higher levels—such as between genes within a genome and on the level of individuals. Similarly, single-celled organisms appear to be undergoing Darwinian evolution, competing with each other for molecular resources. Cells, however, interact, exchange genetic material and may form intricate spatial patterns. It is thus clear that one cannot define evolutionary equations of motions that only have “single-cell types” as variables. One may also attempt to formalize Darwinian evolution on the level of different *species* of single-celled organisms that compete with each other. At an even higher level, collections of interacting cells form multicellular organisms that can be considered to undergo Darwinian evolution as well. Finally, one may even formalize evolution on the level of different *groups* of organisms. See for instance [144] and references therein for an overview of the *group selection* versus *individual selection* debate.

One typically finds that the theorist attempts to circumvent these issues by introducing the mathematical formalization directly at the level that happens to be of immediate interest to him. The equations of motion are then constructed at this level and the behavior at the level of interest can be directly studied. However, as the examples above point out, the dynamics might not have *integrity* at this chosen level, since it depends on interactions with the dynamics on different levels. That is, evolution takes place at different levels at the same time, and the evolutionary *equations of motion* on different levels of description are often interdependent [73].

To give an illustrative example of such cross-level interdependence, which can be found in [13]: In a spatial environment, hypercycles [35] of replicating molecules self-organize into rotating spiral waves of different molecular species. Because of this, selection not only takes place on the level of the individual molecules, but also on the level of the rotating spirals. The competition between these spirals feeds back on the evolutionary dynamics on the level of the individual replicating molecules. That is, due to the selection on the level of spirals, the selection on the level of the replicating molecules is changed. In a completely mixed environment, molecules with a high decay rate are out competed by molecules with a low decay rate. In a spatial environment, however, these fast decaying molecules give rise to spirals that rotate faster than those made up of molecules with a low decay rate. Since on the level of spirals the faster spirals out compete slow spirals, the selection on the level of molecules is reversed, i.e. the molecules with a high decay rate now out compete those with a low decay rate.

From the viewpoint of mathematical formalization such situations present fundamental problems. The dynamics on the level of the individuals cannot be understood without including the dynamics the level of the higher-order structures, i.e. the spir-

1.1 Formalizing the Darwinian Paradigm

als. At the same time, the dynamics on the level of these higher-order structures cannot be separated from the dynamics on the level of the individuals making it impossible to define “closed” equations of motion on either of these levels.

This problem seems to occur frequently in evolutionary theory. It would in principle be possible to define the evolutionary model at some fundamental level—i.e. involving such things as distributions of genotypes, where in space the individuals are located, what the states of the local environments are, which individuals interact with which, what the internal states of the different individuals are, etc. However, one is tempted to immediately move to the level of description of interest and invent ad hoc evolutionary equations of motion at this level. For instance, in mathematical population genetics, one is often interested in the evolution of the frequencies of certain genes in the population. The complicated influences of the rest of the genomes and the dynamics on the level of the individuals are often ignored and the evolutionary dynamics is simply defined on the level of the gene frequencies only. On the other end of the spectrum, models are constructed in which the dynamics on the level of genes is ignored and evolution is described at the level of frequencies of certain *adaptive phenotypic traits* in the population, see for instance [72, 129].

As said, I do not believe there is any one level that is *the* right level for describing evolutionary dynamics. The problem in evolution theory is not to formulate *new* dynamical laws for how natural systems behave at a *fundamental* level. Such laws already exist: they are the laws of physics. The problem is to investigate the consequences that emerge out of these physical laws in the context of biological systems. That is, we need to discover what higher level behavior is induced in biological systems by the underlying laws. The Darwinian evolution theory provides a theoretical paradigm that seems to capture the essence of most of this higher level behavior.

Fortunately, it is not necessary to formalize the Darwinian evolutionary dynamics at *one particular* level exclusively. Defining the equations of motion at a certain microscopic level doesn’t necessarily mean that one can only study its behavior *at that level*. It is possible to start at a certain microscopic level of description and to *infer* the dynamics at a higher level of interest. Precisely this reconstruction of the behavior at the level of interest from the underlying microscopic equations of motion is the goal of *statistical mechanics* in physics. The main mathematical methodology for systems in thermodynamic equilibrium was developed by Gibbs [57] and Boltzmann [14]. In the 1950s Jaynes generalized this methodology and put it on a stronger conceptual basis, which includes exposing its close connections with information theory [84]. This general methodology allows one to start at the fundamental microscopic equations of motion of a system, and derive macroscopic equations of motion for the dynamics at some other level of interest. In the opinion of this author, this methodology is important for evolutionary theory since it may bridge the conceptual gap between the precise microscopic equations of motion of an evolutionary system that do not address the questions that interest us, and the ad hoc macroscopic evolutionary models that are not properly founded in a realistic microscopic dynamics. Although this methodology is conceptually straightforward, its technical implementations for different situations are far from fully developed.

In this thesis, the benefits of this methodology are demonstrated by application to some very simple evolutionary systems that are of interest in their own right. In these

models, the microscopic evolutionary dynamics is precisely defined on the level of distributions of genotypes in the population. At the same time, we are only interested in the dynamical behaviors of the population at some much coarser level—such as the average and best fitness in the population. Although the exact microscopic dynamics is almost impossible to solve analytically in detail, we show how the approximate macroscopic dynamics may still be constructed from the underlying microscopic equations of motion by using the statistical mechanics methodology. Since we focus on dynamics as opposed to the stationary states typically studied in statistical mechanics, we denote this approach as *statistical dynamics*¹ The methodology in itself is general enough that we expect that it can be applied in more or less modified form to many different evolutionary systems including much more complicated and realistic ones.

1.2 Simple Evolutionary Systems

Apart from providing illustrative examples of how one can, starting from the microscopic equations of motion, derive the evolutionary dynamics on a macroscopic level of interest, the models in this thesis also have some more direct relevance for understanding the evolutionary dynamics of simple evolutionary systems. The particular evolutionary systems analyzed in this thesis, and some of the assumptions that we make in defining them, will now be briefly introduced.

All evolutionary systems that are studied in this thesis consist of populations evolving in a constant and homogeneous external environment. That is, for each individual in the population its reproductive success is only a function of its “type”. This assumption excludes a vast array of possible complications that typically occur in more realistic situations. To name a few of these: in many situations the reproductive success of an individual is not only a function of its own type but also of its interactions with individuals of other types. In these situations one often says that the selective environment is not externally imposed but is a function of the internal state of the system. One also talks of *frequency dependent selection*—see for instance [17, 39]—when the reproductive success of one type of individual depends on the frequencies with which other types occur in the population. In case where the different types evolve and determine each other’s reproductive success, the word *co-evolution* is also used to describe these situations—see for instance [124]. Even if selection is entirely externally determined one may still have that the reproductive success of an individual depends on where in time and space the individual finds itself, since the selective environment may vary in space and time. On top of these complications, individuals cannot generally be considered to be characterized by a particular “type”. That is, individuals can occur in many different states themselves and their reproductive success may depend on when and where the individual finds itself in which states. In other words, although an individual can be mainly identified with its *genotype*—the particular DNA sequence that constitutes its genome—it goes through a complicated developmental process of interactions with its environment that eventually leads to a *phenotype*—the actual shape, form, and behavior of the organism. These phenotypes generally show complicated behaviors and interactions with

¹This terminology was first coined in an evolutionary context by Crutchfield. See also [131, 132] for a development of approximately the same analytical approach under the same name.

their environment, which in turn determine the reproductive success of the genotype.

As noted, none of these complications are studied in this thesis. We will look at populations of individuals whose fitness is a constant function of their genotypes only and that reproduce with certain definite genetic operators acting on their genotypes. Such simple systems occur in practice only in very specific situations. For instance, one may think of recent experiments where bacteria of a single species are evolved in a constant environment for many thousands of generations [97]. *In vitro* experiments with replicating biomolecules are also examples of the kinds of systems that may be modeled by the systems studied in this thesis. Particular examples are the *in vitro* experiments with replicating RNA molecules [10, 102] that are closely related to the Eigen model of molecular evolution [33]. Even simpler cases of relevance are examples of *in vitro* evolution where biomolecules are evolved to perform a certain biochemical task, such as binding a particular ligand [135, 147]. Similar *in vitro* evolution experiments have also been simulated in detail on computers [46, 47, 76, 81]. These artificial evolution experiments are closely related to the systems studied in this thesis. In all these cases the dynamics is fairly well understood on a microscopic level: simple “individuals”, i.e. molecules, that can be identified with their genotype, reproduce and undergo simple genetic operators—such as mutation and recombination—in a constant selective environment.

Finally, the models studied here are examples of what are generally called *evolutionary algorithms*. Evolutionary algorithms are stochastic search methods that are based on the Darwinian evolution paradigm. Populations of candidate solutions, represented as “genotypes” in a computer, evolve under selection for solving the problem under consideration and undergo (random) genetic operations that introduce variation; see for instance [4, 18, 94, 103]. Although most of the interest in evolutionary algorithms is focused on the performance of these algorithms as search methods, they can also be regarded as simple prototype models for evolutionary dynamics. This is the way in which such evolutionary algorithms are studied in most parts of this thesis. Understanding the dynamics of such evolutionary algorithms will of course also lead to a better understanding of their optimization and search properties. This type of analysis is also presented, see chapters 5 and 6 in this thesis.

1.3 Epochal Evolutionary Dynamics

Now that some of the conceptual motivations behind the analysis in this thesis, as well as the natural and artificial systems that are most directly related to the models studied here, have been introduced, I finally introduce briefly the type of dynamical behaviors that this work focuses on.

As emphasized above, we are not interested in the precise dynamics on the microscopic level of distributions of genotypes at which the evolutionary dynamics is defined. Especially for the kinds of natural and artificial systems that are related to the evolutionary models that we study, it is natural to be interested only in the dynamics on the coarser level of average or best fitness in the population. For instance, when running an evolutionary algorithm, one is interested in the quality of solutions that are evolved and related statistics—such as the average time to evolve a solution of a certain quality.

In most of this thesis, the analysis thus attempts to predict the dynamics on the level of fitness in the population. In principle, the same type of analysis could be used to study the dynamics on the level of phenotypes instead of fitness or, even, on the level of certain phenotypic traits.

This thesis further focuses on a dynamical behavior in evolving populations that we have called *epochal evolution*. In epochal evolution, long periods of stasis in the population average of a macroscopic variable are punctuated by sudden “innovations” of rapid change. That is, the evolutionary dynamics shows metastable behavior on the level of the macroscopic variable of interest. In cases where fitness is the macroscopic variable of interest, innovations generally lead to a sudden increase of the average fitness in the population. On the level of phenotypes in the population, innovations can often be identified with the emergence of new phenotypic functionality.

Such epochal evolution has been observed in a wide class of evolutionary dynamical systems. First there is the common occurrence of *punctuated equilibria* in the fossil record [64]. On macro-evolutionary time scales the fossil data seems to indicate that phenotypic evolution occurs in short bursts that are separated by long periods of phenotypic stasis. Such behavior has recently also been observed on much shorter time scales in the evolution of bacterial colonies [37]. There, the epochal behavior occurred both on the level of fitness (reproduction rate) and a phenotypic trait (cell size). The behavior has also been observed in the evolutionary dynamics of replicating computer programs [1], where innovations correspond to the sudden increase of the reproduction rate of these computer programs. In computer simulations of *in vitro* evolution of biomolecules, such epochal evolution seems to be a common occurrence as well. For instance in cases where RNA sequences are evolving towards a certain secondary structure [47, 79]. And finally, it is also commonly observed in the behavior of evolutionary search algorithms, for instance in the evolution of cellular automata toward performing a certain computational task [22]. In those evolution experiments, the innovations correspond to sudden changes in the dominant types of computational strategies that occur in the population.

In general, the epochal behavior in natural and artificial evolutionary systems can be the result of a variety of different mechanisms. In this thesis, we focus on some mechanisms that we believe may be responsible for much of the observed metastability in evolutionary systems. These mechanisms for epochal evolution are now briefly introduced.

On an abstract level, the different mechanisms take on the same mathematical form, which results from an interplay between the finite sizes of evolving populations, and the expected *flow* of these populations through the state space of macroscopic variables. That is, from the original microscopic dynamics on the level of the genotypes one constructs an expected dynamics on the level of the macroscopic variables of interest. The state space on this macroscopic level then consists of all possible population distributions over these macroscopic variables—such as, fitness or phenotypic traits. For a finite population such a state space is a discrete lattice; the percentage in the population of a certain macrovariable must be a multiple of the inverse population size. Metastability occurs when the expected flow through this state space is *small* in comparison to the state space lattice spacing, which is equal to the inverse population size.

On a more concrete level, such small flows occur when the expected number of individuals of a certain macroscopic type that is produced per unit time is much smaller than

one. This is quite likely to happen for macroscopic types that are rare in genotype space, since genotype spaces are generally much larger than population sizes. That is, even if the population spreads uniformly through genotype space, it can only visit a minute fraction of this space, so that rare macroscopic types will not be generated for long periods of time. Still, there might be different reasons for the low rate of production of these rare macroscopic types. Most prosaically put, their production may be inhibited because evolutionary forces are keeping the population from approaching them or, alternatively, they may just be hard to find since the evolving population has no way of directing itself *towards* them. Let us briefly review these two distinct mechanisms in the context of populations evolving under selection and mutation in a constant environment.

The first of these two mechanisms is the most commonly evoked explanation for metastability in evolving populations and involves the notion of a *fitness barrier*. Assume that the most-fit genotype occurring in the population is fitter than any of its single mutant neighbors. Selection then concentrates the population around this *local optimum* in genotype space. In this way, the population is kept from even higher-fitness genotypes by a “valley” of low-fitness genotypes around the local optimum at which it is currently located. Note that in this fitness barrier mechanism, the population is metastable on the microscopic level of the genotypes themselves. The evolutionary dynamics of crossing such fitness barriers is the topic of chapter 8. In most of this thesis, however, it is the alternative mechanism—that in which higher fitness types are simply hard to find—that is studied.

The central concept involved in this alternative mechanism is the occurrence of *neutrality* in the genotype-to-phenotype and genotype-to-fitness mapping over the microscopic genotype space. The concept will be discussed in more detail later. For now it suffices to point out how it is involved in bringing about the epochal evolutionary dynamics. The main idea is that there are large degeneracies in the genotype-to-fitness mappings. That is, there are many more genotypes than there are distinct fitness types. Neutrality refers to the fact that (almost) any genotype in genotype space has one or more mutant neighbors that are *neutral* to it with respect to fitness. Instead of the fittest individuals in the population occurring at a local optimum in genotype space, trapping the population genotypically around this local optimum, neutrality causes the population to diffuse through neutral *subbasins* of equal fitness genotypes. Since genotypes with even higher fitness might be rare, the population may have to search through this subbasin for a long time until a *portal* is discovered to a genotype of higher fitness. Nothing keeps the population from visiting these higher fitness genotypes, they are just hard to find. We refer to this situation as an *entropy barrier*. Note also that in this case the metastability does not occur on the microscopic level of genotypes, but only occurs on a coarser level of phenotypes or fitness.

This mechanism for metastability received increasing attention recently [6, 47, 77, 79, 108]. A growing consensus has developed that neutrality is a common occurrence in genotype-to-phenotype and genotype-to-fitness mappings that typically occur in biological systems. The idea goes back to Kimura’s contention that most evolutionary change at the genotype level is selectively neutral [91]. More recently, “biological” genotype-to-phenotype mappings have been studied in much more detail. Large scale neutrality was observed both in the sequence to secondary-structure mapping of RNA [48, 65, 126] and in sequence to structure mappings of proteins [3, 9]. Neutrality seems to commonly oc-

cur in combinatorial optimization problems [22] as well, especially when there is some redundancy in the coding of candidate solutions.

Since neutrality seems to be so common in the fitness functions of evolutionary systems, we believe that the epochal evolutionary dynamics caused by the entropy barriers that it induces might be a common occurrence in experimental and natural systems, and play a role in the metastable behavior of many of the different evolutionary systems mentioned above. Much of the analysis in this thesis is concerned with studying in detail the dynamics of epochal evolution in selective environments that possess neutrality.

1.4 Outline of the Thesis

Chapter 2 gives a general introduction to the conceptual foundations of deriving macroscopic laws from underlying microscopic laws. Such “macroscopic” laws are very common in physics; any *useful* physical law that has direct applications is almost always a macroscopic (approximate) law. Historically, statistical mechanics might be considered the first theory that attempts to formalize deriving macroscopic laws from the underlying microscopic ones. It has been very successfully applied to systems in thermodynamic equilibrium. Unfortunately, the conceptual foundations of the statistical mechanics approach have been, and to a certain extent still are unnecessarily muddled. Since the statistical mechanics framework is going to be used in this thesis on systems relatively far removed from the kind of systems normally studied in statistical physics, it is important to get these conceptual issues laid out very transparently. This is what is being attempted in chapter 2. The exposition is based on the idea that statistical mechanics is a special example of a general method of statistical inference that goes under the name of the *maximum entropy method*. As will also be discussed, this methodology is intimately related to information theory. The chapter ends with some comments on the (ir)relevance of the second law of thermodynamics for evolutionary systems.

In chapter 3 we first describe in more detail the kind of evolutionary systems that will be studied in this thesis. In particular, we describe the general type of microscopic equations of motion that is obeyed by these systems. After that, we discuss what kind of macroscopic variables seem reasonable candidates in the context of evolutionary systems, and how the maximum entropy methodology is applied in these situations. In particular, we derive the general form of the macroscopic equations of motion that are used throughout the applications in this thesis. We then describe how metastability is induced by finite populations in this framework. Finally, this chapter describes the general view of the dynamics through macroscopic state space in epochal evolution that has arisen from the analysis in this thesis. Specifically, the sudden innovations in epochal dynamics correspond to the *unfolding* of new macroscopic, i.e. fitness or phenotypic, dimensions. This unfolding of a new dimension in macroscopic state space can, to a certain extent, also be identified with *phase transitions* through a dynamic symmetry breaking on the microscopic level of the dynamics, i.e. on the level of genotypes. In this way, the epochal evolution unfolds its macroscopic state space over time through evolutionary innovations. This view on the evolutionary dynamics seems especially nice since it can accommodate more naturally the conception of evolution as a dynamic open ended process that doesn't take place in an a priori defined state space, but builds its

effective state space as it goes along. As an aside, it is also consonant with the literal meaning of the word evolution as unfolding.

Chapter 4 presents in detail our statistical dynamics approach to a model of a simple evolutionary system. In particular, we study the dynamics on the level of fitness of a “genetic algorithm” on a set of fitness functions known as *Royal Road* fitness functions. These fitness functions divide the genotype space into a relatively small number of nested neutral subbasins, and the epochal dynamics is the result of the entropy barriers that are generated in this way. We analyze where in the space of fitness distributions the epochs occur and study the stochastic finite population dynamics in these areas using a linearization of the dynamics together with a diffusion equation approach. This allows us to calculate many quantitative features of the epochal dynamics on these sets of fitness functions such as fitness fluctuations during epochs, the shape of innovation curves between epochs, epoch stability and destabilization times. We also study and give approximations to the average durations of the metastable epochs. The quantitative theory is then used to understand a wide range of qualitatively different dynamical behaviors that evolving populations exhibit on this class of fitness functions under changes of the evolutionary parameters.

This analysis is then used and extended in chapters 5 and 6 to derive optimal evolutionary parameter settings for epochal evolutionary search. That is, we derive how to set mutation rates and population sizes such that evolving populations will on average reach the global optimum as quickly as possible. In these chapters we thus study the evolutionary dynamics particularly in the context of evolutionary search algorithms. To model the epochal evolution, we use a class of fitness functions that we have called *Royal Staircase* fitness functions, which are closely related to the Royal Road functions studied in chapter 4. These functions too consist of nested neutral subbasins of increasing fitness. Apart from providing quantitative predictions, the understanding of the general mechanisms involved in parameter optimization should have relevance to more general evolutionary-search settings. In chapter 5 we also study in more detail the role of crossover for evolutionary search in systems that show epochal evolution. Comparison of the *scaling* of the total number of fitness function evaluations needed to reach the global optimum with those for hill climbing algorithms are also made. In chapter 5 we focus mainly on the optimization of the mutation rate, while in chapter 6 we focus on the population size effects which turn out to be mainly related to destabilization of epochs due to finite population size. At the end of chapter 6 we discuss in more generality how the evolutionary dynamics effectively induces a *coarse graining* of the fitness function on which the population evolves. We relate this to the question of what kinds of fitness functions may be efficiently searched by an evolutionary algorithm.

In chapter 7 we move away from studying the dynamics of populations *between* different neutral subbasins and study the dynamics of an evolving population *within* a single neutral subbasin. Such neutral subbasins are generally referred to as *neutral networks* within the theory of molecular evolution. We describe a general neutral subbasin/network architecture by a graph embedded in genotype space and study how an evolving population distributes itself over such a neutral network. The analysis shows that this distribution is *independent* of evolutionary parameters such as mutation rate and population size and is a function only of the topology of the neutral network. Additionally, the theory shows that the population will evolve a *mutational robustness* (the

insensitivity of the phenotype to mutations) that is higher than if the population were to spread randomly over the neutral network. The independence of this mutational robustness to the evolutionary search parameters then allows us to classify neutral network topologies according to the mutational robustness that they generate. Finally, we discuss the possibility of *in vitro* or *in vivo* experiments that would allow one to infer the topology of neutral networks of biomolecules from simple population statistics.

In most of this thesis, we have studied metastability in the context of entropy barriers in genotype space. The general occurrence of fitness barriers and local optima are however the more traditionally invoked explanations for the occurrence of metastability in the population dynamics. In chapter 8 we turn to studying this mechanism for metastability so as to be able to compare it with the metastability caused by entropy barriers. In the first parts of the paper we define simple fitness functions that consist of a single local optimum, a fitness “valley” and a portal genotype. We then study in detail the dynamics of crossing this single fitness barrier between the local optimum and the portal using a branching process approach. We derive accurate analytical expressions for the barrier crossing time as a function of the evolutionary parameters and the parameters defining the fitness function, i.e. barrier height and width. We then also derive the rough scaling behavior of these barrier crossing times. Notably, we show that these barrier crossing times scale very differently from the barrier crossing times typically found in physics (such as the waiting time for a physical system to cross an energy barrier). We discuss how these different scalings arise and clarify the danger in the “landscape” metaphors so often used in evolution theory. We then study in more detail the dynamics of entropy barrier crossing and in particular show that the waiting times for entropy barrier crossing exhibit anomalous scaling in both population size and mutation rate.

In the second half of the paper we combine our previous statistical dynamics analysis with the fitness barrier crossing analysis to study the general metastable population dynamics in fitness functions that contain both fitness and entropy barriers. These fitness functions are called *Royal Staircase with Ditches* and provide a further extension of the class of fitness functions that we study. This wide class of fitness functions allows us to compare the fitness and entropy barrier crossing times directly within the same model setting. The analysis shows clearly how much faster entropy barrier crossing occurs than fitness barrier crossing. The analysis thus quantifies the intuitive idea that the possibility for further evolution of a population is largely determined by where neutral evolution may take it.

Finally, a short comment on the methodology followed in most of this thesis. Generally, we first define a set of microscopic evolutionary equations of motion whose macroscopic behavior on the level of, for instance, fitness we want to understand. After that, we present our theoretical approach that attempts to quantitatively describe and predict this macroscopic behavior. Finally, we compare our predictions with results from simulations of the precise microscopic equations of motion. In particular, apart from showing parameter values for which our predictions are in accordance with simulation data, we also study where in parameter space our predictions break down, and why they do so. We feel that this last step is essential for giving our results the power of generalization. By analyzing the regime of validity of the theory, we are immediately pointed to the *restrictions* on the theoretical results. Additionally, these restrictions may point the way to further improvements of the theory. Without this, it seems that theoretical developments

1.4 Outline of the Thesis

may start to suffer from inbreeding; it may become a theoretical game that relates to nothing whatsoever but itself.

