Stochasticity and variability in the dynamics and genetics of populations

Perez de Vladar, Harold

RIJKSUNIVERSITEIT GRONINGEN

Stochasticity and Variability in the Dynamics and Genetics of Populations

Proefschrift

ter verkrijging van het doctoraat in de
Wiskunde en Natuurwetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr F. ZWARTS,
in het openbaar te verdedigen op
vrijdag 25 september 2009
om 13:15 uur

door

HAROLD P. DE VLADAR

geboren op 18 september 1976
te Caracas, Venezuela

Promotores:

Prof. dr Ido Pen
Prof. dr Franz J. Weissing

Beoordelingscommissie:

Prof. dr Reinhard Bürger
Prof. dr Johan A.J. Metz
Prof. dr Eörs Szathmáry

To Leonardo Mateu

To the Organic Intellectuals

To the Sewer Poets.

Contents

List of Figures **xiii**

Samenvatting (NL) **xv**

Summary (EN) **xvi**

Resumen (ES) **xviii**

Acknowledgements **xx**

General Introduction **1**

I Population Dynamics **19**

1 Density Dependence as a Size-Independent Regulatory Mechanism **21**

1.1 INTRODUCTION 23
1.2 "MECHANICS" OF SIZE REGULATION 25
1.3 GENERALIZED RATES OF GROWTH 30
1.4 STABILITY ANALYSES 31
1.5 SCALING LAWS 41
1.6 DISCUSSION 46

1.7 CONCLUDING REMARKS 53

2 Determinism, Noise, and Spurious Estimations in a Generalized Model of Population Growth 57
2.1 INTRODUCTION 59
2.2 DISTRIBUTION OF THE GROWTH RATES 62
2.3 DENSITY REGULATION BY NOISE 66
2.4 ESTIMATIONS OF GROWTH PATTERNS 69
2.5 CONCLUDING REMARKS 73

II Population Genetics 75

3 Statistical Mechanics and the Evolution of Polygenic Quantitative Traits 77
3.1 INTRODUCTION 79
3.2 GENERAL ANALYSIS 84
3.3 DIRECTIONAL SELECTION, MUTATION, AND DRIFT 92
3.4 LOW MUTATION RATES: $4N\mu < 1$ 107
3.5 DISCUSSION 114

4 $\mathcal{G}$ spotted in *Rana temporaria*! 121
4.1 INTRODUCTION 123
4.2 THEORETICAL BACKGROUND 125
4.3 MATERIALS AND METHODS 129
4.4 RESULTS 132
4.5 DISCUSSION 136

5 Evolution of Polygenic Traits: Adaptation at Maximal Entropy. 143
5.1 INTRODUCTION 145
5.2 MECHANISMS OF QUANTITATIVE EVOLUTION . 147
5.3 ENTROPY MAXIMIZING THEORY 151

5.4 CONCLUDING REMARKS 165

6 Perspectives: Pleiotropic effects on the $\mathcal{G}$-matrix 169
6.1 STANDING AND MUTATIONAL VARIATION 170
6.2 SOME CONSEQUENCES OF PLEIOTROPY 175
6.3 MOVING ON WITH THE $\mathcal{G}$ MATRIX 181
6.4 CONCLUDING REMARKS 183

7 Perspectives: The evolution of quantitative characters under stabilizing selection, mutation, and drift. 187
7.1 INTRODUCTION 188
7.2 MAX-ENTROPIC APPROACH 191
7.3 SINGLE LOCUS DYNAMICS 196
7.4 FORMULATING POLYGENIC DYNAMICS 206
7.5 POSTSCRIPT ON STABILIZING SELECTION 213

III Synthesis 221

8 Synthesis 223
8.1 AN OVERVIEW OF THE CONCLUSIONS 224
8.2 FURTHER EVOLUTIONARY IMPLICATIONS 227
8.3 RESEARCH PERSPECTIVES 244

Bibliography 253

A Maximization of entropy or of mean fitness? 283
A.1 MAXIMUM ENTROPY 284
A.2 MAXIMUM FITNESS 285

B Notes on the analogy with Statistical Mechanics in Physics 291

C Estimations of the genetic architecture of *Rana temporaria* 297
C.1 GENETIC EFFECTS AND EFFECTIVE NUMBER OF LOCI 298
C.2 GENETIC ARCHITECTURE AND M.S.D. EQUILIBRIUM 303
C.3 PLEIOTROPIC STRUCTURES AND FITNESS DIFFERENTIALS (β) 305
C.4 EIGENSTRUCTURES OF THE $\mathcal{G}$ MATRICES 307

D Formulas for the quantitative variables 311
D.1 DIRECTIONAL SELECTION 312
D.2 STABILIZING SELECTION 322

List of Figures

1.1 Bifurcation diagram of growth rates 34
1.2 Size-rate dynamics; $\rho > 0, \theta > 0$ 36
1.3 Size-rate dynamics; $\rho > 0, \theta < 0$ 38
1.4 Size-rate dynamics; $\rho < 0, \theta < 0$ 40
1.5 Size-rate dynamics; $\rho < 0, \theta > 0$ 41

2.1 Distribution of the stochastic rates 60
2.2 Phase space of stationary regimes 65
2.3 Realizations of noisy rate dynamics 66
2.4 Size of populations under stochastic rates 70

3.1 Evolution of a trait with arbitrary initial allele frequencies 81
3.2 Equilibrium values of the observables 95
3.3 Dynamics of the mean trait with a single locus; directional selection and $N\mu > 1/4$ 99
3.4 Dynamics of the mean trait with a single locus; directional selection and $N\mu \sim 1/4$ 100
3.5 Dynamics of the mean trait with a single locus; directional selection and $N\mu < 1/4$ 102
3.6 Dynamics of the mean trait with a single locus; directional selection and changing μ 103
3.7 Dynamics of a polygenic trait with unequal effects 104

3.8 Accuracy of the dynamics of traits with unequal effects 105
3.9 Polygenic trait under fluctuating selection 106
3.10 Polygenic trait under rapidly fluctuating selection 106
3.11 Mismatch of the max-entropic distribution 108
3.12 Accuracy of the dynamics at low μ 111
3.13 Dynamics of polygenic traits at low $N\mu$ 113

4.1 Evolution of genetic variances with distinct mutation rates 134
4.2 Drift and sampling on the $\mathcal{G}$-matrix 139

5.1 Distributions of allele frequencies 153
5.2 Adaptive landscapes 159
5.3 Fisher's fundamental theorem with drift 161

6.1 Genetic effects used in the $\mathcal{G}$-matrix computations 172
6.2 Indirect selection and $\mathcal{G}$-matrix evolution 174
6.3 Apparent stabilizing selection and $\mathcal{G}$-matrix evolution 176
6.4 Effects of pleiotropy over selection-mutation-drift balance 179
6.5 Effects of pleiotroy over $\mathcal{G}$'s evolutionary dynamics 180
6.6 Error in the $\mathcal{G}$-matrix's evolution 182
6.7 Averages and expectancies of $\mathcal{G}$ 183

7.1 Dynamics of the mean trait with a single lo- cus; stabilizing selection 199
7.2 Dynamics of the mean trait with a single locus; stabilizing selection with shifted optimum 200
7.3 [Dynamics of the mean trait with a single locus; stabilizing selection with shifted optimum at low $N\mu$202
7.4 Dynamics of the mean trait with a single locus; stabilizing selection with shifted mutation rates . . 203

7.5 Dynamics of the mean trait with a single locus; stabilizing selection with shifted optimum and Low $N\mu$ 204
7.6 Dynamics of the mean trait with a single locus; directional selection with dominance 205
7.7 Genetic variance of polygenic traits under stabilizing selection 210
7.8 Tri-locus trait under stabilizing selection with a shifted optimum 211
7.9 Macroscopics of polygenic traits under stabilizing selection (equal effects) 215
7.10 Polygenic trait under stabilizing selection with a shifted optimum (equal effects) 217

8.1 Life-history invariants related to growth parameters 233
8.2 Evolution of population patterns 235

B.1 Distribution of allele frequencies with wrong base distributions 293

C.1 Bootstrap distributions of the traits 301
C.2 Bootstraped average maxima 302
C.3 Number of pleiotropic structures 304
C.4 Pleiotropic structures 307
C.5 Distribution of selection coefficients (β) 308
C.6 Distribution of the eigenvalues of $\mathcal{G}$ 309

Samenvatting (NL)

Populatiedynamica beschrijft hoe het aantal individuen in een populatie verandert in de tijd. Er bestaan veel verschillende modellen het proces van populatiegroei beschrijven. Eén van de doelen van dit proefschrift is om een aantal van die modellen samen te brengen in één simpel algemeen model. Hoewel het concept "carrying capcity" of draagkracht vaak wordt gebruikt als mechanisme van dichtheidsafhankelijkheid (de regulatie van groeisnelheid afhankelijk van populatiegrootte), is het algemene model onafhankelijk van dit concept. Maar onder bepaalde condities is draagkracht toch een uitkomst van het model. Onder andere omstandigheden onstaat juist explosieve groei. Het blijkt ook dat bepaalde random fluctuaties de dynamiek zo kunnen beïnvloeden dat het lijkt alsof er dichtheidsafhankelijkheid is. Dit geeft een alternatieve verklaring voor de demografische diversiteit in natuurlijke populaties.

De evolutie van kwantitatieve eigenschappen hangt af van veel genfrequenties die zelden gemeten kunnen worden. Een benadering wordt ontwikkeld gebaseerd op methoden uit de statistische mechanica om de dynamiek te voorspellen van meetbare variabelen, zoals gemiddele en variantie van een kwantitatieve eigenschap. Er wordt aangetoond dat populaties evolueren naar een maximale entropie, afhankelijk van bepaalde statistische randvoorwaarden aan de eigenschappen in kwestie.

Deze neiuwe methode voorspelt evenwichtstoestanden exact en is zelfs accuraat onder plotselinge veranderingen in directionele of stabiliserende selectie. Toepassingen op selectie voor meedere eigenschappen worden bestudeerd, alsmede een analyse van evolutie in de kikker *Rana temporaria*. Tenslotte, ter discussie komt in hoeverre entropiein een evolutionair perspectief gezien kan worden als informatie gecreëerd door natuurlijke selectie.

Summary (EN)

Population dynamics is a temporal description of the number of individuals in a population. Many models exist that quantify the process of population growth. In this thesis, the subject is approached with the goal of bringing together several of these models into a simple general one. Although carrying capacity is often invoked as a mechanism for density dependence (the regulation of growth rates according to the population's size), the general version does not resort to this mechanism. For certain conditions, however, carrying capacities emerge. Other conditions lead to populations that explode in size. However, certain kind of random perturbations can drive the dynamics of the population in a way that they mimic density dependence. This provides alternative testable explanations for the demographic diversity in natural populations.

The evolution of quantitative traits depends on the frequencies of many alleles involved, which can rarely be measured. An approximation is developed borrowing methods from statistical mechanics to predict the dynamics of observable quantities, such as the mean and variance of a trait. Populations are shown to evolve to an entropy maximum, subject to constraints on the expected values of observable quantities. The method gives the equilibrium state exactly and is accurate even when there are abrupt changes in directional or stabilizing selection.

Applications to selection for multiple characters are also studied, and data of evolving traits of the common frog *Rana temporaria* is analyzed in order to gain insights on the limitations of the method. We also initiate a discussion on the interpretation of entropy in evolution as information created by natural selection.

Resumen (ES)

La dinámica poblacional es una descripción temporal del número de individuos en una población. Existen varios modelos que cuantifican el proceso de crecimiento poblacional. En esta tesis, el tema se aborda con la finalidad de unir varios de estos modelos en uno más general pero simple. Aunque la capacidad de carga es frecuentemente empleada como parte del mecanismo de denso-dependencia (regulación de las tasas de crecimiento según el tamaño poblacional), el modelo generalizado no acude a este mecanismo. En ciertas condiciones, sin embargo, las capacidades de carga emergen del modelo. Otras condiciones resultan en poblaciones cuyo tamaño "explota" (crece al infinito). Sin embargo, existen cierto tipos de perturbaciones aleatorias que imitan procesos de denso- dependencia. Estos proveen alternativas plausibles de la diversidad demográfica en poblaciones naturales.

La evolución de rasgos cuantitativos es dependiente de las frecuencias de muchos alelos, los cuales raramente son medibles. Una aproximación es desarrollada para predecir la dinámica de variables observables, como la media del rasgo o la varianza genética, utilizando métodos de mecánica estadística. Se demuestra que las poblaciones evolucionan hacia un máximo de entropía, restringida a los valores de las esperanzas de las variables observables. El método concuerda exactamente con

los estados de equilibrio, y es preciso incluso cuando hay cambios abruptos de presiones selectivas, direccionales o estabilizadoras. También se estudian aplicaciones para selección sobre múltiples rasgos, y se aplican al análisis de datos de rasgos en evolución de la rana común *Rana temporaria*, para comprender las limitaciones del método. También se inicia una discusión de la interpretación de laentropía en evolución como información creada por selección natural.

Acknowledgments

What an interesting experience is to make a PhD thesis these days. Not that I know how was it before, or how will it be in the future, but certainly the conditions in which I worked were great. Still, at several points in the PhD training, I wish I would have followed Freeman Dyson's advice of *not* pursuing a PhD. But for four years I did interesting things -and got paid for it! Nevertheless, these 'down' moments were also a great push to develop more independent thinking and maturity, in the personal sense, as well as in scientific thinking. This path that I followed was assisted by several, to whom I am grateful.

Ido Rolf[1], to whom I am his first (non-shared) PhD student, demonstrated scientific openness when venturing in directing and supporting my project, despite how unrelated my subjects were to his own interests. I learned a great deal from him, from his criticism and his 'down to the point' thinking. With much appreciation I received his support and advice when venturing in new aspects of my research, and in draining the poisonous and stiff views that I had from my years as a [wannabe] physicist (which except for the methods, physicists' approach is much useless in evolutionary biology!).

[1] Prof. Dr. Ido Pen.

Although his involvement in my project was much more passive, I would like to thank 'Herr Professor' Franz Joseph[2]. His continued support from the academic side was a key factor for completing this PhD. Although in many aspects we have opposing views, I admire his maturity in respecting this, and his insistence in using divergent opinions as a method of thinking and analyzing ideas, which has been enriching.

With as much appreciation as to my advisor and with ample admiration, I thank Nicholas Harry[3], from whom I learned so much. I am also very glad for his kindness, hosting at the University of Edinburgh, and the support that I have received from him, even when he had no academic responsibility over me. He would deserve to be a co-advisor of my project, if I had the choice.

Also thanks to the people around me in the Theoretical Biology Group (The *TheoBio*'s) during my project, who with our discussions sharpened the blade of scientific thought. Martin Hinsch, with whom I shared office for 4 years, Tomás Revilla, Thomas Berngruber, Max Wolf, Tim Fawcet, Thor Veen, Bram Kuipers, Vouter Vahl, Charlotte Hemmelrijk, as well as the other *TheoBio*'s to whom I had the pleasure to torture with my boring seminars and literature discussions. Special thanks to Gudrun Ferber, and Joke Bakker for their help. Also thanks to those non-*TheoBio*'s around in the university, Liliana Ballesteros (who dared to be my student in a master project!), Verena Brauer, Francisco Encinas, Rampal Etienne, Anna Harts, Silvia Martínez, Saleta Perez-Vila, Joanna Reszka, Julia Schroeder, Kuke Bijlsma (now a *TheoBio*!), and the rest of the friends and colleagues in CEES.

[2]Prof. Dr. Franjo Weissing.
[3]Prof. Dr. Nick Barton

Likewise, thanks to the old friends around the world who were always willing to discuss, in particular Roberto Cipriani, Omar Cornejo, and Claudia Moccia. And with big appreciation thanks to other friends and colleagues with whom I had great discussions, and/or brought attention of interesting problems: Reinhard Bürger, Julián Chela-Flores, Ananías A. Escalante, Maciej Jan Ejsmond, Brian Enquist, Simon A. Levin, Konrad Lohse, Hans Metz, María Elena Márquez, Eladio Márquez, Rosa Moscarella, Jafet Nassar, Eörs Szathmáry, Leonardo Trujillo, Bruce Walsh Stuart West, Jochen B.W. Wolf, and Ziheng Yang. In particular, thanks to Prof. Dr. Jan Kozlowski and Krzysztof Argasinski, who hosted me for a month and a half in their group in Jagiellonian University in the amazing city of Krakow.

I also would like to acknowledge Dr. J. Merilä for sharing with me the data in *Rana temporaria*, and Drs. C. Boesl, B. Enquist, A. Heinsbroek, J. Hobbs, S.Leigh, A. Ojanguren, and T. Sakata, for providing Liliana with the data for her research, results which are valuable of publication, hopefully in a near future.

My personal thanks for those friend who stood close to me, physically and/or mentally, my mother, uncles and grandparents, the 'cousinhood', and the rest of my family. Thx to the Dharma-friends all around the world, to Ula, and to my Lamas Ole and Hannah Nydahl for their words of support and activity which quite helped to keep my strength in critical periods.

Last I gratefully acknowledge the funding from the university, the Ubbo Emius Burse for the fellowship, that allowed 100% freedom in my projects; the ConGen grant of the European Science foundations, which funded my 3-month (that ended in a 4-month) visit to Edinburgh; the University of Edinburgh, for supporting my stay there (and Nick's projects for supporting other short visit); as well as the other grants for meetings and the organizers of: the European Society of Evolutionary Biology (*ESEB*), the European Meeting of PhD Students in Evolutionary Biology (*EMPSEB*), the staff of the Mathematical Models in Ecology and Evolution (*MMEE*), and the Dutch Meeting of Theoretical Biology (*NVTB*).

HPdV
London
November 17, 2008

General Introduction

Life, as such does not exist. This was Szent-Gyorgy's (1972) perspective about what is life; *The question, in itself its wrong*, he wrote. The view nowadays, is that we cannot actually define life, since there is no physical or intrinsic component that makes it happen, and rather is conceived that life itself is a process (Mayr, 1982). Lwoff (1965) stated *life is a state of the organisms*, which at first read does not say too much, since organisms are those capable of being alive. However it sets the question in such a way that we can have a working definition of life, that is, listing the properties that we associate with 'living entities' (Maynard-Smith and Szathmáry, 2000, p. 3). These properties are (Maynard-Smith and Szathmáry, 2000): (i) multiplication, (ii) variation, and (iii) heredity (MVH). This definition, as more specific ones, suffers from Sagan's "fundamental handicap of biologists" (Sagan, 1973; Emmeche, 1998), that is that we define life only on basis of the organisms that we know, and that are subject to the same "laws" of evolution, ecology, physiology, genetics, chemistry, etc. Ernst Mayr's (1982) list of properties that define life, picture it as an evolutionary and dynamic process, rather than as an intrinsic property of organisms. Notice that the notion of variability and growth are fundamental elements, as well as that the several definitions of life apply to populations, not to individuals (although of course we can extend the definition that an individual is alive if it descends or belongs to a population that has those properties; Maynard-Smith and Szathmáry, 2000).

These definitions, in particular the above mentioned ones, are a reflection of our conception on how the process of life is. Multiplication (population growth) and variability in populations -the two main subject of this thesis- are fundamental in evolutionary biology, and is no coincidence that they constitute -at least in part- the definition of life. Thus somehow (and this is an existentialist argument) the questions that we ask in evo-

lutionary biology, contribute to build an approximation to this unanswerable question. The abstractions of the vital processes that are used to model biological evolution (in this thesis, mathematically), help to slightly widen our handicapped notions.

For example, what we conceive as a replicator (or for these matters, a reproducer), need not to be as physical as we regard it in an organism, or its hereditary material need not to be genes, nor their composition be of DNA (e.g. memes). The evolutionary models and the modern evolutionary synthesis assume certain properties and explain in terms of processes that do not depend on these biological details. Thus the handicap becomes less pronounced (although it is still there).

Part of this flexibility on our re-interpretation of the evolutionary processes out of our own formulation, is because we are modernly regarding the evolving systems as transducers of information (from a generation to another). Hence the limits of our abstract interpretations can be widened without tormenting ourselves with mechanistic details of a high level of complexity, which is to some extent and from this evolutionary perspective, unnecessary. Essentially, in studying evolutionary processes, we are abstracting what life is by addressing the question *how can this evolve?* in both instances: of a particular biological aspect (e.g. a particular mechanism, a particular trait, etc.), or of a system as a whole (without neglecting the environment, of course). In answering this question, (how can this evolve?) we are forced to invoke MVH (plus other things).

Motivation for this research

Putting aside the philosophical aspects, there are of course more specific question or subjects that we (me and co-authors) will deal with. There are also practical needs to understand how populations adapt to new environments. We will disect

MVH into separate processes. This is still another simplification, perhaps not the most natural, but the most historically parsimonic.

On the one hand, the techniques for studying population dynamics have improved, and in consequence have improved their applications. We do not consider all populations as logistic and exponential any more. But along developments in non-linear dynamics, and stochastic processes, as well as advances in classical and Bayesian statistical tools for estimation and forecasting, we have had a more comprehensive view of population growth.

On the other hand, a great deal of the approaches to problems in genetics, specially in the *age of molecular biology* came from the Neutral Theory of Molecular Evolution (Kimura, 1985). Much of the ongoing process of constructing the evolutionary synthesis was halted by the arrival of the neutral theory. Along with the central dogma of molecular biology, neutrality ruled the view of world. In the last two decades, with access to so much genetic data, along with the development and application of bioinformatic tools and the study of epigenetic factors, we have found that the molecular world is not neutral, and that there is no linear relation between the genetics and the phenotypes.

So its time to get back to the basics. We need to understand the dynamics of populations in a better -perhaps more fundamental- approach. Not necessarily to construct more complicated models (although why not!), but to understand what the simple models have given and why, and how this relates to the genetic structures in a population. These are not separated problems. We just view them like that. The study of the roles of selection in natural and artificial populations has also been of great importance in the last years, since have given further insight in the processes of evolution.

Nowadays biologists are more prepared for integrative and fundamental questions, since the use of mathematical and computational tools have (fortunately) permeated to even the most practical aspects of biology, and even triggered the creation of new mathematical tools.

This thesis aims to complement the theories of population dynamics and quantitative genetics. Seeking a synthetic approach to each of these two subjects, where there is plenty of research to discover new facts but not so much to bring them together.

The first subject: Population dynamics (PD) is among the oldest subjects in theoretical and mathematical biology, dating back to Malthus (1798); Gompertz (1825), and Verhulst (1838), and originated in the study of demography. We don't concisely know how to 'derive ' population dynamics from first principles. But nowadays it conforms to one of the central elements of the theory of evolution, both as a subject to study and as a tool to study other subjects.

As Gilpin and Ayala (1973) put it: *"Biology is at a Keplerian stage"*; this was more than three decades ago, and much advance came until today. We could say that nowadays it is rather at a Newtonian stage[4]: there are many ways populations grow, or "growth laws"; some of them have mechanistic explanations, some of them only phenomenological justifications. But the applications of these dynamics is widespread in biology. However there is not a consistent theoretical background that leads to the understanding why populations grow in particular ways.

This is of course not coming as a surprise. If we give it a thought, there are so many factors that determine growth that it is hard to think where to start. For the sole purpose of exer-

[4]Although I. Pen suggests is rather a Laplacian stage.

cising I made a mind map of the factors that I would consider to directly affect growth. The result, was a long list of factors which besides growth, affect each other in a fully connected network fashion. Not too good for a start. But it reflects that distinct disciplines have considered the action and effect of biotic and abiotic factors on growth, so there are several pieces of the puzzle. But how growth has been tuned by the undergoing evolutionary process, is not a question with a trivial answer. There are two classes of hindrances in the search for this answer. From a perspective, and to the main concern in this thesis, the way in which we model population growth is to some extent arbitrary. This has resulted in having a battery of models (Henle et al., 2004a) that are 'adapted' to their use in particular problems. This is interesting, in that it might reveal that the diversity of growth strategies is big. Often these models are chosen by distinct criteria, and do not necessarily reflect biological factors that might be of relevance. Furthermore, it is common to find distinct models that result in very similar –or identical– growth patterns, but which have radically different biological implications. But there is no criterion that suffices for a choice of biological significance. Even if we were to apply statistical methods for model selection, a set of hypotheses is likely to be biased by mathematical easiness. Between lines, I am assuming a reductionist position. The problem might well be how to better 'explain' a growth pattern with minimal set of parameters. But I am referring first, how to understand the factors determining growth, and second, how evolution can shape these factors to result in an evolutionarily stable growth strategy.

There is plenty of work on the evolution of population growth, and how populations adapt to particular conditions.I identified five main trends in the study of the evolution of growth strategies. I am considering models that do not take into account

competition (e.g. Lotka-Volterra types are excluded), and these may or may not comprise age structure. I will briefly go over them, and bear in mind that this classification is arbitrary and the classes are not exclusive. First, following the unexpected results that the simple discrete logistic equation shows complex (chaotic) behaviours (May, 1976), there was a rush to study this new phenomenon, and its consequences. To the big regret of many, studies revealed that in most cases, evolution would tend to tune the dynamics in such a way that they result in a stable equilibrium (Doebeli and Koella, 1995; Ebenman et al., 1996; Schliekelman and Ellner, 2001), although some special conditions would allow the chaotic dynamics (Ferriere and Gatto, 1993; Gatto, 1993; Doebeli and Koella, 1995; Gonik et al., 2005). Second, the influence of stochastic factors on population growth has been a stereotypic model with vast applications (Tuljapurkar, 1990; Lande et al., 2003). In this context, it has also been studied how distinct strategies would evolve to cope with these fluctuating realms (Tuljapurkar and Orzack, 1980; Tuljapurkar, 1982; Orzack and Tuljapurkar, 1989; Yoshimura and Jansen, 1996). Third, the most biologically comprehensive approach is of optimizing growth rates that are determined by life-history traits. To begin with, the proper fitness measure has been debated (Murray, 1997; Metz et al., 1992; Rueffler et al., 2006), and it seems that optimizing the Malthussian parameter (exponential growth rate) is the most consistent option. To follow, the dynamics are coupled to the evolutionary benefits of individuals that maximaze their fitness (Metz et al., 1992; Mylius and Diekmann, 1995; Coulson et al., 2006; Pelletier et al., 2007). This approach is very versatile, and allows to modeling of specific situations for which biological details are included with easiness (Orzack and Tuljapurkar, 1989; Charnov, 1993; Shertzer and Ellner, 2002; White et al., 2006). Fourth, ecological variables and the spatial structure also de-

termine growth rates of populations (Lion and van Baalen, 2008). Fifth, there are genetic determinants to growth (Travis and Greenwood, 1990; Hastings and Harrison, 1994; Doebeli, 1996b; Doebeli and de Jong, 1999) which may be assumed to act directly on growth rates, or indirectly on any other life-history traits.

Most, if not all of these approaches employ specific growth models, parameters of which are tuned by evolution. A classical example, is the logistic model, and the notions of r and K selection. This analysis presumes that growth is logistic, and the evolutionary reasoning sets the details on the values of these parameters (MacArthur and Wilson, 1967; Pianka, 1970). But any alternative strategy that would result in growth dynamics different from logistic, is of course disregarded. A synthetic approach, is missing (though see Metz et al., 1992; Meszéna et al., 1992; Page and Nowak, 2002). Thus we need a way to understand the growth of populations from a wider view. That is to understand what do the different growth patterns have in common. There are two common assumptions, which are (i) growth at low densities is approximately exponential, and (ii) growth at high densities is limited. From the mathematical side, this is equivalent as considering the first two terms in a series approximation of *the* growth law. So naïvely we could adopt the third term, and so on. Surely we can do better than that. Indeed a notable advance was initially achieved by Ayala et al. (1973); Gilpin and Ayala (1973); and Gilpin et al. (1976). They introduced a model that accounted for the internal competition of a population. This has been later applied to study global patterns of growth (Sibly et al., 2005). Although not yet in the evolutionary context, this application of a generalized model reveals another dimension of evolutionary possibilities. How much can we extend this, without invoking more parameters or artificial models? This is the research subject of the first part of this thesis.

The second subject: the theory of quantitative genetics (QG) aims to explain the evolution of quantitative traits and characters based solely on measurable quantities. That is, without making reference to variables that we can not measure, like allele frequencies, genetic effects, or number of loci. The theory of population genetics (PG), on the other hand, studies the evolution of the frequency of alleles, and of the genetic effects of allelic substitutions over phenotypic traits. Still, the relation between both approaches is obscure. Whether it is possible or not to bridge both, we still don 't know, but quantitative genetics relies on the ansatz or conjecture that it is.

Mechanistic approaches from population genetics have, in a sense, failed to achieve this bridge. Only some approximations have been fruitful for specific situations. Still it looks from the experimental and empirical perspective that it is indeed possible. Thus the subject persists.

Why do we need at all the integration of both sub-disciplines of genetics? In *The Origin of Species* Darwin (1859) recognized that the mechanisms of natural and artificial selection were essentially the same. Artificial selection is applied on phenotypic traits, and in a sense ignores what is behind it, in the genetic composition (whose nature was at the time unknown to Darwin). This points out the primary role of understanding the nature of selection. Even when mathematical tools started to be applied to compute the response to selection, the approach was entirely phenotypic (Pearson, 1896), and to certain degree, it has remained like that. It was not until Fisher (1918) when the relation between the Mendelian nature of phenotypic traits under selection was addressed.

But Kimura's (1985) *Neutral Theory of Molecular evolution* went totally in the opposite direction. It required molecular data, and assigned the major evolutionary cause to point mutations and random drift, rather than to natural selection. As

the saying goes, the truth is the intersection of two independent lies, hence given that these three factors, selection, mutation, and drift (SMD) are potential causes for the evolution of virtually any trait, we are thus interested in their relative importance to generate and maintain the diversity that nowadays exists in populations.

It is actually not possible that phenotypic traits change without consequent evolution of their genetic composition. But as breeders have shown, it is possible to predict –to a certain degree– the average values of the offspring's traits of breeding individuals selected for a trait. Thus it is clear that there are immediate applications. Yet, this will only work for some generations, thus for purposes of understanding and explaining biological diversity, these predictive capabilities are not enough. This is because the predictability of the trait values in a population depends on the amount of variation that is available at breeding time. And this variation is changing. How is this change? *That is the question*.

Predicting this change of the genetic variance is not a trivial matter, since it depends on several biological factors, many of which are not measurable quantitatively. PG plays a role here, indicating the factors that influence the evolution of the genetic variance. The down side, is that it necessarily depends on these non-obervable elements (Barton and Turelli, 1987).

A somewhat more realistic situation is the various traits coevolve. Lande (1979) studied this scenario, showing how the genetic co-variances ($\mathcal{G}$ -matrix) are influenced by pleiotropy (Lande, 1980). In general, linkage, epistasis, (Turelli and Barton, 1994), environment, sexual selection (Barton and Turelli, 1991; Turelli and Barton, 2004), and other genetic and ecological causes (Arnold et al., 2008) will affect the co-variant structure of any trait.

It is possible to study this situation in a bottom-up fashion,

that is considering the evolution of allelic effects, and how it leads to a change in the traits. But it is not easy to identify these elements form quantitative data alone, in order to give a fulfilling explanation of evolutionary response from a QG approach alone. Thus we are in need of a way to relate the non-observable factors to the observable variables. Then it might be possible to predict the long term evolution of quantitative variables with accuracy.

Although this remains an important subject, which we will address in the second part of the thesis, we might still wonder why the question of the integration of PG and QG is important, given the advances in molecular biology techniques, which allow us to screen the genetics effects over any kind of trait.

First, of course is the fact that these empirical analysis have certain limitations. Recent theoretical studies (Sella and Hirsh, 2005) have shown that even if an equilibrium between SMD is maintained, the rates of molecular substitution are equal. This is a result which was reserved to, and interpreted as, neutrality. The implications are not yet studied. But clearly we might be missing something. The easiness with which sequence data is analyzed under a neutral model assumption (Li, 1997) might prove misleading, compared to the view where selection is considered.

Second, the identification of quantitative trait loci (QTL) is of major relevance to quantitative genetics. It gives an idea of the amount of loci that might be contributing to the quantitative trait and its variation, as well as their effect. However, QTLs have resolution to discern only those loci of major effect. Furthermore, the technique is able to identify only two alleles. Thus we still are uncertain of the number of loci which are actually contributing, and the overall effect of the contribution of those loci with smaller effect over the quantitative variables. This is critical, since many models assume a contin-

uum of alleles and/or infinite number of loci (Kimura, 1965a; Bulmer, 1972; Lande, 1975; Kingman, 1978; Bulmer, 1980, all reviewed by Turelli, 1984). In these cases, the distribution of the trait will be essentially Gaussian (Turelli and Barton, 1994). Although this situation in practice is that with few di-alleilc loci (say five to eight, as will be shown in chapter 5) normality is already a good approximation. In addition, the estimation of the QTL effects are a statistical matter. Xu (2003) has shown that the sampled population size might substantially bias the estimation of the effects (the *Beavis effect* Beavis, 1998), and its likely that many QTLs that have been reported suffer from this oversight. We know that the distribution of genetic effects is highly skewed, as stated above, with many QTLs of small effect and few with large effect. The problem of underestimating the variation contributed by the many loci with small effect, is that these will actually compensate the variation that is rapidly lost by selection on those loci of high effect (Barton and Turelli, 1989; Barton and Keightley, 2002), so forecasting of evolutionary response is dependent on these underestimated genetic elements.

Third, the architecture of the trait, that is how the allelic effects affect simultaneously and non-additivelly different traits, plays a crucial role in the response to selection (Orr, 2000; Cheverud, 2006; Wagner et al., 2008). For simplicity, most works assume that the contribution of the genetic effects over a trait can be decomposed into the additive and non-additive factors. The former just adds the effects of all the genes contributing to the trait, whereas the latter comprises the interacting factors among all these (or other) genes. With this division in mind, it is possible to make appropriate design to identify QTLs not only for the traits, but also for their interaction (Lynch and Walsh, 1998; Cheverud, 2000). Yet the nature of these interactions is uncertain (Hansen, 2006). Theoreticians typically

assume the influence over the traits comes out of pairwise epistasis (Kondrashov and Turelli, 1992; Turelli and Barton, 1994; Carter et al., 2005). In any case, the issue is that the complex essence of evolving traits, or alternatively, the complex background on which additive or non-additive traits evolve, will by all instances affect the change of genetic variation (Gavrilets and de Jong, 1993; Goodnight, 1995; de Brito et al., 2005), and the knowledge on how this variation will change, is by no means obvious even when knowing pleiotropic and epistatic QTLs.

Fourth, from the genetic point of view it is ambiguous to gauge the evolutionary causes of quantitative characters. Not only for the reasons above, but also because the action of selection at the level of phenotypes might be of a distinct nature than selection at a given locus (for example, selection for a specific amino acid in a protein, Yang and Swanson, 2002; Yang et al., 2005; promoter regions, Haygood et al., 2007; Kawabe et al., 2007; or any other specific molecular unit Rand, 2001; Hoede et al., 2006; Haddrill et al., 2008; Kim and Wiehe, 2008). If selection is acting over these genetic elements, or if selfish genes are inducing genetic conflict (Werren et al., 1988; Hatcher, 2000), at the same time as they are affecting traits under selection, then the net effects of selection over each of these two units will be hard to discern on a particular locus. This kind of effects, consonant with the theory of multilevel selection (Okasha, 2006), have not been studied for the evolution of polygenic traits. Nevertheless, it will not be long before this happens.

To conclude, personally I think that this issue should be interpreted in the opposite direction. That is, how can this genetic information help us to refine the synthesis of PG and QG. Combining these two fields in one thinking seems to give much more than what each field give on its own. The availability of genetic and molecular data helps to refine the quantitative

studies, since it allows a clear view of the assumptions that we can actually make in achieving a successful quantitative theory that can support the empirical facts.

It is recognizable that these are two classical subjects of the 'modern evolutionary synthesis' which are still debated and require completion, and it is desirable that their foundations are solid, so they can support the study of complex interactions in the micro and macro scales. subjects to which in the meantime we are moving further.

Guide to this thesis

The thesis is divided into two parts and a synthesis. Each part contains some chapters with the original results. These are published or (almost) submitted for publication in peer-reviewed journals. They are followed by perspectives chapters that includes research that has not yet been published for distinct reasons, and speculations about the future prospects of the results. I will end with a synthesis, which builds the 'big-picture ' of my results for both the specific subjects, as well as the integrative view of them.

Part I. Population dynamics As discussed above, in PD we find a variety of models that typically describe various phenomenologies, yet a full integration of these models is absent. We pursue such an integrative theory, for a class of growth patterns which are common in the literature of PD, and that describe non-interacting populations. This consolidation is achieved by following the dynamics not only of population size, but also of the per-capita rate. Surprisingly, at least for the class of density dependent patters that are studied in Chapter 1, the per-capita

growth rate is independent of the population size and carrying capacity, yet describes patterns of density dependence. The carrying capacity is determined by the initial growth rate of the population, thus it is rather a consequence of a population's trait, rather than a purely environmental property. (Needless to say, the initial rate is by no means independent from the environment!).

In Chapter 2, environmental stochastic effects on the growth rate are considered. A first prediction is that even those growth patterns that would lead to a population explosion are controlled by the stochastic effects. Infinite-sized population are thus avoided without the need to invoke, but not excluding the presence of, carrying capacity. But there is a second prediction, much more quantitative. The patterns of growth will result in a logistic form, irrespective of the deterministic properties. The equilibrium will be attained at random population densities. Even if the deterministic population is logistic, its carrying capacity will not be a predictor of the equilibrium density.

Part II. Population Genetics PG describes the mechanisms of QG, but it is not entirely clear whether the phenomenological framework of QG can be derived from PG. The subject is full with fuzzy results, with few (or limited) final statements. In Chapter 3 this question is undertaken. This is achieved by considering not the quantitative variables themselves but expectations of these. The distribution used to compute the expectations is obtained through the maximization of entropy (ME), restricted to the quantitative observable variables. This distribution coincides with the exact solution to the diffusion approximation in equilibrium, and approximates very well the distribution when evolution is taking place.(This method is equivalent, or analogous to the coupling of statistical mechanics between microscopic and macroscopic variables). The main applications

in this chapter is to polygenic traits of arbitrary number of diallelic loci with distinct effects, but the applications to other situations, like stabilizing selection and other schemes inducing epistatic effects is also addressed.

But polygenic traits tend to be correlated among each other. In Chapter 4 the results of the previous one are extended to include pleiotropic effects. Under these circumstances, the evolutionary dynamics involve the genetic co-variances matrix, $\mathcal{G}$, which contains in every entry the genetic covariance between any pair of co-evolving traits. We can thus, as above, calculate the expectancies of the genetic co-variances. In this chapter the analyses are not restricted to theoretical constructs, but are employed to analyse previously published data on *Rana temporaria*, for which differences of $\mathcal{G}$ have been quantified across two different locations on four covariant traits. We contrast the results to the scenarios where we assume that directional or stabilizing selection is maintaining the observed genetic variability.

These specific results from the ME method, are accompanied by an interesting conceptual system. These ideas are explored in Chapter 5. The evolution of expected values are of a different nature and we must think of them in different terms than of the quantitative (which are stochastic) variables themselves. These are of course not unrelated. We discuss and apply the concepts above in relation to fitness landscapes, and how to retrieve information of n genetic variables (allele frequencies) using only m quantitative variables, where $n \gg m$. Another result concerns an extension of Fisher's Fundamental Theorem of Natural Selection in the expected values of mean fitness, which will always increase, under the effects of selection and drift. This is an extension that appears only at a statistical level. We also discuss the possibility of employing the ME framework to test distinct hypotheses of the mutation-selection-drift circumstances that

determine the structure of the G matrix. In particular we apply it to contrast the scenarios where directional and stabilizing selection are determining observed empirical $\mathcal{G}$'s

In Chapter 6, I come back to correlated evolution, but from the genetic side. The ME estimations show that the $\mathcal{G}$ matrix is very much constrained by pleiotropic effects. Although selection will induce change in the mean trait, the changes in the $\mathcal{G}$ matrix are delayed to latter stages, perhaps hundreds of generations (for typically low mutation rates), remaining practically constant until then. The effects of apparent stabilizing selection and of the amount of pleiotropic loci over $\mathcal{G}$'s evolution are also studied.

As a last pivotal example, the extension of the theory to stabilizing selection (SS) is presented in detail (chapter 7). The situation is more complicated in that SS is inherently non-linear over the trait under selection. Thus introducing mathematical complications that are limiting. We investigate ways on how to overcome these technical difficulties. The case of directional selection is revisited, with allelic effects with dominance, which can be viewed as stabilizing (or disrupting) selection over each locus.

In the appendices it is shown that maximizing entropy with constant fitness is equivalent to maximize fitness at constant entropy. A discussion about the analogy between statistical mechanics in physics and the ME methods herein presented, is addressed. And finally a battery of the most important formulas for the quantitative variables, in all the above situations is provided.

Part I

Population Dynamics

Published as: H.P. de Vladar – *Density Dependence as a Size-Independent Regulatory Mechanism* Journal of Theoretical Biology. vol. 238, no. 2, pp. 245–256, 2006.

Chapter 1

Density Dependence as a Size-Independent Regulatory Mechanism

> I know it is the fashion to talk about groups, the mass, the race, as though the individual had no importance at all, but in any creative action it is the individual who matters.
>
> Jiddu Krishnamurti

Abstract

The growth function of populations is central in bio-mathematics. The main dogma is the existence of density dependence mechanisms, which can be modelled with distinct functional forms that depend on the size of the population. One important class of regulatory functions is the θ-logistic, which generalises the logistic equation. Using this model as a motivation, this paper introduces a simple dynamical reformulation that generalises many growth functions. The reformulation consists of two equations, one for population size, and one for the growth rate. Furthermore, the model shows that although population is density-dependent, the dynamics of the growth rate does not depend either on population size, nor on the carrying capacity. Actually, the growth equation is uncoupled from the population size equation, and the model has only two parameters, a Malthusian parameter ρ and a competition coefficient θ. Distinct sign combinations of these parameters reproduce not only the family of θ-logistics, but also the van Bertalanffy, Gompertz and Potential Growth equations, among other possibilities. It is also shown that, except for two critical points, there is a general size-scaling relation that includes those appearing in the most important allometric theories, including the recently proposed Metabolic Theory of Ecology. With this model, several issues of general interest are discussed such as the growth of animal population, extinctions, cell growth and allometry, and the effect of environment over a population.

1.1 INTRODUCTION

The logistic equation is a paradigm for population biology. This simple model, in its continuous (Verhulst, 1838; Pearl, 1927) or discrete (May, 1976) versions describes two fundamental issues of population biology, which are (i) the initial exponential rates of growth, and (ii) density-dependent effects, like competition under limited resources, indicated by saturation values. The discrete logistic equation, in itself opened a new and broad field in biology related to chaotic behaviours, and for which some empirical evidences exist (Hanski et al., 1993; González et al., 2003). The continuous version of logistic growth, although sharing properties with its discrete analog, differs in some aspects. It does not show intrinsic bifurcations as the discrete version does, and is much more simple to treat analytically.

Gilpin and Ayala (1973) and Gilpin et al. (1976) introduced a model that "slightly" generalises the popular logistic equation. Their model, consists on modifying the term corresponding to the density-dependence with an exponent θ. Compared to the logistic equation, their "global model" describes a population that converges in time to the same size as the logistic growth, i.e. to the carrying capacity. However, the exponent θ gives new interpretations to this sigmoid model of growth. If $\theta > 1$ then intra-specific competition is high, and the population takes more time to reach its asymptotic value, termed carrying capacity. If $0 < \theta < 1$ then competition is lower and the carrying capacity is reached earlier than in the corresponding logistic dynamics (Gilpin and Ayala, 1973; Gilpin et al., 1976).

The θ-logistic model, as it has been termed afterwards, introduced a new concept on population ecology that is the θ-selection strategies(Gilpin and Ayala, 1973; Gilpin et al., 1976). Originally, they proposed the model to explain data from com-

peting *Drosophila* systems after failure to use a Lotka-Volterra-like model (Ayala et al., 1973). Afterwards, non-competitive versions of the system (i.e. one "allele" or one "species" model) has been used in conservation ecology to model avian population dynamics and calculate extinction times (Saether et al., 2000), and also to estimate the effects of environmental stochasticity on population growth (Saether and Engen, 2002). Other population models have included stochasticity to aid parameter estimation and study the effect of environmental changes in caprine populations (Saether et al., 2002). This model, has also been employed in community ecology to estimate species abundance (Diserud and Engen, 2000). The θ-logistic equation is a "slightly more complicated model [that] yields significantly more accurate results", using the original words of Gilpin and Ayala (1973).

There are, however, other kinds of regulation terms that have been successfully employed to model other kinds of populations and growth. Sigmoid curves in particular are attractive for biologists, but are not necessarily described by θ-logistic equations. The von Bertalanffy (1966) equation, for example, is a sigmoid curve that is frequently used in allometric and ontogenetic modelling, as well as the recently proposed (and controversial) curve derived from bioenergetic considerations by West et al. (2001). Another kind of sigmoid is given by the Gompertz equation (Gompertz, 1825), which was originally formulated to model human demographic data. The Gompertz equation has become an important tool in modelling tumour growth (Norton et al., 1976), although applications include a wider range of topics.

Among non-saturated growth for population, there is the classical exponential growth, typically employed to describe bacterial clonation (Hershey, 1939), or simply as descriptors for non-regulated conditions of growth. However, a "general ver-

sion" of the exponential is potential growth, appearing in tumour biology (Hart et al., 1998), life history theory (Calder, 1984; Roff, 1986; Day and Taylor, 1997; Stearns, 2004), as well as in allometry (Peters, 1983; Calder, 1984; Brown and West, 2000), and has been used also to model the growth of populations of prebiotic replicators (Szathmary, 1991; Scheuring and Szathmary, 2001) and simple approaches to sexual reproduction (Szathmary, 1991).

The form of the θ-logistic model is actually more general than it seems if it is interpreted from a "wide perspective". This paper introduces an alternative way to interpret and formulate population dynamics models. Although it is strongly motivated by the θ-logistic equation, the description explained through out this paper reduces *exactly* to most common population models, including the above-mentioned growth dynamics. This formulation provides a very simple way to manipulate dynamical equations, and it depends only on two parameters.

Another important feature is that with this formulation it is possible to derive general scaling behaviours of populations to their initial sizes and carrying capacities, in a similar but more general way than that of West et al. (2001).

1.2 "MECHANICS" OF SIZE REGULATION

One of the central issues in population dynamics is to determine the growth function that describes a particular population. Growth dynamics is in general of (or can be expressed in) the form

$$\frac{dx}{dt} = xr(x) \ . \tag{1.1}$$

The growth rate $r(x)$ is an explicit function of x. Depending on the nature of the self-regulation, $r(x)$ has different functional forms, the most common continuous time functions are listed

in table 1.1. For a wide review of density dependence functions, including (mainly) discrete dynamics, the reader can refer to Henle et al. (2004b).

Natural choices for $r(x)$ are functions that include two terms, one describing replication, usually of first order in x, and another one describing interaction and/or growth inhibition, of higher orders in x.

The logistic model, for example, describes growth inhibition with a second order term, i.e. x^2. The θ-logistic generalises this second order term to one of an arbitrary order greater than one, expressed by $x^{\theta+1}$. In a biological sense, this non-linear term is proportional to the frequency of contacts that an individual must have in order to produce population growth inhibition –anergy– or promotion –synergy– (Szathmary, 1991; Rueffler et al., 2006).

The idea in this paper is simply to express the dynamics of a population not anymore using its size, x, as the variable of interest, but rather studying a decomposition of it considering the rate r at which the population grows as a separate variable, that is focusing in how each individual is contributing –in average– to the population number (Rueffler et al., 2006)

1.2.1 Exponential Growth

In general the methodology consists of studying a two dimensional dynamical system $(x, r) \in \mathbb{R}^+ \times \mathbb{R}$ describing the population size and replication velocity, respectively. In the case of the exponential growth, because $r = \text{const} = \alpha$, its time derivative is zero. Thus the following trivial system:

$$\dot{x} = xr \ , \tag{1.2a}$$

$$\dot{r} = 0 \ , \tag{1.2b}$$

Table 1.1: Common regulation functions for different population growth models. In all these equations ρ is the Malthusian parameter, θ is the competition coefficient, and α is a parameter determined from environmental conditions. When populations grow to a saturation, α is related to the carrying capacity.

Model	Growth Rate, $r_{(x)}$	Malthusian Parameter, ρ	Interaction Parameter, θ	Initial Rate, $r_{(0)}$
Exponential	α	0	0	$\neq 0$
Logistic	$\rho(1-\alpha x)$	>0	1	>0
θ-Logistic	$\frac{\rho}{\theta}(1-\alpha x^{\theta})$	>0	>0	>0
Gompertzian	$-\rho\log(\alpha x)$	>0	0	>0
Hyperbolic	αx	0	1	>0
Parabolic	$\alpha x^{-1/2}$	0	$-\frac{1}{2}$	>0
von Bertalanffy	$-3\rho(1-\alpha x^{-1/3})$	>0	$-\frac{1}{3}$	>0
West et al. (2001)	$-4\rho(1-\alpha x^{-1/4})$	>0	$-\frac{1}{4}$	>0

is equivalent to exponential growth.

Note that the original expression in terms of one variable, i.e. $\dot{x} = x\alpha$, can be obtaining simply integrating Eq. (1.2b) and substituting it into Eq.(1.2a) . Integration of the second equation simply gives the constant α, which is determined by the initial conditions of the system $(x(0), r(0))$.

1.2.2 Logistic Growth

For the logistic growth, it is necessary to *define* the new variable r as

$$r(x) = \rho(1 - \alpha x) \ , \tag{1.3}$$

where ρ is the Malthusian parameter, and $\alpha > 0$ is the inverse of the carrying capacity. Thus the rate equation again is expressed implicitly as $\dot{x} = xr$, and the time derivative for r is:

$$\dot{r} = -\rho\alpha\dot{x} = -\rho\alpha x r \ ,$$

Regrouping, and then summing and subtracting 1 in the parenthesis, it is possible to write:

$$\dot{r} = \rho(1 - \alpha x - 1)r = (\rho(1 - \alpha x) - \rho)\, r \ .$$

The inner parenthesis of the last expression has the *explicit* form of r. After replacing it with Eq. (1.3), the rate equation becomes:

$$\dot{r} = (r - \rho)r \ . \tag{1.4}$$

Therefore, to solve the dynamical system equivalent to the logistic equation only one parameter and an initial condition are needed. Actually, the initial condition automatically defines the carrying capacity of the population.

1.2.3 θ-Logistic Growth.

The rate for the θ-logistic is defined as

$$r(x) = \frac{\rho}{\theta}(1 - \alpha x^{\theta}) \ . \tag{1.5}$$

Following the same methodology as with the logistic growth, it is not difficult to demonstrate that the the implicit form for the rate equation is:

$$\dot{r} = (\theta r - \rho)r \ . \tag{1.6}$$

Although derived from the θ-logistic, this last equation is general. In the limit $\theta \to 1$ the logistic equation is recovered, and taking *jointly* the limits $\theta, \rho \to 0$ Eq. (1.6) reduces to the simple form of exponential growth. Note that from Eq. (1.4) it is not possible to formally take the limit to the exponential, since it does not show an explicit dependence on θ $(= 1)$. To recover exponential growth from the explicit form of the logistic, the limit would have to be taken as $\alpha \to 0$. But then the rate of the exponential growth will be ρ instead of α. In this formulation α and ρ have distinct properties. On the one hand, ρ is defined as a parameter of the system, and as such may have a role in bifurcations and global stability, while on the other hand α is defined as an initial condition, so it does not play any role in local or global stability. Also a particular population grows following a predefined replication constant ρ, which is considered to be determined by intrinsic factors, while α is determined extrinsically by environmental conditions which define the carrying capacity of the system (MacArthur, 1962). Thus in this mechanistic interpretation where r determines growth response, the environmental issues play no dynamical role *unless they are explicitly and dynamically affecting growth rate*.

1.3 GENERALIZED RATES OF GROWTH

The three versions of the model studied above, namely exponential, logistic and θ-logistic, conform just a part (in fact a minority) of the possible outcomes of the system. They were generated by some non negative combinations of parameters θ and ρ. These, and other dynamics admitting also negative values for θ and ρ, conform a dynamical system that generalises most classic types of population growth (Table 1.1). In other words, the model presented herein is a unification of several growth dynamics. Resuming, population growth can be described in a general form by the two equations:

$$\dot{x} = xr \ , \tag{1.7a}$$

$$\dot{r} = (\theta r - \rho)r \ , \tag{1.7b}$$

refereed to from now on as *growth equation* and *rate equation*, respectively.

Although these two equations entirely describe the dynamics of a population, the rate equation (1.7b) is not coupled to the growth equation (1.7a), thus the entire dynamic is determined by the rate equation - actually this is true even considering the explicit form of $r(x)$. The growth equation (1.7a) can be written in a *per capita* form:

$$\frac{1}{x}\dot{x} = r \ . \tag{1.8}$$

It is straightforward that the *per capita* growth is a function only of the rate equation (1.7b), which determines solely and entirely the individual reproduction.

Although the rate equation (1.7b) indicates that regulation mechanisms are independent of population size, *Per capita* response is a contribution of both, individual reproduction (related to the parameter ρ) and interaction with other individuals (related to the parameter θ).

The units of the parameter ρ are inverse of time (frequency), and it gives the characteristic time scale at which individuals down-regulates the reproduction rate when, for example, the population approaches an equilibrium state like carrying capacity or extinction.

The parameter θ, is non-dimensional, but it sets the density scale at which the interaction of an individual with the population affects its reproduction rate.

1.4 STABILITY ANALYSES

The rate equation (1.7b) encloses all the information of the fixed points of the population dynamics. The rate equation has two fixed points, namely, $r_0 = 0$ and $r_1 = \rho/\theta$. Intuitively r_0 corresponds to the non-trivial equilibrium point of the growth equation, i.e. when the rates become zero the population is in a stationary equilibrium between reproduction and mortality. This means that r_0 is a steady state under balanced regulation.

Take for example the explicit form of the rate for the $\theta-$logistic equation

$$r_0 = 0 = r(x^*) = \frac{\rho}{\theta}\left(1 - \alpha(x^*)^\theta\right)$$

that implies $x^* = \alpha^{-1/\theta}$, and which corresponds to the carrying capacity. In this case, the population has a finite size, regulated by reproduction (replication at the individual level) and mortality (competition at the population level).

The biological meaning of the second equilibrium point, $r = r_1$, is not so obvious. From well-known cases, like the $\theta-$logistic equation, it is possible to realise that the population has a fixed point in $x^* = 0$. However, the fixed point given by the rate equation (Eq. 1.5) means that

$$r_1 = \frac{\rho}{\theta} = r(x^*) = \frac{\rho}{\theta}\left(1 - \alpha(x^*)^\theta\right) \tag{1.9}$$

which implies directly $x^* = 0$. Thus $r^* = r_1$ is equivalent to $x^* = 0$

The system (2.1) suggest a third fixed point: $(x^*, r^*) = (0, 0)$. This point, however is a paradoxical point, since both, the rate and the population size cannot be simultaneously zero, unless $\rho, \theta,$ and α are zero.

The stability of these fixed points, can be studied with the eigenvalues method. The Jacobian matrix of the system (2.1) is

$$\mathbb{J}(x, r) = \begin{pmatrix} r & x \\ 0 & 2\theta r - \rho \end{pmatrix} \tag{1.10}$$

Evaluating the Jacobian in the first fixed point, $P_0 = (x_0, r_0) = (x^*, 0)$ (where x^* is the asymptotic value obtained by equating $x^* = r^{-1}(0)$), leads to the eigenvalues:

$$\begin{cases} \lambda_0^x &= 0 \\ \lambda_0^r &= -\rho \end{cases} \tag{1.11}$$

Now, evaluation of the Jacobian matrix in the second fixed point, i.e. $P_1 = (x_1, r_1) = (0, \rho/\theta)$ gives the eigenvalues:

$$\begin{cases} \lambda_1^x &= \rho/\theta \\ \lambda_1^r &= \rho \end{cases} \tag{1.12}$$

The notation for the eigenvalues, λ_i^j refers to the eigenvalue associated with the j $(= x, r)$ coordinate of the fixed point P_i $(i = 0, 1)$.

Since ρ and θ can take any real value, the stability of the fixed points P_0 and P_1 depends on the signs of these two parameters.

However, some properties are already evident. On the one hand, none of the fixed points can be foci, since the eigenvalues cannot take imaginary values. The first fixed point, P_0, has a null eigenvalue, λ_0^x, indicating that there is a invariant set

$x^{inv} \in (x, r)$ that has "null" stability. The meaning is that every trajectory that intersects x^{inv}, (a) depends entirely on the initial conditions $(x(0), r(0))$, and (b) is fixed. Actually, the dynamical equations imply this invariant set corresponds to $x^{inv} = 0$.

In the growth dynamics, the explicit rate equations involve a constant α which does not appear in the implicit form of the dynamical system (2.1). Integrating the growth equation leads naturally to the constant α, as it did in the previous calculations of the exponential and θ-logistic equations. The value of α is related to the carrying capacity in sigmoid dynamics. By itself α plays no particular role in the stability of the system. It is a consequence of the dynamics of r rather than its cause, and thus it depends on the extrinsic factors that determine the initial conditions of the rate equation. The point at which the trajectories intersect x^{inv} correspond to the carrying capacity of the system, and it is a function only of α and θ.

According to the signs of ρ and θ, equations (1.11-1.12) indicate that there are four distinct possible sign combinations for the eigenvalues. For each of these combinations, termed regimes, particular patterns in the trajectories occur. This suggests several types of equilibria, convergence to equilibrium (which corresponds to growth dynamics) and transitions between the distinct regimes or types of equilibria (bifurcations).

Figure 1.1 shows how P_1 varies in the parameter space (ρ, θ).

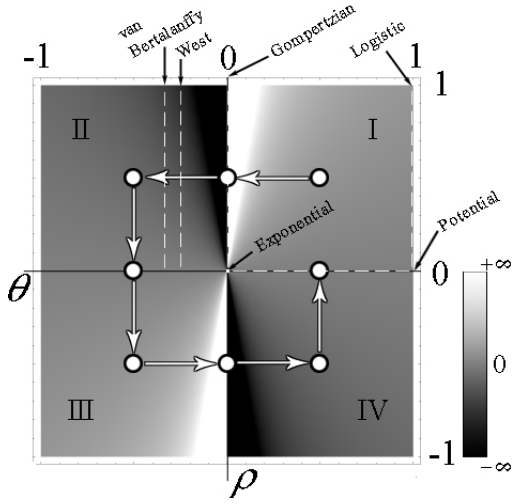


Figure 1.1: (Left) The value of the fixed point r_1 (shown as a density) as a function of the parameters ρ and θ. The four dots in the four quadrants, are the main four regimes, and the four dots over the axes are bifurcation points. The arrows follow the explanation in section 1.4. In quadrant I, r_1 is an unstable node; the point at $\theta = 0$, following the arrow, is a bifurcation point (r_1 is at infinity). In quadrant II, r_1 is a saddle, with the stable variety in the x coordinate. Following the arrow to $\rho = 0$ another bifurcation point is found. The invariant set x^{inv} also changes from attracting to repulsive in this point. Continuing to quadrant III, r_1 continues to be a saddle, but the stable variety is now at the r coordinate. Following the arrow to the point at $\theta = 0$ another bifurcation is found. The stability of this point changes in quadrant IV to a stable node. Finally, returning to quadrant I, another bifurcation a $\rho = 0$ changes the stable node to an unstable node, and x^{inv} changes to be again attractive. Particular examples of the growth rates listed in table 1.1 are represented with white dashed lines Potential growth comprises parabolic ($|\theta| < 1$) and hyperbolic ($|\theta| > 1$) replication.

1.4.1 Stability when $\rho, \theta \geq 0$

Beginning with ρ and θ both positive (first quadrant in fig 1.1), the phase space in this regime shows that P_1 is an unstable node. The line x^{inv}, termed *stable manifold*, attracts the trajectories close to it.

Figure 1.2 represents the phase space for this regime. Orbits with initial conditions such that $r(x_0) < \rho/\theta$ are attracted to x^{inv}. As discussed above, the point at which the orbits intersect x^{inv} correspond to the equilibrium value of x, i.e. the carrying capacity of the population.

This selection of parameters correspond to the θ-logistic equation. Population description based in this kind of growth range from flies (Gilpin and Ayala, 1973) to mammals (Saether et al., 2002), and also includes the classical version of the logistic growth.

If the initial conditions are such that $r(x_0) > \rho/\theta$, then the orbits are upper unbounded, and growth is unlimited. The growth for these region of the phase space is faster than exponential and any potential growth.

The line $(x, \rho/\theta)$ is the separatrix for the two possible dynamics.

Maintaining $\rho > 0$ and decreasing $\theta \to 0$, the systems shows a discontinuous (i.e. first-order) transition (Fig. 1.1). The value of the fixed point P_1 increases as θ decreases, and at $\theta = 0$ the fixed point disappears at infinity. This is a bifurcation point: on its right (positive perturbation to θ), the fixed point becomes an unstable node, and at its left (negative perturbation to θ), the point becomes a saddle (this will be discussed in the following sub-section).

When $\theta = 0$ the stable variety x^{inv} still remains and retains its stability (note that the eigenvalues associated to x^{inv}, λ_0^x does not depend on θ). All orbits in the system, converge to x^{inv}.

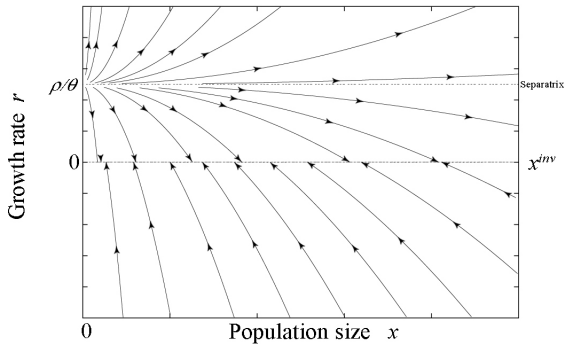


Figure 1.2: This regime of the phase space corresponds to the first quadrant in fig 1.1, where $\rho, \theta > 0$. There are two types of growth. If the initial conditions are below the separatrix at $r = \rho/\theta$, then the population grows as a θ-logistic. The intersection of the orbits at $r = 0$ correspond to carrying capacities. If the initial conditions are above the separatrix, then the growth shows synergistic interactions between individuals which are improve population increase.

This bifurcation point, for any initial condition, corresponds to Gompertzian growth.

This kind of growth, has been widely applied in tumour biology to investigate distinct aspects of tumour response and regression (Norton et al., 1976; Norton and Simon, 1977), as well as microbiological models (Kozusko and Bajzer, 2003). Molski and Konarski (2003) demonstrated that the Gompertz equation can be interpreted as the result of self organisation (cooperativity), in such a way that the individual response is correlated to the state of the whole population. This scale-wide correlation is a signature of criticality (Kadanoff, 2000).

1.4.2 Stability when $\rho \geq 0$, $\theta < 0$

If $\theta < 0$, and maintaining $\rho > 0$ the properties of the system change. In this regime (quadrant II in fig. 1.1), P_1 has one negative and one positive eigenvalue (Eq. 1.12) , meaning that it is a saddle. The invariant manifold x^{inv} still retains it attracting stability for the x coordinate (Fig. 1.3).

If the initial conditions $r(x_0) > -\rho/\gamma$, then the orbits correspond to a saturated growth and intersect x^{inv} at the carrying capacity.

It can be shown (see "Discussion") that ontogenetic growth laws, like van Bertalanffy's equation (von Bertalanffy, 1957), or the model by West et al. (2001), are included in this regime of parameters.

If the initial conditions are $r(x_0) < -\rho/\gamma$, the rate decreases $r \to -\infty$ asymptotically (fig. 1.3) and the convergence to $x \to 0$ is in a finite time t_c given by

$$t_e = \frac{1}{\rho} \log \left(\frac{r(x_0)}{r(x_0) - \frac{\rho}{\theta}} \right) .$$

This extinction happens because the regulation decreases hyperbolically in time to an asymptote at t_e.

Returning to the phase diagram, from this point in the second quadrant (fig. 1.1), decrease ρ to zero while maintaining $\theta < 0$. The transition to $\rho = 0$ is continuous (or a *second order* transition). The stable varieties x^{inv} and r_1 collapse onto each other.

Because at this point $\lambda^x_{(1)} = \lambda^r_{(1)} = 0$, we cannot infer about the dynamical properties of the nullcline at $r = 0$. However, with a perturbation on each side of the fixed point, it is possible to determine the stability.

If Δr is the perturbation (a trajectory slightly displaced from

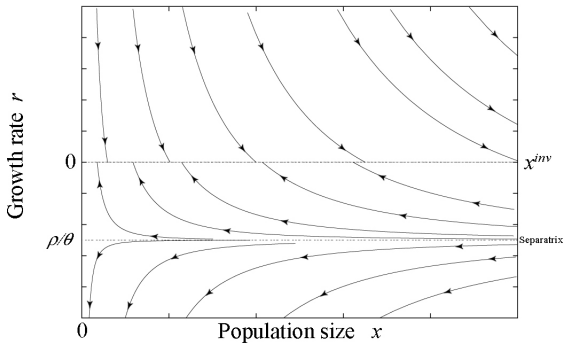


Figure 1.3: In this regime, corresponding to quadrant II in figure 1.1 with $\rho > 0$ and $\theta < 0$ two types of growth are possible. The initial conditions above the separatrix at $r = \rho/\theta(< 0)$ reproduce sigmoid growth curves which converge to a carrying capacity that corresponds to the intersection of the orbits at $r = 0$. If the initial conditions are below the separatrix, then the interactions are anergistic and the population decreases hyperbolically , and become extincted in a finite time.

zero), the rate equation for the point at $\rho = 0$ and $\theta < 0$ is

$$\dot{\Delta r} = -|\theta|\Delta r^2 \ ,$$

thus the system always responds diminishing the rate. Because rates can be negative (decreasing) the result is different if $\Delta r > 0$ or if it is $\Delta r < 0$. Consider the solution to the perturbation to the rate equation (note that there is no "first order" approximation in the rate equation):

$$r(t) = \frac{\Delta r}{1 + \Delta r|\theta|t} \ .$$

If the perturbation is positive, the rate will be damped to zero asymptotically. This means that the population will grow potentially. If the perturbation is negative, the rate will decrease to $-\infty$, in a finite time given by $t_e = (\Delta R|\theta|)^{-1}$.

Thus the stable variety repels the orbits on its left (initial conditions $r(x_0) < 0$, and asymptotically attracts the orbits on its right (initial conditions $r(x_0) > 0$).

Potential growth has been one of the corner stones of allometry (Peters, 1983; Calder, 1984), where pre-reproductive growth is assumed to potential (Roff, 1986; Day and Taylor, 1997, and references therein). Also, certain types of tumours have been inferred to grow potentially (Hart et al., 1998).

This kind of growth law has been sub-clasified into parabolic growth (if $|\theta| < 1$) and hyperbolic growth ($|\theta| > 1$). These situations describe different phenomena. The former models very well the dynamics of prebiotic replicators, as DNA self-replicators and quasi-species (Szathmáry and Demeter, 1987a; Szathmary, 1991; Scheuring and Szathmary, 2001). The hyperbolic case, well models the need of several (more than one) individuals in order to produce offspring. Such is the case of sexual reproduction, and also of hypercicles (Eigen and Schuster, 1979; Szathmary and Gladkih, 1989; Szathmary, 1991)

1.4.3 Stability when $\rho, \theta \leq 0$

In the phase space at the third quadrant, still maintaining $\theta < 0$ and now making $\rho < 0$, implies that r_1 is again positive. In this regime r_1 is a saddle point. It attracts trajectories from both sides in the r coordinate, but at the same time, repels the trajectories to $x \to \infty$.

The stable variety x^{inv} is a separatrix, which repels the orbits on its neighbourhoods. Figure 1.4 shows several trajectories for this regime.

Proceeding in direction to the fourth quadrant, maintaining $\rho < 0$ and decreasing $\theta \to 0$, once again a first order (discontinuous) transition is found, where $r_1 \to \infty$. x^{inv} still repels the trajectories from each side. This dynamic corresponds to

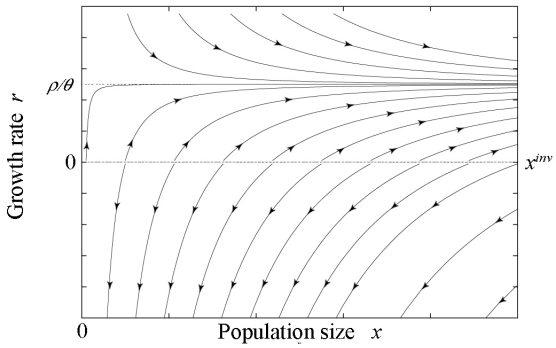


Figure 1.4: An invasion-extinction regime, corresponding to quadrant III of Fig. 1.1, where $\rho, \theta < 0$. This regime represents dynamics that show threshold behaviour. If $r(0) > 0$ the population can invade asymptotically-exponential with rate ρ/θ. If initial conditions are not appropriate, i.e. $r(0) < 0$ the population may go extinct in finite time with the rate decreasing hyperbolically .

a functional form similar to Gompertzian growth, but with a negative rate.

1.4.4 Stability when $\rho < 0$, $\theta > 0$

This last regime is characterised by being the only one having the point r_1 stable (fig 1.5). For those initial conditions such that $r(0) < 0$ (i.e. negative rates), the population decrease to zero sigmoidally.

If the initial conditions are $r(0) > 0$, then the population increases hyperbolically to $x \to \infty$.

The last two regimes are difficult to identify in real life data from their qualitative behaviours. They both comprise two different types of invasion and extinction dynamics, which may have different consequences in ecological contexts. The regime

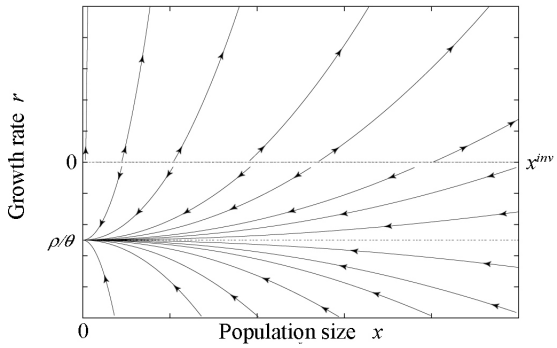


Figure 1.5: An invasion-extinction regime corresponding to quadrant IV in Fig. 1.1 with $\rho < 0, \theta > 0$. This regime is of threshold type. Invasion occurs in a synergistic way when $r(0) > 0$, although when the conditions are met for extinction, i.e. $r(0) < 0$ the population vanishes decreasing exponentially.

in quadrant III of Fig. 1.1 include invasions that are asymptotically exponential, but in quadrant IV, invasions are faster than exponential. The extinctions for quadrant III are in finite time, since they accelerate violently, while the extinctions in quadrant IV are slower, decreasing exponentially.

1.5 SCALING LAWS

Because the rate equation is not coupled to the growth equation (Eqns. 2.1), it is possible to find a general form for the solution to the rate equation (1.7b). It is convenient to replace $r \to \theta r$, which gives

$$\dot{x} = \frac{1}{\theta} r x \ , \tag{1.13a}$$

$$\dot{r} = r(r - \rho) \ . \tag{1.13b}$$

Denoting the solution for the rate equation as $R(t)$, then the solution for the growth equation then becomes

$$x(t) := x(0) \exp\left(\frac{1}{\theta}\int_0^t R(s)ds\right) \ , \tag{1.14}$$

where $x(0)$ is the initial condition for $x(t)$. Rearranging this system, we get

$$\left(\frac{x(t)}{x(0)}\right)^{\theta} = \exp\left(\int_0^t R(s)ds\right) \ . \tag{1.15}$$

The right-hand side of the last equation is independent of θ. Thus populations described by the system (2.1) are always scalable to their initial sizes, and interaction exponent θ.

Further rescaling is possible for the right-hand side. The solution to the rate equation can be written in the form

$$R(t) := \rho\left[1 + e^{\rho t}\left(\frac{r(0)}{\rho} - 1\right)\right]^{-1} \ , \tag{1.16}$$

and changing the time variable as

$$T \to \rho t - \log\left(\frac{r(0)}{\rho} - 1\right) \ , \tag{1.17}$$

and also changing properly the differential in the integral in equation (1.14) to $dT' = \rho dt$, then the result of the integral is, in scales of T

$$\left(\frac{x(t)}{x(0)}\right)^{\theta} = \frac{1 + e^{T_0}}{1 + e^{T}} \ . \tag{1.18}$$

The variable T depends on the initial condition $r(0)$ whose meaning is not so obvious.

Suppose that the population achieves a carrying capacity x_∞:

$$x_\infty = \alpha^{-1/\theta} \ , \tag{1.19}$$

using the explicit form of $r(t)$ given by equation (1.5), it results that the term in the rescaled time:

$$\frac{r(0)}{\rho} - 1 = \left(\frac{x(0)}{x_\infty}\right)^{-\theta} - 1 \ . \tag{1.20}$$

Using the last result in the transformation (1.17) and rearranging terms:

$$\left(\frac{x(t)}{x_\infty}\right)^{-\theta} = 1 - \left(\left(\frac{x(0)}{x_\infty}\right)^{-\theta} - 1\right) e^{-\rho t} \ . \tag{1.21}$$

To conclude, define the new scaled variables as:

$$\begin{cases} \chi &= \left(\frac{x(t)}{x_\infty}\right)^{-\theta} \\ \tau &= \rho t - \log\left(1 - \left(\frac{x(0)}{x_\infty}\right)^{-\theta}\right) \end{cases} \ , \tag{1.22}$$

with which the general scaling law obeys

$$\chi = 1 - e^{-\tau} \ . \tag{1.23}$$

Scaling at the bifurcation points

For the critical points, the scaling law above does not directly hold. As it will be shown at the end of this section, there is a clear relation between the scaling law (1.27) and the scaling laws at the critical points. However, it is first necessary to derive the scaling laws for Gompertzian and Potential growths from the solution of their growth rates.

Potential Growth

The solution for the growth dynamics for the potential growth is given also by equation (1.14). Consider then, that the solution to the rate equation is

$$R(t) := r(0)\left(1 - r(0)t\right)^{-1} \ . \tag{1.24}$$

In this case, the initial condition $r(0)$ is not expressed in terms of a carrying capacity x_∞, but only on the integration constant α. Upon integration and rearranging of terms, the following form is found:

$$\left(\frac{x(t)}{x_\infty}\right)^{-\theta} = 1 - \alpha t \ . \tag{1.25}$$

The scaled variables can then be defined as

$$\begin{cases} \chi &= \left(\frac{x(t)}{x_\infty}\right)^{-\theta} \\ \tau &= \alpha t \end{cases} \ , \tag{1.26}$$

with which the general scaling law obeys

$$\chi = 1 - \tau \ , \tag{1.27}$$

that is simply a decreasing line.

Gompertzian Growth

Consider now the solution to the size equation for the Gompertzian Growth

$$x(t) = x(0) \exp \int_0^t R(t')dt' \ . \tag{1.28}$$

The main difference between equation (1.14) and the last equation is that the former can be scaled with the exponent θ, which for the Gompertzian Growth case is zero.

Consider then the solution $R(t)$ to the rate equation:

$$R(t) := r(0)e^{-\rho t} \ . \tag{1.29}$$

Thus the general solution for Gompertzian Growth is

$$x(t) = x(0) \exp\left[\log\left(\frac{x(0)}{x_\infty}\right)\left(e^{-\rho t} - 1\right)\right] \ , \tag{1.30}$$

where the initial condition $r(0)$is expressed as:

$$r(0) = -\rho \log \frac{x(0)}{x_\infty} \ . \tag{1.31}$$

Equation (1.30) can be rearranged to give

$$\log \frac{x(t)}{x_\infty} = \exp -\rho t + \log\log \frac{x(0)}{x_\infty} \ . \tag{1.32}$$

Thus, defining the dimensionless variables:

$$\left\{ \begin{array}{rcl} \chi & = & \log \frac{x(t)}{x_\infty} \\ \tau & = & \log\log \frac{x(0)}{x_\infty} \end{array} \right. \ , \tag{1.33}$$

then the scaled dynamics results as

$$\chi = e^{-\tau} \ , \tag{1.34}$$

that is a decreasing exponential.

Although the complete scaling behaviour is completely defined for all the dynamics in the phase space, there is still an interesting question to be asked: Is it possible to derive the scaling behaviour for the critical points from the "general" scaling form (1.27) and its non-dimensional variables?

Direct evaluation of $\rho = 0$ or $\theta = 0$ does not give the scaling laws for Gompertzian or Potential Growth. However, Taylor expansion *on the parameters* ρ and θ to the linear term around the

critical values, do give the scaling laws. Note that in order to give a precise meaning to the transformations (1.22), the initial condition for the rate equation was transformed to its explicit form. It is necessary, however, to make this transformation *after* the limits are taken in the scaled variables (1.22).

1.6 DISCUSSION

The simple model derived in this paper is rich in qualitative solutions since it resumes several growth rates that often appear in the literature, which include several levels of biological organisation. Several examples were alluded to in the text. These examples range from the origins of life, cellular populations of procariots, cellular populations in eucariots, in ontogeny and cancer, to population biology of mammals and birds, to community ecology. It is a nice result that all of these kinds of growth can be described by such simple equations that resumes the main features of populations, in the traditional sense of showing density dependence, and in the distinct interpretations introduced in this paper.

The θ-logistic equation has become a paradigm in ecology. Modelling populations with it has been an important tool to confront actual problems about density-dependent ecology. The transformation introduced in this paper gives a good insight into the meaning of the quantities appearing in the equations, namely ρ, θ, and α (either in its interpretations as carrying capacity or not).

Population growth, as proposed in Eqns. (2.1), represents carrying capacity as an initial condition for the rate equation. However, the regulation mechanism is independent of α, meaning that it is dynamically independent of the environmental conditions, and therefore making it a mechanism that is completely

intrinsic of a population. The actual size of the population is then result of both, the environment -determining the initial conditions-, and the growth rate -intrinsic mechanism-. This is a new result in the sense that previous formulations included the carrying capacity as a predefined constant, which was of course assumed to be environment-dependent, but which was not totally separated from the dynamics of the population as (MacArthur, 1962) pointed.

This property of the populations –to have a size-independent regulatory mechanism– can give new insights to evolutionary biology, because it makes population models adaptive, while the mechanism for regulation remains robust against environmental changes.

1.6.1 Extinctions and Invasions

Suppose that a population is in (or fluctuating around) a fixed point of x^{inv}. Environmental changes, are know to "unbalance" some populations. An environmental change can be traduced in translating the dynamics, formerly in a fixed point, to a lower (or higher) value of r, but maintaining the same value for x. As a consequence, the dynamic is placed in another orbit out of equilibrium, and the response is to decrease (or increase) population size to a lower (or higher) carrying capacity. However, if the perturbation is strong enough, then the orbit where the dynamics is placed could be part of a basin of attractions that does not include the stable population size, and which leads either to extinction or invasion.

As an example, consider global warming. This has become an important issue in the last years. There is the open question about how temperature increments may affect populations. With the results of this paper, it is possible to evaluate the consequences of temperature increase on a population.

Brown et al. (2004) proposed a theory in which the carrying capacity of a population is temperature-dependent through a Boltzmann factor. This dependence can be expressed as $K = K_0 \exp(E/kT)$, where K_0 is a parameter depending on mass, resources, etc., E is the energy of the limiting metabolic reaction, k is Boltzmann's constant, and T the absolute temperature.

Consider a population that is in its carrying capacity at a temperature T_0. Then $x^* = K_0 \exp(E/kT_0)$, and the rate at equilibrium is $r^* = 0$. If suddenly the temperature increases to T_1, the rate will change to

$$r = \frac{\rho}{\theta}\left[1 - \exp\left(\frac{\theta E}{kT_0T_1}(T_1 - T_0)\right)\right] < 0 \ .$$

This means that the population is taken out from equilibrium. Its response is to relax to a new (smaller) carrying capacity. However, from this equation it is straightforward that temperature alone, cannot induce a change such that the population goes to extinction. This is because to induce an extinction, the exponential term would have to change sign, thus to have new initial conditions in a basing of attraction decreasing to $(x, r) \to (0, -\infty)$, which is not possible for any temperature.

Similar examples exist with laboratory cultures of protozoans, fungi, and procariots which are limited by available nutrients. These are examples of external factors that determine the values of α.

It is always possible however, to consider such changes that although they theoretically do not imply extinctions, numerically are so small that in real life populations could disappear.

The saturated dynamics, comprised in quadrants I and II of Fig. 1.1, have other co-existing behaviours. In the case of quadrant I, there is the possibility of invasion (Fig. 1.2). To make this possible, it is necessary that external perturbations induce cooperativity among individuals, rather than competition for re-

sources, and would be indicated by $\alpha < 0$.

On the other hand, to induce an extinction in quadrant II, it is necessary that individuals become aggressive and any interaction results in mutual annihilation (Fig. 1.3), indicated by $\alpha > 0$.

The conditions for these situations (i.e. in which basin of attraction are the initial conditions) have a similar interpretation to Hamilton's rule for kin selection (Hamilton, 1963, 1964). Hamilton's rule points out that if the cost of an altruistic behaviour is such that it benefits a genetically related individual, then the strategy can be selected. In this way, a population consisting of cooperative individuals can spread faster than expected by exponential models (Fig. 1.2). However, Hamilton's rule has another "solution". This is that aggressive behaviours can also be selected, provided that damage is induced to "negatively related" (i.e. unrelated) individuals.

Recently, Gardner and West (2004) reported an example of this aggressive behaviours in wasps. It would not be surprising if local populations self-annihilate under certain demographic conditions.

Another example, mentioned in the text above, is the extinction of sparrows (Saether et al., 2000). In order to have the risk of finite-time extinctions from previously stable populations, it is necessary that (a) the population dynamics belongs to quadrant II of Fig. 1.1, and (b) there is a perturbation such as mentioned above. The estimated mean value for the population of sparrows is $\hat{\theta} \simeq 1$, indicating that the population is logistic. However, the estimated distribution allows a small but not negligible probability for $-1.5 < \theta < 0$. If this is the case, then a real risk of finite time extinction exists.

1.6.2 Life histories

In life history theory, the central problem is the allocation of resources for adaptive strategies. Survival and reproduction in distinct stages are the determinants of the growth function (Day and Taylor, 1997; Stearns, 2004).

Resource allocation in non-reproductive stages of life, i.e. before maturity, are greatly devoted to growth. However, how much energy is allocated depends on the size of the body. This dependence is typically described by a potential growth function, where the exponent θ is indicative of some length scale of the physiological processes allocating the energy devoted to growth (Calder, 1984; Stearns, 2004).

According to the rate equation, initial stages in growth have to be dominated by the term θr^2. In order to allow this term to be dominant at low densities, the exponent has to fulfil $\theta < 0$. Actually, this condition is met in in the von Bertalanffy (1957) and West et al. (2001) equations, and well as in most allometric growth relationships (Calder, 1984).

However, when reproduction becomes a priority in the life history, the energy income has to be partitioned according to a survival-reproduction compromise. In this case, that is when the term ρr is not negligible, individuals are distributing the energy between reproduction and survival.

1.6.3 Allometry and scaling

Other important results are the scaling laws derived in section 1.5. These general formulas show that scaling is not a particularly eccentric result. It is rather a rule than an exception that growth can be scaled. This of course does not invalidate the underlying theories of resource allocation, at any level, bioenergetic, or ecological. However it rather gives an broader view to the open discussion of whether the West et al. (2001) equation is legitimate or not (Kozlowski and Konarzewski, 2004). In terms of these formulations, although numerically different, West's equations and its classical competitor, van Bertalanffy's equation, have the same qualitative behaviour.

The model (2.1) reproduces the famous von Bertalanffy (1966) equation. This equation can be written in the following form:

$$\dot{x} = ax^{2/3} - bx \ . \tag{1.35}$$

It is possible to rearrange this equation, to express it as the system (2.1), for which the parameters are then:

$$\rho = \tfrac{b}{3} \quad , \quad \theta = -\tfrac{1}{3} \ . \tag{1.36}$$

The exponent $-1/3$ follows from the hypothesis that mass is proportional to the third power of length, and the parameter b is related to individual reproduction. The parameter a is related to the carrying capacity of the population, thus it does not appear in the transformations.

Another example in this regime is West's ontogenetic growth equation (West et al., 2001), given by

$$\dot{x} = am^{1/4}\left[1 - \left(\frac{x}{x_\infty}\right)^{1/4}\right] \ , \tag{1.37}$$

The parameters for the rate equation are then:

$$\rho = \tfrac{a}{4x_\infty^{1/4}} \quad , \quad \theta = -\tfrac{1}{4} \ . \tag{1.38}$$

The exponent $-1/4$ derived from fractal patterns of fluid transport systems like circulatory system or plant vascularisation (West et al., 1997; Brown et al., 2004) and is also supported by empirical data (West et al., 2001). The parameter ρ in this case depends explicitly on the carrying capacity x_∞. This suggests that in ontogenetic growth, these two quantities could be correlated.

If this is the case, then there would be further important implications for allometry, because it would imply that the macroscopic growth i.e. the cell population, is "transmitting" information to the microlevel, i.e. single-cell dynamics. and thus it could imply existence of self-organisation. This however should be verified from experimental data, and theoretical models.

Both cases, von Bertalanffy and West equations, are used to model ontogeny. Although they are derived with distinct assumptions, under proper scaling, they follow exactly the same law (Eq. 1.27).

1.6.4 Cancer and Diseases

Cancer research is a particular subject in which mathematical models are applied (Wheldon, 1988). The growth of tumours have been very well studied directly *in vitro* and in experimental frameworks. Solid tumours grow according to a Gompertzian Law, as it was shown by Norton et al. (1976). This deterministic description of growth has been of great impact in medicine, and has even lead to important conclusions about treatment scheduling.

The Gompertzian law is a sigmoid curve that grows toward a carrying capacity. In the rate-based scheme the orbits are attracted to the invariant set x^{inv} for any initial condition. Thus the dynamic of solid tumours comprises only growth towards carrying capacity, since all the resources in a tumour are de-

voted to reproduction, as indicated by $\theta = 0$. However, the rate equation admits another Gompertzian solution for $\rho < 0$. In this case, there is regression of the tumour size for initial conditions $r(0) < 0$. This condition is fulfilled if a strong enough therapy is applied (González et al., 2003; de Vladar and Gonzalez, 2004). The analysis of the Gompertzian model by Molski and Konarski (2003) supports this regression solution.

1.7 CONCLUDING REMARKS

There are, of course, more examples for each of the growth types described in this paper. However, more interesting is that there are regimes that have not been reported. This is not surprising, because they conform distinct types of indeterminate growth, which usually is assumed to be "exponential". However, these distinct types of explosions, can have important consequences in disciplines like biotechnology, where a strict control of growth is necessary. If by some reason, a population is wrongly manipulated, such that it spreads "indeterminately" then the distinct types of growths should be managed distinctly.

However, in the opinion of the author, the most important result is that the rate equation is explicitly independent of the population size. The results presented in this paper, are derived from a simple mathematical transformation, which surprisingly result in a very broad class of regulatory mechanism. Although this is a result that may apply only to the simple systems included in this work, it is puzzling why and how the regulatory mechanisms act.

Actually, the two terms of the rate equation (1.7b), from a more abstract perspective, correspond to two processes that constitute regulation: reproduction, which comes from an individual level, and sensibility to population interaction.

Depending on the context, the sensitivity to the population can be of synergistic or anergistic nature. Actually, the two regulation processes, could be thought of as fragmentation and condensation reactions (Fontana and Buss, 1994; Szathmary, 1995):

$$\begin{array}{rcll} x & \xrightarrow{\rho} & 2x & \text{(reproduction)} \\ Nx & \longrightarrow & Mx & \text{(population sensibility)} \end{array} \tag{1.39}$$

If $M > N$ then the population experiences synergy in growth (i.e. population interaction promotes growth). If $M < N$ then the population experiences anergy in growth (i.e. population interactions avoids growth).

The term $-\rho r$ in the rate equation indicates that the population is growing or "relaxing" to a fixed point. Since ρ is the inverse of the relaxation time for the rate, then the bigger ρ is the smaller the time to let the mechanism to relax, and thus the fastest to reach limiting population size at P_1.

The second term of the rate equation relates to the interaction between individuals in the population. The square in the term means that the rate is auto-catalysed. Thus the parameter θ indicates the level of this auto-catalysis. This term can be compared to a "potential" indicating some kind of resource potentiating (either synergistically or anergistically) from the interaction. The kind of interaction, is given by the sign of θ and by the environmental conditions, i.e. by α.

The relationship between ρ and θ determining the distinct types of growth rates, gives distinct types of behaviours for distinct initial conditions. The auto-catalysed reaction can result or cooperative, competitive, or aggressive strategies. These are strategies that can be sought directly from the rate equation. If a population is behaving cooperatively, then it means that the rate begins over a threshold such that it grows unlimited, because there is a benefit improving growth resulting from the

interaction. In this case, the resources have to be unlimited, so cooperativity improves resource allocation for reproduction. In the case of competition, usually the scenario is that where a resources are limited, and there must be an equilibrium between reproduction and survival. But if the population presents aggressive behaviours, then the initial rate of the population is below the threshold where it goes extinct in finite time or exponentially.

The exponential growth is a particular case in which no regulation mechanism (thus interactions) is present in the population. Increase is based only in individual reproduction, and the *per capita* response is totally independent of the state of the system.

The distinct qualitative solutions (regimes) for population growth have an important underlying symmetry. The regression equation for a population shrinking is in general obtained by inverting the time arrow. In other words, changing $t \to -t'$ is equivalent to write the equations at which he population shrinks. However, time-inverting the rate equation does not produce the desired result. In order to obtain the regression dynamics from the size-rate decomposition, besides inverting the time arrow, it is necessary to invert the rate variable $r \to -r'$. Thus a time reversed equation results in the transformed system

$$\dot{x} = xr' \ , \tag{1.40a}$$

$$\dot{r}' = (\theta r' + \rho)r' \ , \tag{1.40b}$$

where it becomes obvious that, in order to obtain a time-regressed equation, we simply need to change the sign of the Malthusian parameter ρ. This same result is obtained changing the sign of ρ in any of the population dynamics equations where density dependence appears explicitly as a function of size x.

This transformation for population shrinking, can be regarded as a reflexion of the parameter space (Fig. 1.1) with respect to the θ-axis. Thus, although most of the relevant population dynamics are in quadrants I and II in fig 1.1, their corresponding regression dynamics are in quadrants III and IV.

The decomposition presented in this work, is a change in the paradigm of population dynamics. The equations (2.1) are very general, but still simplistic because there are many biological aspects left aside. Take for example the Allee effect (Allee, 1931). This density dependent growth mechanism, is not represented in the rate equation in the form presented in here. Actually, including Allee effect in population growth, leads to a polynomial equation for the rate equation (1.7b) that is in general of higher order than 2. There is however no general law that can be derived. It remains to investigate based on life history theory for which kind of resource allocation the density-independent rate equations can be derived. This is a work currently under development that is expected to help to drive conclusions about other biologically relevant aspects not included in this work.

Published as: H.P. de Vladar and I. Pen – *Determinism, Noise, and Spurious Estimations in a Generalized Model of Population Growth* Physica A. vol. 373, pp. 477–485, 2007.

Chapter 2

Determinism, Noise, and Spurious Estimations in a Generalized Model of Population Growth

> Let us therefore agree that the idea of eternal return implies a perspective from which things appear other than as we know them: they appear without the mitigating circumstance of their transitory nature
>
> Milan Kundera

Abstract

In this chapter we study a generalised model of population growth in which the state variable is population growth rate instead of population size. Stochastic parametric perturbations, modelling phenotypic variability, lead to a Langevin system with two sources of multiplicative noise. The stationary probability distributions have two characteristic power-law scales. Numerical simulations show that noise suppresses the explosion of the growth rate which occurs in the deterministic counterpart. Instead, in different parameter regimes populations will grow with "anomalous" stochastic rates and (i) stabilise at "random carrying capacities", or (ii) go extinct in random times. Using logistic fits to reconstruct the simulated data, we find that even highly significant estimations do not recover information about the deterministic part of the process. These results have implications for distinct model-aided calculations in biological situations because these kinds of estimations could lead to spurious conclusions.

2.1 INTRODUCTION

Population dynamics are frequently modelled with simple equations that mimic some aspects of replicating biological entities, such as division (in cells), fission (in modular organisms) or reproduction (in eukaryotes), competition, and population size limiting (saturation). These and other properties are represented by various models. Frequently the validity of these models is a matter of statistical goodness of fit with a specific data set. However, these biological properties are not entirely of intrinsic nature to the individuals, or to the populations themselves, but rather emerging ecological properties, i.e. the interaction between "individuals" and "environment". The models of population growth simplify (whenever it is possible) the potential complexity of a detailed ecological description into simple equations.

Recently we showed that a variety of biological growth models can be unified using a phase-space decomposition using two dynamical variables, population size x and growth rate r (de Vladar, 2006) (analogous to a particle' s position and momentum, respectively):

$$\dot{x} = xr \ , \tag{2.1a}$$

$$\dot{r} = r(\theta r - \rho). \tag{2.1b}$$

The constant ρ is the Malthusian parameter, and θ is the intraspecific interaction coefficient. By varying these two parameters it is possible to reproduce exactly a wide family of growth laws including exponentials, logistics (Ayala et al., 1973; Sibly et al., 2005), Gompertzian (Kozusko and Bajzer, 2003), Potential (Roff, 1986; Szathmáry and Demeter, 1987a), as well as allometric growth laws like Von Bertalanffy's (von Bertalanffy, 1966; Roff, 1986) and West's (West et al., 2001) equations, among others (de Vladar, 2006).

Although the deterministic behaviours can be associated with distinct biological scenarios (de Vladar, 2006), populations are often influenced by some source of noise. Biologically, random "forces" are often related to environmental and demographic fluctuations, as well as to intrinsic complex factor like genetic architecture, mating system, recombination, mutation, segregation, etc. These environmental and genetic factors (and their interaction) can be thought to be the determinants of the parameters that describe the growth of a population. This conjecture means that they are (complicated) measures of phenotypic expression. Thus to a first approximation we can model the effect of phenotypic variation as noise over these parameters.

We can consider two sources of noise $\eta_{\rho\ (t)}$ and $\eta_{\theta\ (t)}$ affecting respectively the parameters θ and ρ. The rate equation (2.1b) linearly perturbed with $\rho \to \rho + \eta_{\rho\ (t)}$ and $\theta \to \theta + \eta_{\theta\ (t)}$ results in

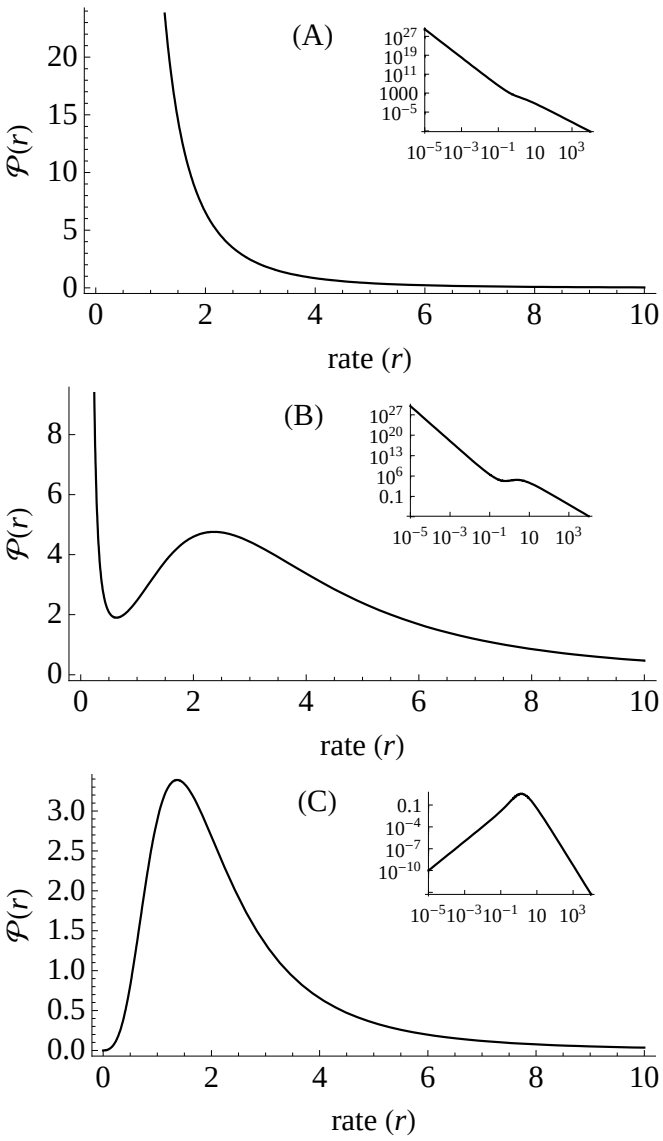
$$\dot{r} = r(\theta r - \rho) + r^2 \eta_{\theta\ (t)} + r\eta_{\rho\ (t)}\ , \tag{2.2}$$

where the $\eta_{i\ (t)}$ have the usual properties of white noise:

$$\begin{aligned} &\langle \eta_{i\ (t)} \rangle = 0, \qquad && \langle \eta_{i\ (t)} \eta_{i\ (s)} \rangle = \delta(t-s)\epsilon_i^2 \\ &\langle \eta_{\theta\ (t)} \eta_{\rho\ (t)} \rangle = \gamma \end{aligned}$$

with $\langle \ldots \rangle$ denoting expectations, and $i = \rho, \theta$.

Figure 2.1: (Right) Potential solutions for the stationary diffussion equation. (A) When $0 < \theta_0^2/8 < \epsilon_0^2(\rho_0 + 1)$, the distribution shows an exit barrier at $r = 0$; in this example $\rho = 2,\ \epsilon_0 = 1,\ \theta = 2$. (B) When $0 < \epsilon_0^2(\rho_0 + 1) \leq \theta_0^2/8$, besides the exit barrier, the distribution also shows an analytic maximum; in this example $\rho = 2,\ \epsilon_0 = 1,\ \theta = 6$. The abscise is in scaled by 10^4 are in (C) When $\epsilon_0^2(\rho_0 + 1) < 0 \leq \theta_0^2/8$, the exit barrier disappears, and the distribution is unimodal with an analytic maximum; the parameters are $\rho = -2,\ \epsilon_0 = 1,\ \theta = 2$. The insets plot the distributions in log-log scale, showing that there are two characteristic scales. The left tail scales with an exponent of $-2(1 + \rho_0)$ whereas the right tail scales with an exponent of -4. The inflection points are close to the maximum.


(A)
$\mathcal{P}(r)$
rate (r)
(B)
$\mathcal{P}(r)$
rate (r)
(C)
$\mathcal{P}(r)$
rate (r)

Here, ϵ_i are the intensities of the noise sources, γ is the correlation between the two noise sources, and $\delta(t)$ is Dirac's delta function.

The resulting system is a stochastic differential equation (SDE) where ρ is the drift term and θr^2 can be thought as the force of an external field, in the mathematical/physical terminology (Gardiner, 2004, notice that what in biology we define *drift* as the random component, whereas mathematicians and physicists refer with this word to the deterministic component). Multiplicative noise, often represents fluctuating barriers or processes of anomalous diffusion (i.e. diffusion where the probability of long steps is higher then in the normal case) (Fleming and Hänggi, 1993; Hänggi, 1994; Kaniadakis and Lapenta, 2000; Fa, 2003; Biró and Jakovác, 2005). Also, multiplicative noise is a process that retains memory (i.e. is non-Markovian), and has been investigated in the context of population growth and extinctions (Halley and Kunin, 1999; Ai et al., 2003; Wichmann et al., 2005).

Equation 2.2) remains uncoupled from the size x. This gives an operational advantage since the analyses of the SDE can be made in terms of r as a 1-dimensional system that is relatively simple to handle.

2.2 DISTRIBUTION OF THE GROWTH RATES

To study the effects of multiplicative noise, and make precise the meaning of "anomalous growth" in populations (in analogy to anomalous diffusion), first consider the probability distribution for the rates. In the Itô interpretation of noise the probability is given by the related diffussion equation (DE, know also

as Fokker-Plank equation):

$$\begin{aligned}\partial_t \mathcal{P}_{(r,t)} = &- \partial_r \left[r(\theta r - \rho)\mathcal{P}_{(r,t)} \right] \\ &+ \frac{1}{2}\partial_{rr} \left[(\epsilon_\rho^2 - 2\epsilon_\rho \epsilon_\theta \gamma r + \epsilon_\theta^2 r^2) r^2 \mathcal{P}_{(r,t)} \right] \ .\end{aligned} \tag{2.3}$$

Setting the time derivative equal to zero makes it possible to calculate the potential solution of the DE on the stationary regime $P(r)$ (i.e. equilibrium solution), which gives

$$\begin{aligned} P(r) \quad &:= \quad \mathcal{N} \left(r^{-2} \right)^{\rho_0 + 1} \left(\epsilon_0^2 - \epsilon_0 \gamma r + r^2 \right)^{\rho_0 - 1} \times \\ &\times \quad \exp \left[2(\epsilon_c \theta_0 - \gamma_c \rho_0) \tan^{-1} \left(\epsilon_c r - \gamma_c \right) \right] \ , \end{aligned} \tag{2.4}$$

where $\mathcal{N}$ is the integration constant, and

$$\begin{aligned} &\rho_0 = \rho/\epsilon_\rho^2 \ , && \theta_0 = \theta/\epsilon_\theta^2 \ , && \epsilon_0 = \epsilon_\rho/\epsilon_\theta \ , \\ &\epsilon_c = \epsilon_0^{-1}(1-\gamma^2)^{-1/2} \ , && \gamma_c = \gamma(1-\gamma^2)^{-1/2}. \end{aligned}$$

Fig. 2.1 shows that there are three distinct kinds of stationary distributions for the rates. The first thing that we note, is that the correlation γ modulates the transition from one distribution to another. Thus, for simplicity for the further analyses we proceed setting $\gamma = 0$.

The first distribution (Fig. 2.1A) is monotonous decreasing, and the other two (Figs. 2.1B-C) have an analytic maximum at

$$r^* = \frac{\theta_0}{4} + \sqrt{\left(\frac{\theta_0}{4} \right)^2 - \epsilon_0^2(\rho_0 + 1)} \ .$$

The condition for having an analytic maximum is

$$\frac{\theta_0^2}{8} \geq \epsilon_0^2(\rho_0 + 1) \ . \tag{2.5}$$

Fig. 2.2 outlines the regions where the inequality (2.5) holds. The parabolic curve (i.e. the boundary given by the equality in

Eq. 2.5) divides the parameter space into three regions with different stationary regimes. The first, when $0 < \theta_0^2/8 < \epsilon_0^2(\rho_0 + 1)$, corresponds to the space under the parabola (Fig. 2.2); in this case the probability density is accumulated at $r = 0$. The second region, characterised by $0 < \epsilon_0^2(\rho_0 + 1) \leq \theta_0^2/8$, corresponds to the space over or under the parabola region (Fig. 2.2). The third region, defined by $\epsilon_0^2(\rho_0 + 1) < 0 \leq \theta_0^2/8$, is the space on the left of the parabola. In the two last regions, the probability mass of rates is distributed along the axis, indicating that the growth rate can be persistent (i.e. non-zero). In the following of the paper, we will show that each of these regions have distinct qualitative solutions in which the deterministic nature of the process is "forgotten", but the resulting dynamics of the population size look like exponential or logistic dynamics. We will demonstrate however, that these two forms are entirely product of noise, hence fitting these models to the realizations -although statistically significant- are spurious.

For particular cases of Eq. (2.2) the stationary distribution has been calculated before. When θ and its noise η_θ term are absent, the equation recovers the geometric Brownian motion (Oksendal, 2002), whose stationary distributions were shown to have power-law tails (Biró and Jakovác, 2005). In this representation, the growth corresponds to a Gompertzian growth. A power-law-tailed distribution is also found for the stationary distribution of an equation where $\rho \neq 0$ but which is not perturbed (Góra, 2005). Also, a logistic case $\theta = 1$ was analysed by Morita and Makino (1986) using perturbation techniques for a time dependent solution. In log-log scale the distribution (2.4) is kinked near $r_c = \epsilon_0 \exp(\pi\theta_0/2\epsilon_0(1-\rho_0))$, with a right-tail decreasing in a power law fashion $\log P(r) \sim -4 \log r$ (insets in Fig. 2.1). This is a result that can be derived directly from the particular case studied by Góra (2005), because the right tail of the distribution is independent of the parameters. Moreover,

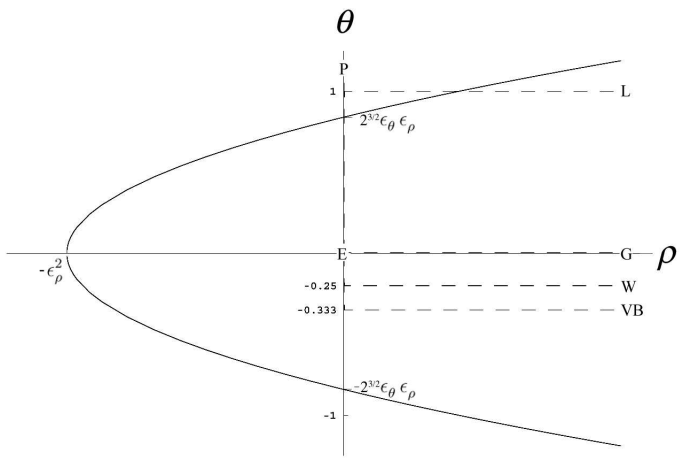


Figure 2.2: The parameter space (ρ, θ) consists of four quadrants, corresponding to their sign combinations. The dotted lines indicate distinct deterministic growth functions known in the literature: Potential (P) $\rho = 0$; logistic (L) $\rho > 0, \theta = 1$; Gompertzian (G) $\rho > 0, \theta = 0$; West (W) $\rho > 0, \theta = -1/4$; von Bertalanffy (VB) $\rho > 0, \theta = -1/3$; Exponential (E) $\rho = 0, \theta = 0$. The solid curve represents the noise-transition points between the three distinct regimes of the distributions of the rate: (a) inside the parabolic region $0 < \theta_0^2/8 < \epsilon_0^2(\rho_0 + 1)$; (b) above or below the parabolic region $0 < \epsilon_0^2(\rho_0 + 1) \leq \theta_0^2/8$; and (c) at the left of the parabolic region $\epsilon_0^2(\rho_0 + 1) < 0 \leq \theta_0^2/8$.

there is also a power law behaviour for small values of r which is given by $\log P(r) \sim -2(\rho_0 + 1) \log r$. These power law tails lead to Tsallis statistics (Anteneodo and Tsallis, 2003; Biró and Jakovác, 2005). Some relationships between exponents have been derived for a system related (but not equivalent) to ours (Genovese and noz, 1999).

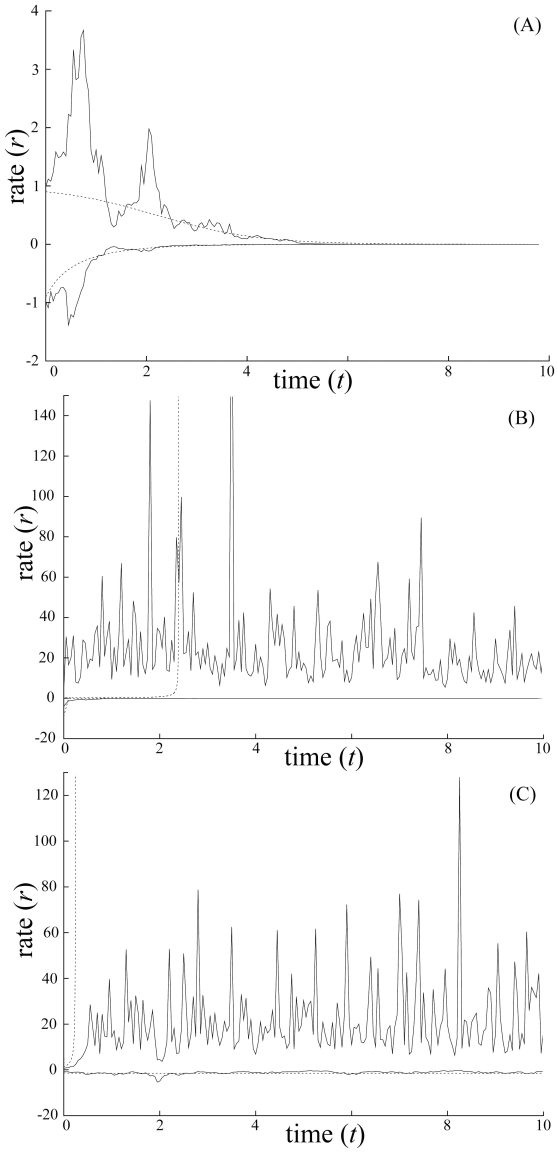
2.3 DENSITY REGULATION BY NOISE

The fast decrease of the right tail has an important consequence, which is the boundedness of the process. In other words, it means that the fluctuations remain finite. The simulations of Fig. 2.3A show that when the probability density is accumulated at $r = 0$ the rates will stochastically reach zero and stay there forever. Whenever this happens, the population freezes at its -random- current size.

Sibly et al. (2005) performed an analysis where they fitted more than a thousand population time series to the θ-logistic model. Their analysis was based on the size-dependent per capita growth rate $r(x)$. The examples they presented, show comparable patterns to the realizations obtained from our model (Fig. 2.3A). However, as we can see in this figure, the stochastic trajectories are notcentredd on the deterministic trajectories, as it is common for multiplicative perturbations. Therefore, the interpretation of the estimations in Ref. (Sibly et al., 2005) differ from the deterministic path, at least in the light of our model. We will return to this discussion later .

Other simulations, using a processes having stationary distribution with maxima, are shown in figure 2.3B-C. In this cases the rate does not explode in the time-window, even when these realizations (the deterministic and the stochastic) have the same initial conditions which would lead to explosions in the absence of perturbations.

Figure 2.3: (Opposite page) Realizations of noisy rate dynamics. The dotted lines show deterministic solutions, while the continuous bold lines show the stochastic realizations. It can be seen that when the deterministic rates decrease to zero, the stochastic dynamics will also decrease to zero. Also, when deterministic rates explode the stochastic dynamics remain finite. These realizations correspond respectivelyy to the regimes and parameters of Fig. 2.1


(A)
rate (r)
time (t)
(B)
rate (r)
time (t)
(C)
rate (r)
time (t)

Recently, Mao et al. (2002) demonstrated that the deterministic explosions of "positive" logistic equations (e.g. of the form $\dot{x} = ax(1 + bx)$) can be controlled with certain types of multiplicative noise sources. When these fluctuations are present, populations will not diverge in finite time, although their purely deterministic analogue does.

The rate-representation introduced in this paper is also of quadratic form, thus the results of Mao et al. (2002) apply to Eq. (2.2). However, the biological interpretations change, because explosions are suppressed in the rate rather than in population size.

As indicated by the distribution of the rates, probability is accumulated near the maximum, thus the rates will be non-vanishing, jumping from very slow to high (but finite) values, making the size of the population increase in bursts, reconstructing a devil's staircase pattern (a staircase where all the steps are of different size and height). Also, because the rate never reaches zero, the population grows unlimited.

The same distributions of Fig. 2.1 appear for negative rates. The course of the population is the opposite, i.e. decreasing, although the distribution is the same (in absolute value): (i)if the distributions have an analytic maximum, the rates will remain finite and fluctuating, meaning that population will decrease erratically but monotonously and therefore populations will become extinct in random times; (ii) if the distributions do not have an analytic maximum, then r reaches zero stochastically (Fig. 2.3A), and then the populations will stabilise, again at a "random carrying capacity".

The distribution (2.4) is not normalisable whenever $\rho_0+1 > 0$, because it diverges when $r \to 0$. This means that $r = 0$ is an exit barrier, and hence once the rate reaches zero it will stay there. This limit is the same if taken from the left, thus the rate cannot either jump to a negative value once it reaches zero.

For instance the rates maintain their sign or become null, but never change sign. The meaning is that when the rates are stationary, an initially growing populations will continue to grow, or at most, cease growing but they will not suddenly shrink. Therefore, the converse is also true: populations that started shrinking, will not suddenly change its course and grow. They will continue to shrink until extinction, or reach a stable value.

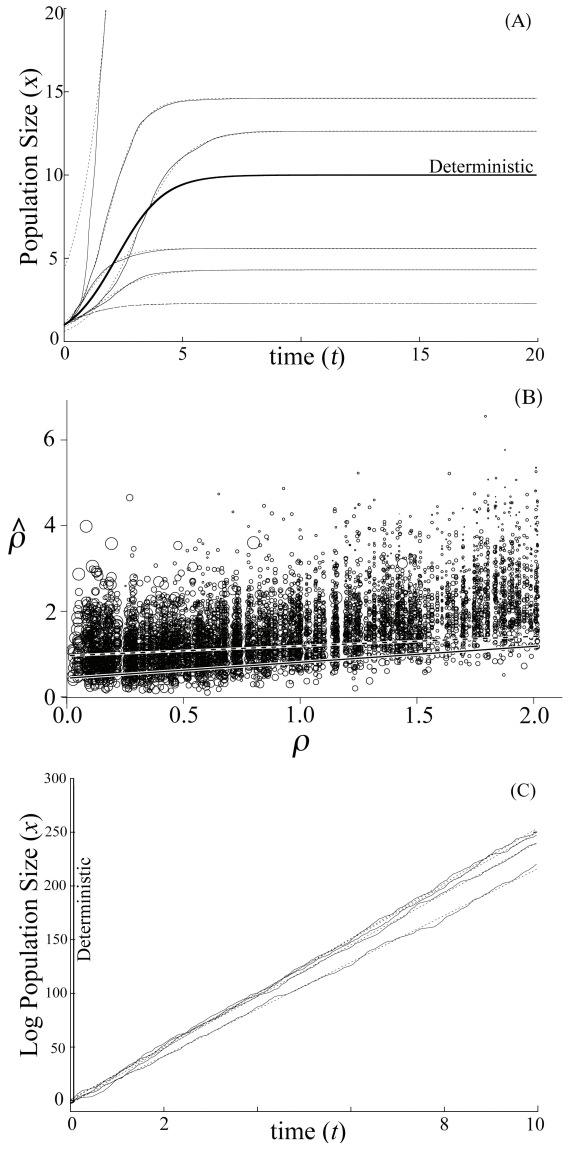
2.4 ESTIMATIONS OF GROWTH PATTERNS

The dynamics results in distinct realizations that can give drastically different solutions, when compared, for example, to the deterministic solution. Fig. 2.4A show that the equilibrium value of the populations can be very different from the deterministic carrying capacity. Thus the observed equilibrium value of the populations is no longer determined by the initial conditions, as in the deterministic case (de Vladar, 2006). Actually, the carrying capacity is now a random variable. For example, in Fig. 2.4A the size equation is solved for several realizations of the process (2.2). For the naive eye, the distinct realizations could be seen as distinct "noisy logistics" with different carrying capacities. Comparing the realizations to a logistic equation gives highly significant fits, even when the data come from a common process having the same values of the parameters.

In order to determine if we can recover deterministic information of the processes, we performed simulations of 250 randomly selected values of ρ, θ and for each we performed 30 realizations. To every growth curve we least-squares-fitted a logistic model, and calculated its parameters $\hat{\rho}$ and $\hat{x}_\infty$. Fig. 2.4B shows a scatter plot of the estimated vs. the deterministic values of the Malthusian parameter, showing a poor relationship. These results show that the reconstructions are totally spuri-

ous since they do not reflect any information of the generating process. But because the dynamics resemble a logistic realization, accepting a null hypothesis that the biological phenomena determining growth are of logistic nature is true, statistically speaking. However, our calculations show that stochastic processes can account for the same qualitative and quantitative description. Therefore, simple analysis like least squares fits are not enough to confirm the logistic hypothesis.

Figure 2.4: (Opposite page) (A) Integrations for population size for distinct realizations of the same process when $\theta_0^2/8 < \epsilon_0^2(\rho_0 + 1)$ using the parameters $\rho = 1,\ \epsilon_\rho = 0.5,\ \theta = 1, \epsilon_\theta = 0.5$. The bold line shows the deterministic dynamics and the thin lines are realization for the population size. The dotted lines are logistic estimations. In this cases, the estimated Malthusian parameters range between $\hat{\rho} \in (0.8, 1.5)$. Carrying capacities range from $\hat{x}_\infty \in (1, 10^3)$, and the deterministic carrying capacity is $x_\infty = 10.0$. All of the estimations have a regression coefficient $R^2 > 0.995$ with p-values less than 10^{-3}. (B) Correlation between the estimated and generating Malthusian parameters, from 7500 simulation spanningg 250 distinct pairs of uniformly distributed values of $\rho \in [0, 2],\ \theta \in \pm\epsilon_0\sqrt{2(\rho_0 + 1)}$, using $\epsilon_\rho = 1.0,\ \epsilon_\theta = 0.1$. The radii of the circles are $10^{-3}\log(SE)$ (*SE* = standard error). The continuous line is the linear trend, which gives $\hat{\rho} = 0.389591 + 0.371044\rho$ with $R^2 = 0.2873$ (p$< 2.2\ 10^{-16}$). The dotted line is a linear trend weighted with the inverse of the standard error of each estimation: $\hat{\rho} = 0.91300 + 0.22528\rho$ with $R^2 = 0.3393$ (p$< 2.2\ 10^{-16}$). Comparing these two estimations we see that even in the best case (the weighted regression) the predictive power is poor. (C) Integration for population sizes for distinct realizations of the process when $\theta_0^2/8 \geq \epsilon_0^2(\rho_0 + 1)$ using the parameters $\rho = -3,\ \epsilon_\rho = 0.3,\ \theta = 2, \epsilon_\theta = 1.5$). The bold line represents the deterministic dynamics, and the thin lines the realizations for population size. The dotted lines are estimations for the exponential growth. The estimated values for the exponential growth parameter are in the range of $(21.8, 26.0)$. All the estimations have a regression coefficient $R^2 > 0.998$ with p-values less than 10^{-16}. The graph is in semi-log scale.


(A)
Population Size (x)
time (t)
Deterministic
(B)
$\hat{\rho}$
ρ
(C)
Log Population Size (x)
Deterministic
time (t)

There is an analogous effect for the case when the rates are persistent. Once the stochastic rates are in stationarity, the resulting population dynamics resembles exponential growth. In the deterministic exponential growth, the initial condition of the rate determines the growth parameter de Vladar (2006). However, under our scheme, the growth process is Markovian, thus the initial conditions of (x, r) do not affect the stationary distribution. As a consequence, the expected or averaged rates are spurious estimators of an exponential dynamic (Fig. 2.4C). A similar problem was described by Renshaw (Renshaw, 1991) when demographic stochasticity is present in an exponentially growing population.

At this point it is necessary to make a distinction between the outcomes of noise sources coming from demographic or phenotypic stochasticity. The first has been studied and experimentally supported (Renshaw, 1991; Lande et al., 2003). This kind of stochasticity is such that randomness affects the population through events of accidental mortality or occasional migrations (and is analogous to energy input coming from a heat bath.) In these cases, the populations would fluctuate, for example, close to carrying capacities, and thus information for the deterministic part of the dynamics can be extracted by averaging. The second type, i.e. parameter stochasticity, is more related to fluctuations in phenotypes, which results from the "superposition" of genetic and environmental processes. However, from the perspective of our model, where carrying capacities are not an intrinsic property of the environment, this averaging might not make biological sense. As we said, an average of the stochastic trajectory does not recover the deterministic path, like in Figs. 2.4A, 2.4A. Of course, populations might still be subject to demographic stochasticity, and therefore show fluctuations around a stable size. In this case, we would be presented with an additional noise source η_M, more

related to the measuring techniques, perturbing the size equation as: $\dot{x} = xr + \eta_M$, that gives the fluctuating pattern over the stable size. (This is a problem known as *filtering*: when the measuring procedure has additional noise sources, not taking them into account in the estimations, may bias the interpretation of the underlying process (Oksendal, 2002).) Considering this source of fluctuations is more related to time series estimation than to the biological aspects of our model (Siefert et al., 2003).

2.5 CONCLUDING REMARKS

To summarise, we have presented an analysis of a novel population growth model that is based in fluctuations in the per-capita growth rate, rather than in the growth variable. The result, is that the rate always remains finite, either because rate explosions are suppressed, or because rate is damped to zero stochastically. As a consequence, and depending on the relationship between the deterministic parameters and noise, the model reproduce patterns that resemble exponential and logistic (sigmoid) growth. It is important to notice that these behaviours are irrespective on how the deterministic population would grow. These forms are determined by the fluctuations and not from the biological processes of birth and death, at least not in the conventional interpretationn and description. When $\theta_0^2/8 \geq \epsilon_0^2(\rho_0+1)$ then the resulting population grows anomalously, but with bounded fluctuations, and resembles an exponential growth. When $\theta_0^2/8 < \epsilon_0^2(\rho_0+1)$ then the populations grow toward saturation. However, this result challenges the idea of a carrying capacity, that is supposed to describe self regulatory processes and an intrinsic property of the environment. Here, it is an emerging property from the fluctuations. In both

cases, and more critically in the second (the logistic), statistical fits to the realizations are highly significant. But since the effects of randomness override the deterministic forces of the system, making these statistical estimations becomes unreliable. In the context of our formulation, the question about rate estimations looses its sense, because forecasting using the classic deterministic models proves useless. Thus fluctuation analysis might prove more informative about the stochastic driving forces. In this way, estimations and forecasting can give other statistical solutions to classical and new problems, using our different perspective, that is, when populations are subject to phenotypic stochastic variability.

Part II

Population Genetics

Published as: N.H. Barton and H.P. de Vladar– *Statistical Mechanics and the Evolution of Polygenic Traits*. Genetics. vol. 181, no. 3, pp. 997–1011, 2009.

Chapter 3

Statistical Mechanics and the Evolution of Polygenic Quantitative Traits

It was a coincidence. The two fields were entirely unconnected, except at one point: Maxwel's Demon.

Thomas Pynchon

Abstract

The evolution of quantitative characters depends on the frequencies of the alleles involved, yet these frequencies cannot usually be measured. Previous groups have proposed an approximation to the dynamics of quantitative traits, based on an analogy with statistical mechanics. We present a modified version of that approach, which makes the analogy more precise, and which applies quite generally to describe the evolution of allele frequencies. We calculate explicitly how the macroscopic quantities (e.g., trait mean and genetic variability) depend on evolutionary forces, in a way that is independent of the microscopic details. We first show that the stationary distribution of allele frequencies under drift, selection, and mutation maximizes a certain measure of entropy, subject to constraints on the expectation of observable quantities. We then approximate the dynamical changes in these expectations, assuming that the distribution of allele frequencies always maximizes entropy, conditional on the expected values. When applied to directional selection on an additive trait, this gives a very good approximation to the evolution of the trait mean and the genetic variance, when the number of mutations per generation is sufficiently high $4N\mu > 1$. We show how the method can be modified for small mutation rates $4N\mu \to 0$. We outline how this method describes epistatic interactions as, for example, with stabilizing selection.

3.1 INTRODUCTION

Predicting the evolution of quantitative characters from first principles poses a formidable challenge. When multiple loci contribute to a quantitative character z, the effects of selection, mutation and drift are difficult to predict from the observed values of the trait; this is true even in the simplest case of additive effects. The fundamental problem is that the distribution of the trait depends on the 'microscopic details' of the system, namely the frequencies of the genotypes contributing the trait. In an asexual population, long-term evolution depends on the fittest genotypes, which may currently be very rare. In sexual populations - the focus of this paper - new phenotypes are generated by recombination in a way that depends on their genetic basis. If selection is not too strong, we can assume Hardy Weinberg proportions and linkage equilibrium (HWLE): this is a substantial simplification, which we make throughout. Even then, however, we must still know all the allele frequencies, and the effects of all the alleles on the trait, in order to predict the evolution of a polygenic trait. In this paper, we seek to predict the evolution of quantitative traits without following all the hidden variables (i.e., the allele frequencies) that determine the course of evolution.

For this purpose, several simplifications have been proposed. The central equation in quantitative genetics is that the rate of change of the trait mean equals the product of the selection gradient and the additive genetic variance (Lande, 1976). This simple prediction can be surprisingly accurate, since the genetic variance often remains roughly constant for some tens of generations (Falconer and Mackay, 1996; Barton and Keightley, 2002). However, we have no general understanding of how the genetic variance evolves, or indeed, what processes are responsible for maintaining it (Falconer and Mackay, 1996; Bürger

et al., 1989; Barton and Keightley, 2002). Even if we take the simplest view, that variation is maintained by the opposition between mutation and selection, the long-term dynamics of the genetic variance still depend on the detailed distribution of effects of mutations.

Lande (1976), following (Kimura, 1965b), approximated the distribution of allelic effects at each locus as a Gaussian distribution. However, this is only accurate when many alleles are available at each locus, and when mutation rates are extremely high (Turelli, 1984). Barton and Turelli (1987) assumed that loci are close to fixation, but again, this approximation has limited application: in particular, it cannot apply when one allele substitutes for another. Some progress has been made by describing a polygenic system by the moments of the trait distribution (Barton, 1986; Barton and Turelli, 1987; Turelli and Barton, 1990; Barton and Turelli, 1991). For additive traits, a closely related description in terms of the cumulants is a more natural way to represent the effects of selection (Bürger, 1991, 1993; Turelli and Barton, 1994; Rattray and Shapiro, 2001). These transformations are exact, and quite general: they provide a natural description of selection and recombination, and extend to include the dynamics of linkage disequilibria as well as allele frequencies. A moment-based description provides a general framework for exact analysis of models with a small number of genes, and for approximating the effects of indirect selection (Barton, 1986; Barton and Turelli, 1987; Lenormand and Otto, 2000; Kirkpatrick et al., 2002; Roze and Barton, 2006); for some problems, simply truncating the higher moments or cumulants can give a good approximation (Turelli and Barton, 1990; Rouzine et al., 2007). However, results are sensitive to the choice of approximation for higher moments, and so the approximation is to this extent arbitrary. The equations must be truncated "*by hand*", guided by mathematical

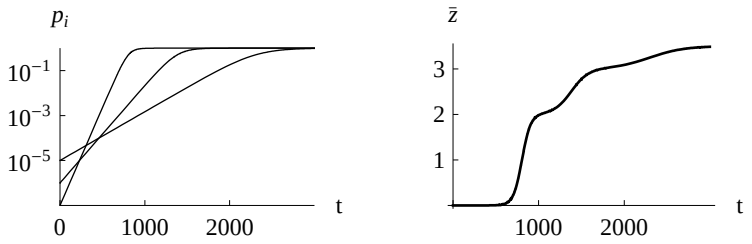


Figure 3.1: The frequency of favorable alleles at three loci (left) and the mean of an additive trait, $\bar{z}$ (right). Initial frequencies are $10^{-5}, 10^{-6}, 10^{-7}$, and effects on the trait are 1, 2, $\frac{1}{2}$, so that $\bar{z} = p_1 + 2p_2 + \frac{1}{2}p_3$; the selection gradient is β=0.01 (i.e. fitness is $e^{\beta\bar{z}}$).

tractability rather than biological accuracy.

The fundamental problem is to find a way to approximate the 'hidden variables' (in this context, the allele frequencies), but we cannot hope to do this in complete generality. Even with simple directional selection on a trait, the pattern of allele frequencies depends on the frequencies of favorable alleles that may be extremely rare ($p \ll 1$), and which will take $\sim \frac{1}{s}\log\left(\frac{1}{p}\right)$ generations to reach appreciable frequency. Thus, undetectably rare alleles can shape future evolution, without much affecting the current state. Fig. 3.1 shows an example where the trait mean changes as three alleles sweep to fixation at different times and rates. By choosing initial frequencies and allelic effects appropriately, we could produce arbitrary patterns of trait evolution. We can only hope to make progress in situations where the underlying allele frequencies can be averaged over some known distribution, rather than taking arbitrary values.

We know that particular frequencies of rare alleles can ultimately lead to peculiar outcomes (Fig. 3.1). However provided that selection does not act directly on individual alleles we ex-

pect that random drift will cause the distribution of allele frequencies to spread out across the full space of possibilities.

Despite this fundamental difficulty, progress can be made in two ways. First, we can include random drift, and follow the distribution of allele frequencies, rather than the deterministic evolution of a single population. Then, we can hope that the distribution of allele frequencies, conditional on the observed trait values, will explore the space of possible states in a predictable way. Second, we can allow selection to act only on the observed traits, and assume that the distribution of allele frequencies spreads out to follow the stationary distribution generated by such selection. That makes it much harder (and perhaps, impossible) for populations to evolve into an arbitrary state with unpredictable and idiosyncratic properties. There is an analogy here with classical thermodynamics, in which molecules might start in a special state, such that after some time they concentrate in a surprising way: all the gas might rush to one corner, for example. However, if all states with the same energy are equally likely, this is extremely improbable. In this paper, we use procedures analogous to statistical thermodynamics, but adapted to population genetics. First, we use an information entropy measure, S_H, which is derived from population genetic considerations, and which ensures an exact solution at statistical equilibrium. This measure, which is proportional to the quantity H, defined by Boltzmann (1872), was proposed by Iwasa (1988), and independently by Sella and Hirsh (2005), see Barton and Coe (2009). Second, we choose to follow a set of observable quantities that include all those acted on directly by mutation and selection. This reveals a natural correspondence between the observables, and the evolutionary forces that act on them, which is analogous to extensive and intensive variables in thermodynamics (Barton and Coe, 2009). These two innovations allow us to set out the method in a very general

way.

Throughout, we will make the usual approximations of population genetics, that populations are at Hardy-Weinberg and linkage equilibrium (HWLE), and that drift, mutation, and selection are weak. Linkage equilibrium is justified if selection and drift are not only weak (s, $1/2\ N \ll 1$) but are also weak relative to recombination (s, $1/2\ N \ll\ r$). We also assume only two alleles at each locus. These assumptions allow us to describe populations solely in terms of allele frequencies at n loci ($\vec{p} = p_1, \ldots, p_n$), and to use the continuous-time diffusion approximation.

We begin by analyzing the stationary distribution, showing the analogy with thermodynamics.Iwasa (1988) showed that the free fitness, which is the sum of the log mean fitness and the information entropy S_H ($\log\left(\bar{W}\right) + \frac{1}{2N} S_H$), always increases through time, and reaches a maximum at the classic stationary distribution of allele frequencies under mutation, selection and drift:

$$\psi\left(\vec{p}\right) = \frac{1}{\mathbb{Z}} \bar{W}^{2N} \prod_{i=1}^{n} \left(p_i q_i\right)^{4N\mu - 1} \quad , \tag{3.1}$$

where $p_i (i = 1, \ldots, n)$ are the allele frequencies at n loci, $q_i = 1 - p_i$, N is the number of diploid individuals, $\bar{W}$ is the mean fitness, and μ is the mutation rate (Wright, 1937a). The normalizing constant $\mathbb{Z}$ plays a key role; it is analogous to the partition function in statistical mechanics, and acts as a generating function for the quantities of interest, in the sense that its derivatives give the expectations of the macroscopic quantities (Barton, 1989). We then show how the rates of change of expectations of observable quantities can be approximated by averaging over this stationary distribution ψ. The crucial assumption here is that the distribution of allele frequencies always has the form of Eq. 3.1. This is accurate provided that the system evolves as a result of slow changes in the parameters,

so that it has time to approach the stationary state. By analogy with thermodynamics, such changes are termed reversible (Ao, 2008; Barton and Coe, 2009).

After setting out the method in a general way, we apply it to directional selection on an additive trait. In this simple case, we can give closed-form expressions for $\mathbb{Z}$, and hence for observables such as the expectations of the trait mean, genotypic variance, genetic variability, etc. We then show that our approximation to the allele frequency distribution gives a good approximation to the dynamical change in the trait distribution, even when selection changes abruptly. However, the method only works for high mutation rates ($4N\mu > 1$), and breaks down when $4N\mu < 1$. Nevertheless, we show how the method can be adapted to the case where $4N\mu$ is small.

3.2 GENERAL ANALYSIS

Defining entropy: The key concept is of an entropy, S_H, which measures the deviation of the population from a base distribution ϕ - in this case, the density under drift alone. Entropy always increases as the population converges towards ϕ under drift. With selection and mutation, a *free energy*- the sum of the entropy and a potential function - always increases (Iwasa, 1988). We show that the stationary distribution maximizes S_H subject to constraints of the expected value of a set of observable quantities. Thus, the dynamics of these quantities can be approximated by assuming that the entropy is always maximized, conditioned on their values.

There is a wide range of definitions, interpretations and generalizations of entropy (e.g. Renyi, 1961; Wehrl, 1978; Tsallis, 1988); these have been applied to biological systems in various ways. Iwasa (1988) introduced the concept of entropy into

population genetics, for a diallelic system of one locus under reversible mutation and with arbitrary selection; he also considered a phenotypic model of quantitative trait evolution. Iwasa (1988) used an information entropy, also known as a relative entropy (Gzyl, 1995, Ch.3, Georgii, 2003), and defined as:

$$S_H[\psi] \equiv -\int \psi \log\left[\frac{\psi}{\phi}\right] d\vec{p}\,. \tag{3.2}$$

This is a functional of the probability distribution of allele frequencies, ψ, that evaluates the average entropy of a function with respect to a given base distribution, ϕ, integrated over all possible allele frequencies, denoted by $d\vec{p} = dp_1.dp_2 \ldots dp_n$. It can be thought of as (minus) the expected log-likelihood of ϕ, given samples values drawn from a distribution ψ, relative to the base distribution ϕ; it has a maximum at ψ=ϕ, when $S_H[\phi] = 0$. We denote it by a subscript H because it is essentially the same as the measure introduced by Boltzmann (1872) in his H-theorem.

The variation of S_H with respect to small changes in ψ is:

$$\delta \mathrm{S}_H[\psi] = -\int \left(\lambda + \log\left[\frac{\psi}{\phi}\right]\right) \delta\psi d\vec{p}\,, \tag{3.3}$$

where λ is a Lagrange multiplier associated with the normalization condition $\int \psi\, d\vec{p} = 1$ (Barton and Coe, 2009). Note that because ψ is normalized, $\int \delta\psi d\vec{p}$ = 0. With no constraints other than this normalization, setting $\delta \mathrm{S}_H = 0$ implies that the entropy is at an extreme only if $\psi = \phi$; this is a unique maximum.

We are interested in a set of observable quantities, A_j, which are functions of the allele frequencies in a population. These might, for example, describe the distribution of a quantitative trait - for example, its mean and variance. We need to find the distribution of allele frequencies, ψ, that maximizes the entropy, $S_H[\psi]$, given constraints on the expected values of these

observables, $\langle A_j \rangle$:

$$\langle A_j \rangle = \int A_j \psi d\vec{p} \, . \tag{3.4}$$

With these constraints, the extremum of Eq. is calculated by including the Lagrange multipliers associated with the A_j's, defined for convenience as $-2N\alpha_j$:

$$\int \left(\lambda + \log \left[\frac{\psi}{\phi} \right] - \sum_j 2N\alpha_j A_j \right) \delta\psi d\vec{p} = 0 \, . \tag{3.5}$$

At the extremum, the term in parentheses should be zero. This implies that the distribution that maximizes entropy, subject to constraints, is the Boltzmann distribution:

$$\psi_{\text{ME}} = \frac{\phi}{\mathbb{Z}} \text{Exp} \left[\sum_j 2N\alpha_j A_j \right] = \frac{\phi}{\mathbb{Z}} \text{Exp} \left[2N\vec{\alpha}.\vec{A} \right] \, , \tag{3.6}$$

where we have expressed the Lagrange multiplier λ as $\mathbb{Z}$ = Exp[λ], and choose $\mathbb{Z}$ to normalize the distribution. We show later that $\mathbb{Z}$ is a generating function for the moments of the observables, A_j. It will play a major role in our calculations:

$$\mathbb{Z} = \int \phi \exp \left[2N \left(\vec{\alpha}.\vec{A} \right) \right] d\vec{p} \, . \tag{3.7}$$

We will show that under directional selection and mutation the α_j can be identified with the set of selection coefficients and mutation rates, and the A_j with the quantities on which selection and mutation act (e.g. trait mean and genetic variability). The potential function $\sum_j A_j \alpha_j = \vec{\alpha}.\vec{A}$ consists of the log-mean fitness, $\log \left(\bar{W} \right)$, plus a term representing the effect of mutation, $2\mu \sum_j \log \left(p_j q_j \right)$. Then, Eq. 3.6 gives the classical stationary density of Eq. 3.1. (Note that although $\vec{\alpha}.\vec{A}$ must equal the potential function, which includes all evolutionary processes, apart from

drift, we still have some freedom to separate this into components in a variety of ways. For example, directional selection on a set of traits could be represented by almost any linear basis. Nevertheless, there will usually be a natural set of components that represent different evolutionary processes. In addition, we are free to include additional observables, that are not necessarily under selection, and so have $\alpha_i = 0$. These extra degrees of freedom will improve the accuracy of our dynamical approximations.)

Notice that there is an alternative measure of entropy, S_Ω, defined by the log-density of states that are consistent with macroscopic variables $\langle \vec{A} \rangle$ (Barton and Turelli, 1989). Barton and Coe (2009) discuss the relation between S_Ω and S_H, and show that these two measures converge when the distribution clusters close to its expectation.

The generating function, $\mathbb{Z}$: The normalizing constant $\mathbb{Z}$, which is a function of $\vec{\alpha}$, acts as a generating function for quantities of interest. Differentiating w.r.t. $2N\vec{\alpha}$ we find that:

$$\frac{\partial \log(\mathbb{Z})}{\partial \left(2N\alpha_j\right)} = \langle A_j \rangle \ . \tag{3.8}$$

Differentiating w.r.t. population size:

$$\frac{\partial \log(\mathbb{Z})}{\partial (2N)} = \left\langle \vec{\alpha}.\vec{A} \right\rangle \ . \tag{3.9}$$

Differentiating again w.r.t. the $\vec{\alpha}$ gives the covariance between fluctuations in the $\vec{A}$:

$$\frac{\partial^2 \log(\mathbb{Z})}{\partial \left(2N\alpha_j\right) \partial \left(2N\alpha_k\right)} = \mathrm{Cov}\left(A_j, A_k\right) \equiv C_{j,k} \ . \tag{3.10}$$

This covariance matrix, which we denote C, will play an important role in the dynamical approximation (Le Bellac et al., 2004, p. 64).

Analyzing the dynamics: As the system moves away from stationarity, it will not in general follow precisely the distribution that maximizes entropy. (This can be seen by substituting the maximum entropy from Eq. 3.6 with time-varying parameters $\vec{\alpha}_{(t)}$ as a trial solution to the diffusion equation). However, the distribution of microscopic variables may nevertheless stay close to a maximum entropy distribution (Nicolis and Prigogine, 1977; De Groot and Mazur, 1984; Goldstein and Lebowitz, 2004). Our key assumption is that the macroscopic variables change slowly enough that the system is always close to a local equilibrium.

We will show that the Lagrange multipliers, $\vec{\alpha}$, correspond to forces that act on the observables, $\vec{A}$: directional selection acts on the trait mean, mutation on the diversity U, and so on. Crucially, we assume that changes occur solely through changes in the parameters $\vec{\alpha}$; arbitrary perturbations that act directly on the allele frequencies could have arbitrary effects (as, for example, in Fig. 3.1)

Assume that changes in allele frequency are determined by a potential function $\vec{\alpha}.\vec{A}$, which can be written as a sum of components $\alpha_k A_k$. (In physics, energy acts as a potential; in population genetics, mean fitness plays an analogous role; it defines an adaptive landscape such that allele frequencies and their means change at rates proportional to the fitness gradient Wright, 1967; Lande, 1976). Our method only works for systems whose dynamics can be described by a potential in this way (Ao, 2008, though see). In an infinitesimal time δt, the mean and mean square changes are:

$$\langle \delta \mathrm{p}_i \rangle = \frac{p_i q_i}{2} \frac{\partial \left(\vec{\alpha}.\vec{A} \right)}{\partial p_i} \ , \tag{3.11a}$$

$$\left\langle \delta \mathrm{p}_i \delta \mathrm{p}_j \right\rangle = 0 \text{ for } i \neq j \ , \tag{3.11b}$$

$$\left\langle \delta \mathrm{p}_i^2 \right\rangle = \frac{p_i q_i}{2N} \ . \tag{3.11c}$$

The first equation, for $\langle \delta \mathrm{p} \rangle$, is just Wright's (1967) formula for selection, modified to include mutation. The variance of allele frequency fluctuations, $\left\langle \delta \mathrm{p}^2 \right\rangle$, is the standard formula for random drift. Under the diffusion approximation, this leads to the stationary distribution of Eq. 3.1, provided that the base distribution is defined as:

$$\phi = \left(\prod_{i=1}^{n} p_i q_i \right)^{-1} . \tag{3.12}$$

Under the diffusion approximation, the rate of change of $\langle A_j \rangle$ is:

$$\begin{aligned} \frac{\partial \langle A_j \rangle}{\partial t} &= \sum_{i=1}^{n} \frac{\partial A_j}{\partial p_i} \langle \delta \mathrm{p}_i \rangle + \ \frac{1}{2} \sum_{i=1}^{n} \vec{n} \sum_{k=1} \frac{\partial^2 A_j}{\partial p_i \partial p_k} \langle \delta \mathrm{p}_i \delta \mathrm{p}_k \rangle \\ &= \sum_k B_{j,k} \alpha_k + \frac{1}{2N} V_j \ , \end{aligned} \tag{3.13}$$

where

$$\begin{aligned} B_{j,k} &= \left\langle \sum_{i=1}^{n} \frac{\partial A_j}{\partial p_i} \frac{p_i q_i}{2} \frac{\partial A_k}{\partial p_i} \right\rangle \ , \\ V_j &= \left\langle \sum_{i=1}^{n} \frac{p_i q_i}{2} \frac{\partial^2 A_j}{\partial p_i^2} \right\rangle \ . \end{aligned} \tag{3.14}$$

This relationship is exact, provided that the matrix B and the vector V are evaluated at the current distribution of allele frequencies. Eq. 3.13 can also be derived directly, by making a diffusion approximation to multivariate observables, where the deterministic terms are $a_i = \sum_k \frac{\partial A_j}{\partial p_i} \frac{p_i q_i}{2} \alpha_k$, and the diffusion terms are $b_i = \sqrt{\frac{p_i q_i}{2N}}$ (Ewens, 1979; Gardiner, 2004). If our system is described by only one observable, we directly recover the formula derived by Ewens (1979, pp. 136-137).

The local equilibrium approximation: In general, as the system moves away from stationarity, it will not precisely follow the distribution that maximizes entropy. (This can be seen by substituting the maximum entropy form, Eq. 3.2, with time-varying parameters $\vec{\alpha}_{(t)}$ as a trial solution to the diffusion equation). However, the distribution of microscopic variables may nevertheless stay close to a maximum entropy distribution if the macroscopic variables change slowly enough such that the system remains close to a local equilibrium at all times (Prigogine, 1949; Klein and Prigogine, 1953; Nicolis and Prigogine, 1977; De Groot and Mazur, 1984; Goldstein and Lebowitz, 2004).

We now approximate $B_{j,k}$andV_j by $B^*_{j,k}$, V^*_j, assuming the distribution in Eq. 3.6 evaluated at $\vec{\alpha}^*$. We know that at the stationary state, under parameters $\vec{\alpha}^*$, expectations are constant, and so from Eq. 3.13, $\sum_k B^*_{j,k}\alpha^*_k + \frac{1}{2N}V^*_j = 0$. Therefore:

$$\frac{\partial \langle A_j \rangle}{\partial t} \approx \sum_k B^*_{j,k} \left(\alpha_k - \alpha^*_k\right) \tag{3.15}$$

The matrix $B_{j,k}$ is closely related to the additive genetic covariance matrix. Making the link with quantitative genetics is not quite straightforward, because the A_j are arbitrary functions of the allele frequencies, and need not be the means of actual traits carried by individuals. Nevertheless, if we do regard them as the means of some quantity, then $\partial A_j/\partial p_i$ is twice the average effect of alleles at locus i. (Since we assume HWLE, average effect is equal to average excess; Falconer and Mackay, 1996). Therefore, $B_{j,k}$ is the expected additive genetic covariance between $A_j$andA_k, the expectation being taken over the distribution of allele frequencies. Moreover, if the α_k contribute to the log-mean fitness (rather than to the component of the potential that describes mutation), then they can be interpreted as selection gradients in the usual way. Equation 12 thus gives the rates of change of the expected trait means as the product

of the expected additive genetic covariance, and the difference between the actual selection gradient, α_k,and the gradient that would give stationarity at the current expectations, α_k^*. This interpretation will become clearer when we consider specific examples, below.

In thermodynamics, equations similar to Eq. 3.15 are called phenomenological equations (van Kampen, 1957; De Groot and Mazur, 1984, Ch. IV). They were postulated as approximations to processes that are close to equilibrium. In such cases, the variables α^* represent the deviation from an equilibrium defined by α. These equations are valid as long as a local equilibrium exists, and (as suggested by Eq. 3.13) it holds in general that $B_{k,j} = B_{j,k}$ (Onsager, 1931; Prigogine, 1949). For theoretical purposes, we can follow either the expectations $\langle A_j \rangle$themselves, or the parameters α_k^* that would give those expectations at stationarity. In numerical calculations, the latter is more convenient, because that avoids calculating the α_j^* from the $\langle A_j \rangle$ (a tricky inverse problem). The rates of change of the α_k^* are related to the rates of change of the $\langle A_j \rangle$ via the matrix $\partial \langle A_j \rangle / \partial \alpha_k^*$. Now, since $\langle A_j \rangle = \partial \log(\mathbb{Z}) / \partial (2N\alpha_k^*)$, we have

$$\partial \langle A_j \rangle / \partial \alpha_k^* = 2N \left[\partial^2 \log(\mathbb{Z}) / \partial \left(2N\alpha_j^*\right) \partial (2N\alpha_k^*) \right] = 2NC_{jk}$$

thus, the relation between the $\langle A_j \rangle$ and the α_k^* is via the covariance of fluctuations, $C_{j,k}$ (Eq. 3.10). In matrix notation (equivalent to De Groot and Mazur, 1984, p. 36):

$$\frac{\partial \vec{\alpha}^*}{\partial t} \approx \frac{1}{2N} C^{-1}.B.\left(\vec{\alpha} - \vec{\alpha}^*\right) \ , \tag{3.16}$$

where C is the matrix of covariances of fluctuations in the $\vec{A}$, and B is analogous to the additive genetic covariance matrix. Both these are evaluated at the stationary distribution defined by $\vec{\alpha}^*$. For given $\vec{\alpha}^*$, we find the $\left\langle \vec{A} \right\rangle$ by integrating using the density in Eq. 3.6, or application of Eq. 3.8.

3.3 DIRECTIONAL SELECTION, MUTATION, AND DRIFT

3.3.1 Analysis

The stationary distribution: We now apply this method to a quantitative trait under directional selection, mutation, and drift. We first define a measure of genetic variability (Barton and Coe, 2009):

$$U = 2\sum_{i=1}^{n} \log\left(p_i q_i\right) \ , \tag{3.17}$$

which is $2n$ times the log-geometric mean heterozygosity across loci (plus a constant); n is the number of loci. The rate of change of p_i dues to symmetric mutation is $\mu\left(q_i - p_i\right) = \frac{p_i q_i}{2}\frac{\partial(\mu U)}{\partial p_i}$, as required by Eq. 3.11a. Under our assumption of linkage equilibrium, the rate of change of p_i due to selection is $\frac{p_i q_i}{2}\frac{\partial \log\left(\bar{W}\right)}{\partial p_i}$ (Eq. 3.11a). The log mean fitness, $\log\left(\bar{W}\right)$, is a natural potential for the system, and will be expressed as a sum of components $\vec{A}.\vec{\alpha}$, where the $\vec{\alpha}$ are a set of selection coefficients. We deal with the very simplest case of exponential (directional) selection, but note that the derivation applies to any form of selection for which a potential function can be defined - most obviously, the case where genotypes have fixed fitnesses. If individuals with trait value z have fitness $e^{\beta z}$, then to leading order in β, the mean fitness is $\bar{W} = e^{\beta \bar{z}}$. Wright's equilibrium density can then be written in the form of Eq. 3.6, with $\vec{A} = \{\bar{z}, U\}$ and $\vec{\alpha} = \{\beta, \mu\}$:

$$\psi = \frac{\phi}{\mathbb{Z}} e^{2N(\beta \bar{z} + \mu U)} \ . \tag{3.18}$$

Thus, the stationary distribution under mutation, selection and random drift is given by maximizing the entropy subject to constraints on the expected genetic diversity $\langle U \rangle$,and the expected

trait mean, $\langle \bar{z} \rangle$. The entropy is defined by Eq. 3.2, with baseline distribution $\phi = \prod_{i=1}^{n} (p_i q_i)^{-1}$. Then, Eq. 3.18 is the stationary distribution, and is equal to Eq. 3.1.

We have shown that a population evolving under mutation, multiplicative selection and drift will converge to a stationary distribution that has maximum entropy, S_H, given the expected trait mean and genetic diversity. As we will see below, other forms of selection can be represented by introducing other observables. Each constrained observable will be conjugated with a natural variable: in this example, the expected mean $\langle \bar{z} \rangle$ corresponds to the strength of directional selection β, and the expected diversity $\langle U \rangle$ to the mutation rate μ. Information about the full distribution of the observables is contained in the normalizing constant $\mathbb{Z}$, which is a generating function that depends only on the natural variables α_j. In the next section, we calculate an explicit expression for it.

The generating function for an additive trait: We have not yet made any assumptions about the genetic basis of the trait, z; in general, there might be arbitrary dominance and epistasis. We now assume that it is additive, with locus i having effect γ_i:

$$z = \sum_{i=1}^{n} \gamma_i \left(X_i + X_i^* - 1 \right) \ , \tag{3.19}$$

where X_i and X_i^* represent the allelic states (labelled 0 or 1) of the two copies of each of the n loci. With additivity, exponential selection on the trait corresponds to multiplicative selection on the underlying loci. If we average over the population, where p_i represents the frequency of the $X_i = 1$ allele, and $q_i = 1 - p_i$, then the mean and genetic variance are:

$$\bar{z} = \sum_{i=1}^{n} \gamma_i \left(p_i - q_i \right) \ , \ v_z = \vec{n} 2 \sum_{i=1} \gamma_i^2 p_i q_i. \tag{3.20}$$

More than two alleles could be allowed, but only for special mutation rates that give detailed balance.

The normalization function $\mathbb{Z}$ can now be calculated explicitly, using Eq. 3.7. In this simple case of directional selection on an additive trait, the integrand separates out as a product over allele frequencies, and so:

$$\begin{aligned}\mathbb{Z} &= \int \exp\left[2N\left(\beta\bar{z}+\mu U\right)\right]\left(\prod_{i=1}^{n} p_i q_i\right)^{-1} d\vec{p} \\ &= \prod_{i=1}^{n}\left(\int_0^1 e^{2N\beta\,\gamma_i(p-q)}(pq)^{4N\mu-1}dp\right) \\ &= \prod_{i=1}^{n}\left(\sqrt{\pi}2^{1-8N\mu}\Gamma[4\mathbf{N}\mu]\,{}_0F_1\left[\frac{1}{2}+4N\mu,(\mathbf{N}\beta\gamma_i)^2\right]\right) \\ &= \prod_{i=1}^{n}\left(\sqrt{\pi}\,(4\mathbf{N}\beta\gamma_i)^{\frac{1}{2}-4\mathbf{N}\mu}\,\Gamma[4N\mu]I_{4N\mu-\frac{1}{2}}\left(2\mathbf{N}\beta\gamma_i\right)\right)\ ,\end{aligned} \tag{3.21}$$

where $\Gamma(\cdot)$ and ${}_0F_1(\cdot,\cdot)$are the gamma and the regularized confluent hypergeometric functions, respectively. We have also given an equivalent form, in terms of the modified Bessel function of order ν, $I_\nu(\cdot)$.

Finding the expectations $\langle U\rangle, \langle\bar{z}\rangle$**:** The expectations, variances and covariances of $\bar{z}$ and U can be calculated either by direct integration, or by taking derivatives of $\log(\mathbb{Z})$ w.r.t. β and μ (Eqns. 3.8, 3.10). Explicit formulae are given in Appendix D.1 (Eqns. D.13 and D.5).

Figure 3.2 shows how the expected values change for a range of mutation rates and selection pressures, for a population of individuals with n loci of equal effect, $\gamma_i = 1$. As selection becomes strong relative to mutation, the allele $X = 1$ tends to fixation, and $\langle\bar{z}\rangle$ tends to n (top right of Fig. 3.2). As mutation becomes strong relative to drift, allele frequencies tend to $\frac{1}{2}$,

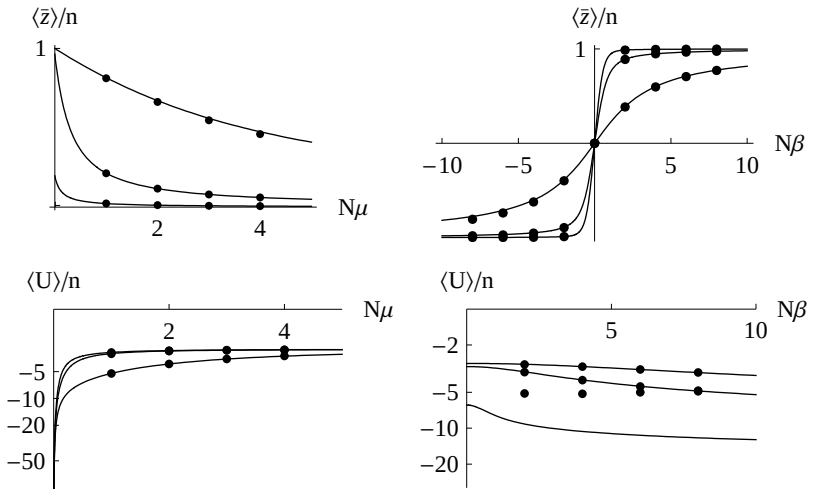


Figure 3.2: Dependence of $\langle \bar{z} \rangle, \langle U \rangle$ on $\mathrm{N}\mu, \mathrm{N}\beta$. The solid curves show the statistical mechanical approximation, while the dots show exact values for the Wright-Fisher model with $N = 100$. The plots against $\mathrm{N}\mu$ (left column) show $\mathrm{N}\beta = 0.1, 1, 10$; those against $\mathrm{N}\beta$ (right column) show $\mathrm{N}\mu = 0.1, 1$. Agreement between discrete and continuous models is close for $\langle \bar{z} \rangle$ and for $4\mathrm{N}\mu > 1$, but statistical mechanics fails to predict $\langle U \rangle$ when $4\mathrm{N}\mu = 0.1$ (lower series of dots at lower right). (For the discrete model, $\langle U \rangle$ is calculated omitting fixed classes).

and $\langle \bar{z} \rangle$ tends to zero (top left of Fig. 3.2). The expected diversity, $\langle U \rangle$, increases with mutation rate (bottom left of Fig. 3.2), and decreases slightly with the strength of selection (bottom right of Fig. 3.2). Figure 3.2 compares the statistical mechanics expectations (Eqns. D.13 and D.5 in Appendix D.1) with the Wright-Fisher model for $N = 100$. There is close agreement for $\langle \bar{z} \rangle$ for all $4N\mu$, and for $\langle U \rangle$ when $4N\mu > 1$. In the discrete model, $\langle U \rangle$ must be calculated excluding the fixed classes, since U would otherwise be infinite. This has negligible effect when $4N\mu > 1$ because fixation is unlikely. However, when $4N\mu < 1$,

there is a substantial probability of being fixed, even when fixed classes must be dropped. Thus, $\langle U \rangle$ depends on population size, and differs substantially from the diffusion approximation (compare lower series of dots with lower curve in Fig. 3.2 , bottom right). The stationary density is still close to the diffusion approximation for polymorphic classes, and so for very large N, when the probability of actually being fixed becomes small $\left(\sim \int_0^{1/2N} p^{4N\mu-1}\, dp \ll 1\right)$, $\langle U \rangle$ in the discrete Wright-Fisher model does converge to the diffusion approximation. However, for population sizes in the hundreds, there is still a very large discrepancy. We consider the implications of small $4N\mu$ for the maximum entropy method below.

For an additive trait, and equal allelic effects, the distribution of allele frequencies is the same at each locus, and so this simple case is essentially a single-locus analysis. However, this is no longer the case when we allow unequal allelic effects; more generally, if there is epistasis for fitness, the allele frequency distributions at each locus are no longer independent, and if there is epistasis for the trait, we can no longer treat macroscopic variables as sums over loci.

Covariances of fluctuations, C, and additive genetic variance, B: In order to approximate the dynamics, we need the covariances of fluctuations, C, and the additive genetic covariance, B, defined above. The matrix C, that gives the variances and covariance of U and $\bar{z}$, is calculated by taking derivatives of the generating function (Eq. 3.10; Appendix D.1, Eqns. D.15 (or D.39), D.7, and D.22).

The additive genetic covariance matrix, B, is defined in Eq. 3.13, in terms of the derivatives $\partial A_j/\partial p_i$. For the observables $\{U, \bar{z}\}$, these are $\{2\left(q_i - p_i\right)/\left(p_i q_i\right), 2\}$). Using the relation $\left(q_i - p_i\right)^2 = 1 - 4p_i q_i$:

$$\begin{aligned} B &= \left\langle \sum_{i=1}^{n} \frac{p_i q_i}{2} \begin{pmatrix} \left(\frac{\partial U}{\partial p_i}\right)^2 & \frac{\partial U}{\partial p_i}\frac{\partial \bar{z}}{\partial p_i} \\ \frac{\partial U}{\partial p_i}\frac{\partial \bar{z}}{\partial p_i} & \left(\frac{\partial \bar{z}}{\partial p_i}\right)^2 \end{pmatrix} \right\rangle \\ &= \left\langle \sum_{i=1}^{n} \begin{pmatrix} 2\left(\frac{1}{p_i q_i} - 4\right) & 2\gamma_i\left(q_i - p_i\right) \\ 2\gamma_i\left(q_i - p_i\right) & 2\gamma_i^2 p_i q_i \end{pmatrix} \right\rangle . \end{aligned} \tag{3.22}$$

Note that $B_{\bar{z},\bar{z}} = \left\langle \sum_{i=1}^{n} \frac{p_i q_i}{2}\left(\frac{\partial \bar{z}}{\partial_{p_i}}\right)^2 \right\rangle = \sum_{i=1}^{n} \left\langle 2\gamma_i^2 p_i q_i \right\rangle$ is just the expected genetic variance for the trait z, $\langle v_z \rangle$, consistent with our interpretation of B as a genetic covariance matrix. For this model, B has a simple form:

$$B = \begin{pmatrix} \frac{2(2n + 4\mathrm{N}\beta\langle\bar{z}\rangle)}{4N\mu - 1} & -2\left\langle \bar{z} \right\rangle \\ -2\left\langle \bar{z} \right\rangle & \frac{2(N\mu)\langle\bar{z}\rangle}{\mathrm{N}\beta} \end{pmatrix} . \tag{3.23}$$

Remarkably, B depends only on $\langle\bar{z}\rangle$, and not directly on the distribution of allelic effects, γ. Note that the expected genetic variance, $\langle v_z \rangle$ is equal to $2\frac{\mathrm{N}\mu}{\mathrm{N}\beta}\left\langle \bar{z} \right\rangle$, even with unequal allelic effects. This can be understood by seeing that the rates of change of $\langle\bar{z}\rangle$ due to mutation, -2$\mu\langle\bar{z}\rangle$, and due to selection, $\beta\langle v_z\rangle$, must balance at statistical equilibrium. In the limit where selection becomes weak, both $\langle\bar{z}\rangle$ and $N\beta$ tend to zero, and the expected genetic variance tends to a definite limit: all frequencies are at $\frac{1}{2}$, and so $\langle v_z \rangle$ tends to $\frac{n}{2}$.

The coefficient $B_{U,U}$ includes the expectation of $1/(p_i q_i)$, which diverges when $4N\mu < 1$. Because the rate of change depends on $B_{U,U}(\mu^* - \mu)$ (Eq. 3.23), that implies that μ^* must be held fixed at its actual value (i.e., $\mu^* = \mu$). In effect, therefore, $\langle U \rangle$ can no longer be included in the approximation. We discuss the implications of this constraint below.

3.3.2 Approximating the dynamics

Evolution of the expectations: We can now use Eq. 3.15 to approximate the rates of change of the expectations, $\langle U \rangle, \langle \bar{z} \rangle$:

$$\frac{d}{dt}\begin{pmatrix} \langle U \rangle \\ \langle \bar{z} \rangle \end{pmatrix} \approx \begin{pmatrix} \frac{2(2n+4N\beta\langle \bar{z} \rangle)}{4N\mu - 1} & -2\langle \bar{z} \rangle \\ -2\langle \bar{z} \rangle & 2\frac{N\mu}{N\beta}\langle \bar{z} \rangle \end{pmatrix} \begin{pmatrix} \mu - \mu^* \\ \beta - \beta^* \end{pmatrix} \tag{3.24}$$

These equations are proportional to the difference between the actual parameters $\{\mu, \beta\}$, and the parameters that would give a stationary distribution with the current expectations, $\{\mu^*, \beta^*\}$. To iterate these recursions, we would need to find $\{\mu^*, \beta^*\}$ from $\langle U \rangle, \langle \bar{z} \rangle$, which is troublesome. It is more straightforward to work with the rates of change of $\{\mu^*, \beta^*\}$, which are found by multiplying the rates of change of the expectations (Eq. 3.24) by the inverse of the covariance of fluctuations, C (see Eq. 3.16 and Appendix D.1, Eqns.D.36, D.27, and D.52, for $k = 1$). However, because C depends on the allelic effects in a complex way, the full dynamics do depend on the distribution of allelic effects, γ_i.

In the following sections we test the accuracy of this local equilibrium approximation against two situations: an abrupt change in β or μ, or a sinusoidal change in μ or β. An abrupt change seems the strongest test of our approximation, whilst a sinusoidal change allows us to find how the accuracy of the approximation decreases as changes become faster. For the

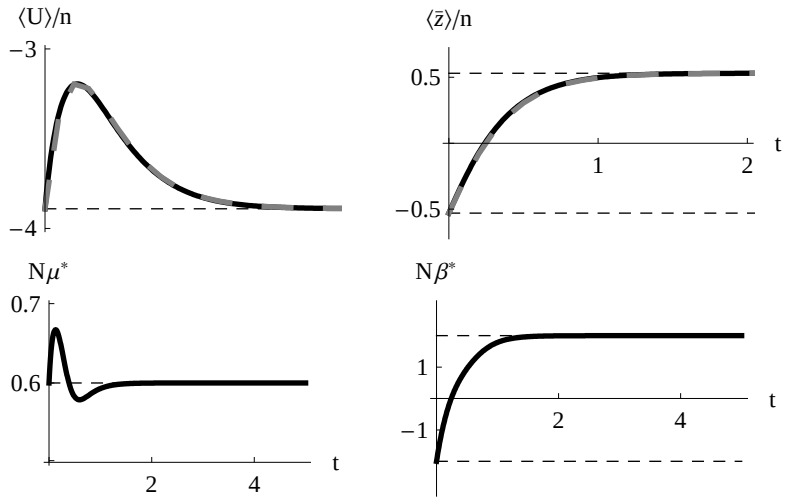


Figure 3.3: (Top panel) Calculation of genetic variability $\langle U \rangle$ (left) and trait mean $\langle \bar{z} \rangle$ (right) and over time, with $N\mu = 0.6$ as Nβchanges from -2 to +2 at time $t = 0$. The horizontal lines show the stationary values. The solid curves show the approximation, and the dashed curves, numerical solutions to the diffusion equation; these are not distinguishable on this scale. (Bottom panel) Changes over time in the parameters μ^*(left)andβ^* (right), calculated using the approximation of Eq. 3.16. Time is scaled to $2N$ generations.

moment, we focus on high numbers of mutations $(4N\mu > 1)$. We begin by considering the case of equal effects, where the distributions at all loci are the same. We also discuss results for the case where most loci have small effect, but some have large effect: the patterns are similar to the symmetric case of equal effects, and so we detail them below. Throughout, we compare the approximation with numerical solutions of the diffusion equation: these are close to solutions of the discrete Wright-Fisher model provided that $4N\mu > 1$ (Fig. 3.3).

Equal allelic effects: If all loci have equal effects on the trait,

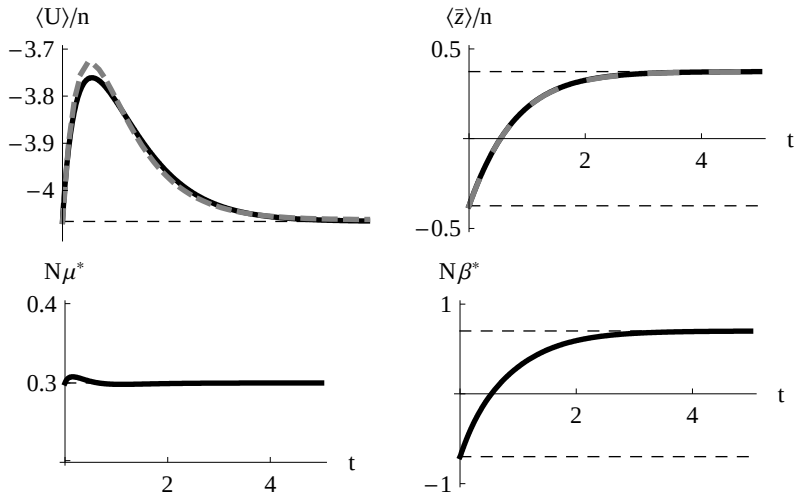


Figure 3.4: The accuracy of the approximation for N$\mu = 0.3$; Nβchangesfrom -0.7 to +0.7 at time $t = 0$. Otherwise, as for Fig. 3.3.

and if selection only acts on the trait, and not on the individual genotypes, then under directional selection the distribution of allele frequencies will be the same at each locus, and will be independent across loci. Thus, we only need follow a single distribution, whose time evolution is given either by numerical solution of the diffusion equation, or as an expansion of eigenvectors (Crow and Kimura, 1970, p. 396). However, the maximum entropy approximation is still non-trivial, even in this highly symmetric case, since it approximates the full distribution by a few degrees of freedom, such as $\{\langle 2\log(pq)\rangle, \langle p-q\rangle\}$. Also, note that with other forms of selection, the allele frequency distribution is not independent across loci: for example, with stabilizing selection populations cluster around states where the sum of allele frequencies is close to the optimum.

Abrupt change in $N\beta$: First, assume that a system is at equilibrium with evolutionary forces β_0 and μ_0. These forces are then abruptly changed to new values β and μ, and the system moves towards its new stationary state. Figure 3.3 shows that for moderately high mutation rates (N$\mu = 0.6$), and for an abrupt change of selection from $N\beta = -2$to $+2$, the approximation is extremely accurate, as compared with theresults of the diffusion equation. The expected genetic diversity, $\langle U \rangle$, increases as the allele frequencies pass through intermediate values, but returns to its original value as $\langle \bar{z} \rangle$ moves from -2 to +2 (top left). This transient increase is mainly due to the change in mean allele frequencies: there is only a small transient change in μ^*(bottom left). The distribution of allele frequencies predicted by the approximation is always close to the actual distribution (not shown).

For a lower mutation rate of $N\mu = 0.3$, close to the critical value of $1/4$, the effective mutation rate hardly changes: it is held close to the actual value of $N\mu$ = 0.3 (Fig. 3.4, lower left). The approximation is still accurate, but there is an appreciable discrepancy in $\langle U \rangle$ (upper left). For a still lower mutation rate of N$\mu = 0.1$, below the threshold where $B_{U,U}$ diverges, μ^* must necessarily be held equal to the current mutation rate (Fig. 3.5, upper left). Then, there is a poor fit to the transient increase in expected diversity, $\langle U \rangle$, but the dynamical approximation to $\langle \bar{z} \rangle$ remains accurate (Fig. 3.5, upper right). (Because μ^* must be held fixed at its actual value when $4N\mu < 1$, $\langle U \rangle$ is not now included in the approximation).

Abrupt change in $N\mu$: Figure 3.6 shows the effects of an abrupt change in mutation rate from $N\mu = 0.3$ to 1. Here, the approximation does poorly when mutation rate increases abruptly, even when 4Nμ is always larger than 1 (left of Fig. 3.6). It does perform better when the mutation rate decreases abruptly, however ($t > 5$ in Fig. 3.6).

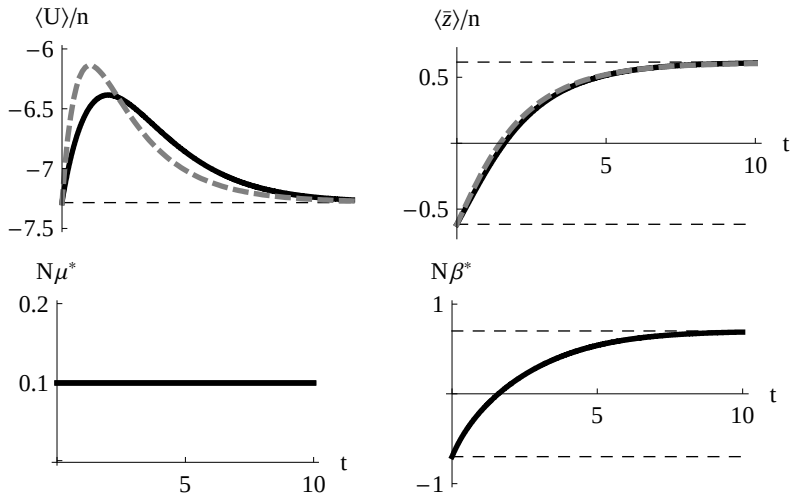


Figure 3.5: The accuracy of the approximation for N$\mu = 0.1$; Nβchangesfrom -0.7 to +0.7 at time $t = 0$. The effective mutation rate, $N\mu^*$, is held fixed at $N\mu$ (lower left). Otherwise, as for Fig. 3.3.

Unequal allelic effects: So far, we have assumed equal allelic effects. This ensures that the allele frequency distribution is the same at each locus, so that we are essentially analyzing a single-locus problem. This is not entirely trivial, since we are approximating the full allele frequency distribution by two variables, $\{\langle \bar{z} \rangle, \langle U \rangle\}$. However, we now turn to the more challenging case of unequal allelic effects at n loci: now, we are summarizing n distinct distributions by two variables. We do, however, assume that the allelic effects are known.

We draw allelic effects at ten loci from a Gamma distribution, with mean 1 and standard deviation $\frac{1}{2}$:

$$\gamma_i = \{1.69, 1.47, 1.15, 1.05, 1.04, 1.03, 1.01, 0.81, 0.500, 0.401\} \quad (3.25)$$

The maximum range of the trait is $\pm \sum_i \gamma_i = 10.15$, and the max-

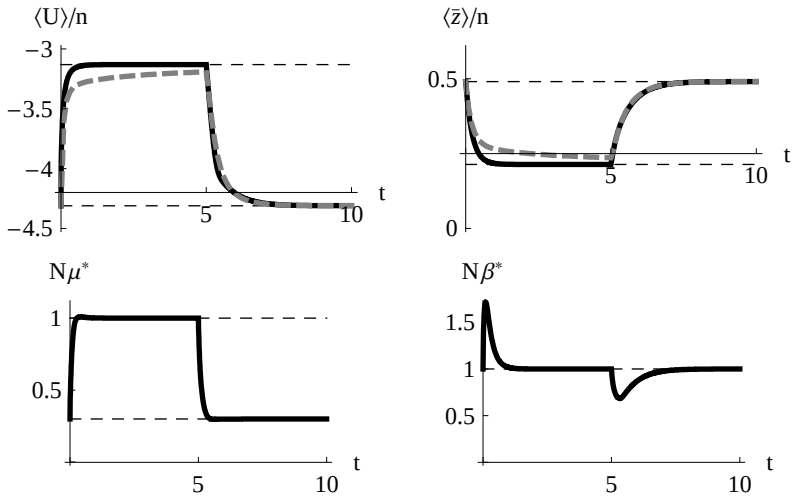


Figure 3.6: The mutation rate increases abruptly from $N\mu = 0.3$ to $N\mu = 1$ at $t = 0$, and then changes back at $t = 5$; throughout, $N\beta = 1$. The horizontal dashed lines show values at the stationary states, the dashed curves show numerical solutions of the diffusion equation, and the solid curves in the top row show the approximation.

imum genetic variance is $v_{\max} = \frac{1}{2}\sum_{i=1}^{10}\gamma_i^2 = 11.66$. 25% of this is contributed by the locus of largest effect, and 54% by the largest three loci.

Figure 3.7 shows the response of the mean and the genetic variance, as selection changes from $N\beta = -2$to $+2$, with $N\mu = 0.3$ throughout: the approximation matches well. There is a transient increase in the genetic variance as allele frequencies pass through intermediate values. In Fig. 3.7, the shift is by 3.08 genetic standard deviations.

If mutation rate is not low, then the statistical mechanics methods apply very well. As in the case for equal effects, as mutation rate approaches 1/4N there will be deviations from

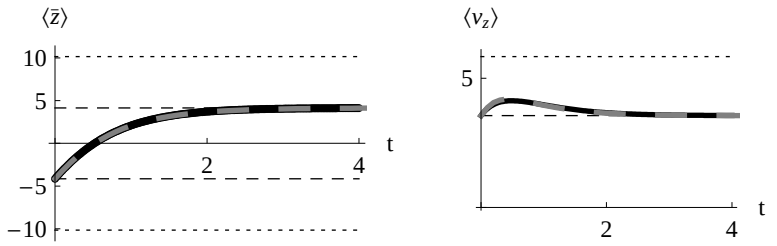


Figure 3.7: The accuracy of the approximation with unequal allelic effects, γ, given by Eq. 3.25. Nβ changes from -2 to +2 at time $t = 0$; $N\mu = 0.5$ throughout. Otherwise, as in Fig. 3.3

the true values, as compared with the diffusion equation. The accuracy of the approximations diminishes as Nμ approaches 1/4 (Fig. 3.8)

Fluctuating selection: A step towards a realistic scenario is to consider that the selective pressure is not a fixed quantity, but that it is subject to fluctuations. Besides abrupt changes in selection, we have also looked at the effects of oscillating selection. If fluctuations are sufficiently slow, then the maximum entropy approximation converges to the exact solution.

We can model this situation considering that selection is time-varying, for example as:

$$\beta(t) := \beta_0 \cos[\omega t + \varphi], \tag{3.26}$$

and study what happens for different frequencies ω. If these frequencies are low, we expect the macroscopics to respond smoothly, as it is shown in Fig. 3.9.

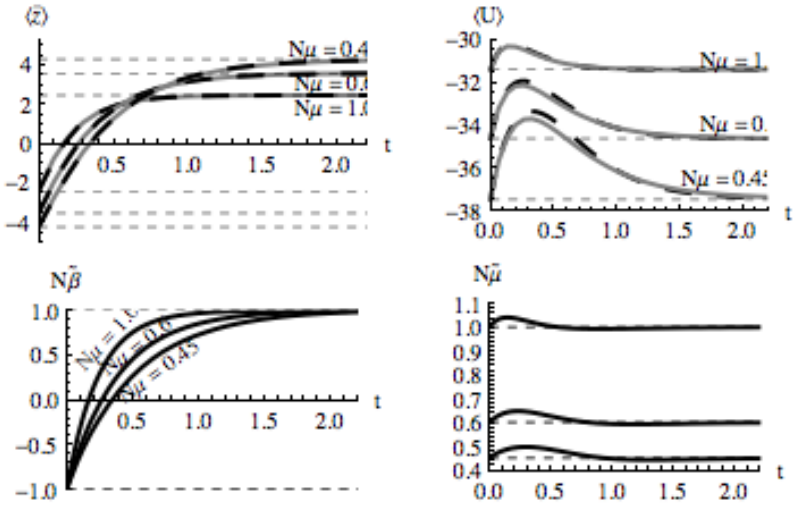


Figure 3.8: Response of the macroscopic variables composed of multiple loci, for different selective pressures (labels in the panels). Each macroscopic is determined by 10 loci of effects given by Eq. 3.25. For a sudden change in $N\beta = \pm 1$, and for mutation rates of $N\mu$ = 1, 0.6, and 0.45 we show how the predictions of the statistical mechanics work. Otherwise, as in Figs 3.3.

On the other extreme if ω is high, the variables would effectively perceive an average intensity. Notice how for high values, in Fig. 3.10, the dynamics already resemble those of a sudden change in the selective intensity from $N\beta = -1$ to $N\beta = 0$. Still, in both examples the amplitude β_0 of the oscillations is the same, but the response of the macroscopics have downscaled amplitudes. (Notice that even if ω is small, the trait never reaches $\pm\beta_0$.) Regardless of the frequency of oscillation of β, the local equilibrium holds satisfactorily (as long as $4N\mu > 1$).

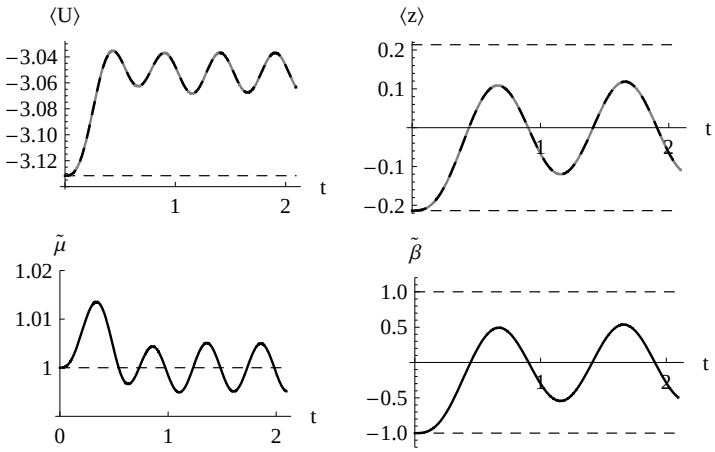


Figure 3.9: Response of the macroscopics and local variables to a low frequency of fluctuating selection intensity, as in Eq. 3.26 with $\beta_0 = -1$ and $\omega = \pi/3$. The macroscopics are given by 10 loci of different effects, following Eq. 3.25

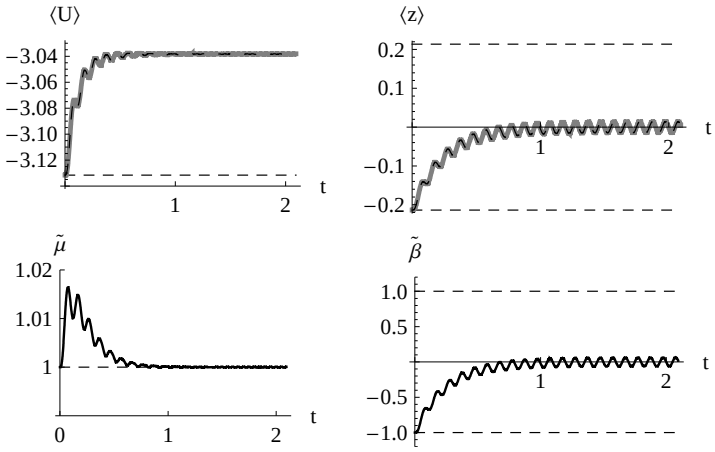


Figure 3.10: Response of the macroscopics and local variables to a fast frequency of fluctuating selection intensity, w= 20 p. Otherwise as in Fig. 3.3.

3.4 LOW MUTATION RATES: 4Nμ < 1

Failure of the maximum entropy approximation: When the number of mutations produced per generation is small ($4N\mu < 1$), populations are likely to be close to fixation. The diffusion approximation still works surprisingly well: it predicts the allele frequency distribution accurately even adjacent to the boundaries $\left(p = \frac{1}{2N}, 1 - \frac{1}{2N}\right)$. The maximum entropy approximation also makes accurate predictions for the change in trait mean, provided that the mutation rate is kept fixed (Fig. 3.5, top right). However, the approximation does not allow changes in μ^* when $4N\mu < 1$. Formally, the coefficient $B_{U,U}$ (Eq. 3.22) diverges, which implies that the effective mutation rate must always be held equal to the actual mutation rate ($\mu^* = \mu$). Thus we lose one degree of freedom from the dynamics. What causes this pathological behavior?

The key point is that near the boundary, the allele frequency distribution changes on a much faster time scale than in the centre: the characteristic time scale of random drift is determined by the number of copies of the allele in question. Thus, the shape of the distribution at the centre and at the edge is uncoupled, so that it may be impossible to adequately approximate the whole distribution as being close to the stationary state. Near the boundaries, selection is negligibly slow relative to mutation and drift, and the allele frequency distribution rapidly takes the form $p^{4N\mu-1}$, even while the bulk of the distribution remains unchanged (Fig. 3.11). For example, suppose that $4N\mu$ changes from smaller than 1 to greater than 1. The density at the boundaries immediately falls to zero, and the distribution takes on a two-peaked shape that cannot be approximated by any of the family of stationary distributions. Conversely, when $4N\mu$ falls below the threshold, small singularities immediately develops at the boundaries, representing fixed

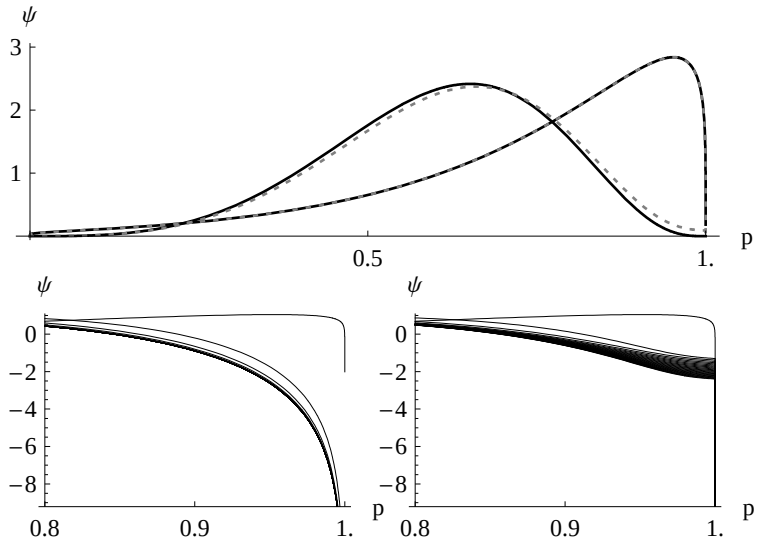


Figure 3.11: Failure of the max-entropic distribution of allele frequencies at the borders for changing mutation rates. Top panel: dotted curves: the genuine distribution, given by the diffusion equation; solid curves the max-entropic distribution. The bulk of the distributions is well approximated initially (curves towards the right; t=0) and close to equilibrium (bell-shaped curves; t=5). Lower left panel: the max-entropic distribution at different times (from t=0 to t=5, top to bottom) near the edge p=1, incorrectly predicts that there is no fixation, compared the diffusion equation solutions at different times (from t=0 to t=5, top to bottom) near the edge p=1, which shows that some genotype fix (lower right panel). Numerics as in Fig. 3.6.

populations, but it takes a long time for the bulk of the populations to approach fixation. This asymmetry explains why the maximum entropy approximation is much more accurate when $4\mathrm{N}\mu$ falls than when it rises (Fig. 3.6, $t > 5$).

We can gain some insight by analyzing the limit of $4N\mu \to 0$, when populations are almost always fixed for one of the 2^n genotypes. With directional selection, the probability of fixation of one or other allele is independent across loci, and equals $P_i = 1/(1+\exp(-4N\beta\gamma_i))$, where γ_i is the effect of alleles at the i'th locus. Populations will jump from fixation for '0' to '1' as a result of the fixation of favorable mutations, at a rate $4N\mu\beta\gamma_i/(1-\exp(-4N\beta\gamma_i))$, and in the opposite direction due to fixation of deleterious alleles, at a rate that is slower by a factor $\exp(-4N\beta\gamma_i)$. In this simple case, it is easy to write down the dynamics at each locus:

$$\frac{\mathrm{dP}_i}{\mathrm{dt}} = 4N\mu\beta\gamma_i\left(\frac{Q_i}{1+e^{-4N\beta\gamma_i}} - \frac{P_i e^{-4N\beta\gamma_i}}{1+e^{-4N\beta\gamma_i}}\right)$$
$$= 4N\mu\beta\gamma_i\frac{\left(\hat{P}_i - P_i\right)}{\left(\hat{P}_i - \hat{Q}_i\right)}\mathrm{where}\hat{P}_i = \frac{1}{1+e^{-4N\beta\gamma_i}}\ , \qquad (3.27)$$

noting that this does have a sensible limit as $\beta \to 0$: $\partial_t P = \mu(1-2P)$ which is correct for neutral alleles. The trait mean changes as:

$$\frac{d\langle\bar{z}\rangle}{dt} = 4N\mu\beta\sum_{i=1}^{n} 2\gamma_i^2\frac{\left(\hat{P}_i - P_i\right)}{\left(\hat{P}_i - \hat{Q}_i\right)}\ . \qquad (3.28)$$

The maximum entropy approximation simplifies the problem by assuming that the P_i always follow a stationary distribution, determined by a single parameter β^*, with $P_i = 1/(1+\exp(-4N\beta^*\gamma_i))$. Thus, provided we know the allelic effects, we can deduce the P_i from the observed $\langle\bar{z}\rangle$, without knowing the distribution at the n loci individually. From Eq. 3.24, assuming that $\mu = \mu^*$, we have:

$$\frac{d\langle\bar{z}\rangle}{\mathrm{dt}} = 2\frac{N\mu}{N\beta^*}\langle\bar{z}\rangle(\beta-\beta^*)\ . \qquad (3.29)$$

This can be understood by seeing that at equilibrium, selection must balance mutation, so that $\beta^* \gamma_i \langle p_i q_i \rangle = \mu \langle p_i - q_i \rangle$ at each locus if the P_i follow a stationary distribution with parameter β^*. The rate of change of the trait mean is

$$\sum_i 2 \left(\beta \gamma_i^2 \langle p_i q_i \rangle - \mu \gamma_i \langle p_i - q_i \rangle \right) =$$

$$\sum_i 2 \frac{\mu}{\beta^*} \gamma_i \langle p_i q_i \rangle - 2\mu \langle \bar{z} \rangle = 2\mu \langle \bar{z} \rangle \left(\frac{\beta}{\beta^*} - 1 \right) \quad ,$$

equal to Eq. 3.29.

It is easy to show that the maximum entropy approximation, Eq. 3.29, converges to the exact solution, Eq. 3.28, for small N$\beta\gamma_i$; this is confirmed by Fig. 3.12, for $N\beta = 0.2$, in an example with equal effects, $\gamma_i = 1$. However, for stronger selection ($N\beta = 2$ (thick lines in Fig. 3.12), the maximum entropy approximation underestimates the initial rate of increase. That is because the approximation is that the initial state, in which $P_i = 0.02$ at all loci, is caused by strong selection against the '1' allele; such selection would necessarily cause low standing variation, and so the prediction is for a slow response when the direction of selection is reversed. However, as soon as selection is reversed, populations fix new favorable mutations at a rate that is *independent* of the previous standing variation. Thus, the method that led to Eq. 3.24, which was developed for polymorphic populations, fails as $4N\mu \to 0$.

Maximum entropy for $4N\mu \to 0$: When mutation is rare, populations are almost always fixed for one of the 2^n genotypes, and an ensemble of populations (or equivalently, the probability distribution of a single population) evolves as a result of jumps between genotypes, mediated by fixation of single mutations. The stationary distribution is proportional to $\bar{W}^{2N}$, multiplied

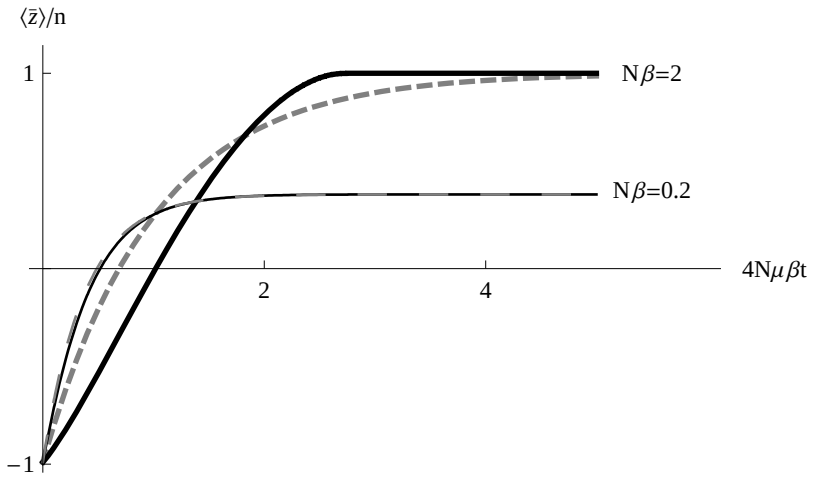


Figure 3.12: Comparison between the exact solution (Eq. 3.28) and the maximum entropy approximation (Eq. 3.29), in the limit of low mutation rates (4 $N\mu \to 0$). Initially, the probability that a locus is fixed for the '1' allele is $P = 0.02$ at all loci, so that $\langle \bar{z} \rangle / n = -0.99$; all alleles have effect γ=1. Selection $N\beta$ =0.2 or $N\beta = 2$ is then applied, and the trait mean shifts to a new equilibrium, in which a fraction $P = 1/(1 + \exp(-4\mathrm{N}\beta))$ of loci are fixed for the '1' allele. When selection is weak ($N\beta = 0.2$), the maximum entropy approximation is barely distinguishable from the exact solution. However, when selection is strong ($N\beta = 2$), the maximum entropy approximation (dashed lines) underestimates the initial rate of change.

by a factor that reflects the pattern of mutation rates (Iwasa, 1988; Sella and Hirsh, 2005); this can be derived as the limit of Eq. 3.1 for small $4N\mu$ (Barton and Coe, 2009). We can go further, and apply the maximum entropy method to this process. This gives an approximation for the dynamics of macroscopic quantities such as $\langle \bar{z} \rangle$, so that we do not need to follow the full distribution across the 2^ngenotypes. In the simplest case of directional selection on an additive trait, with equal allelic effects, this gives no benefit, since the distribution of fixation probabil-

ity is independent across loci, and moreover, is the same at each locus: the problem therefore involves just a single variable, P. However, with unequal effects, the maximum entropy approximation does give a useful simplification, since we do not need to follow the individual P_i. With epistasis for fitness or for the trait, the advantage would be greater, since we would then avoid following the full probability distribution, across the 2^n genotypes. (Note that in the limit of $4N\mu \to 0$, the model applies regardless of the pattern of recombination, because only one locus evolves at a time).

We now apply the maximum entropy approximation to directional selection on an additive trait, assuming that $4N\mu \to 0$, but allowing for unequal allelic effects, γ_i. This is distinct from the previous section, since we now apply maximum entropy to the limiting system, rather than apply the limit of $4N\mu \to 0$ to the full maximum entropy approximation). The system is described by a single local variable, β^*, defined implicitly by $\langle \bar{z} \rangle = \sum_i \gamma_i \tanh [2N\beta^* \gamma_i]$; the assumption is that at each locus, $(P_i - Q_i) = \tanh [2N\beta^* \gamma_i]$, as if the ensemble were at a local stationary state under a selection gradient β^*.

Thus:

$$\begin{aligned}
\frac{d \langle \bar{z} \rangle}{dt} =& \sum_i 2\gamma_i \frac{dP_i}{\mathrm{dt}} \\
=& \sum_i 2\gamma_i \left(4N\mu\beta\gamma_i \left(\frac{Q_i}{1 + e^{-4N\beta\gamma_i}} - \frac{P_i e^{-4N\beta\gamma_i}}{1 + e^{-4N\beta\gamma_i}} \right) \right) \\
=& 4N\mu\beta \sum_i \gamma_i^2 \left(1 - \frac{\tanh [2N\beta^* \gamma_i]}{\tanh [2N\beta\gamma_i]} \right) \ . \qquad (3.30)
\end{aligned}$$

It is easier to work in terms of β^*. Multiplying by $\mathrm{d}\beta^* / d\bar{z}$, we

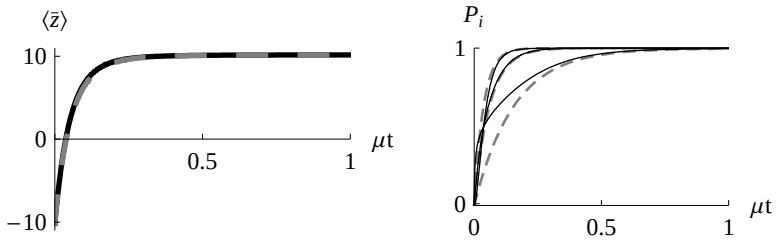


Figure 3.13: The maximum entropy approximation (Eq. 3.30), made assuming that populations jump between fixed states, gives an accurate prediction for the change in mean (left): this is indistinguishable from the exact solution (Eq. 3.28). The population is initially at equilibrium with directional selection N$\beta = -4$; selection then changes sign abruptly. Allelic effects are given by Eq. 3.25. Predictions for the underlying allele frequencies are less accurate. The right panel shows allele frequencies at the locus with the strongest effect ($\gamma_1 = 1.69$), with intermediate effect ($\gamma_5 = 1.04$), and weakest effect ($\gamma_1 = 0.401$), reading left to right. Solid lines show the maximum entropy approximation, (Eq. 3.30), and dashed lines, the exact solution (Eq. 3.28).

obtain a closed equation for β^*:

$$\frac{\mathrm{d}\beta^*}{dt} = 2\mu\beta \frac{\sum_i \gamma_i^2 \left(1 - \frac{\tanh[2N\beta^*\gamma_i]}{\tanh[2N\beta\gamma_i]}\right)}{\sum_i \gamma_i^2 \left(1 - \tanh\left[2N\beta^*\gamma_i\right]^2\right)} \,. \tag{3.31}$$

When selection is weak ($2N\beta^*\gamma_i << 1$), Eq. 3.30 simplifies to $4N\mu \sum_i \gamma_i^2 (\beta - \beta^*)$. Since $2N\beta^* \sum_i \gamma_i^2 \sim \langle \bar{z} \rangle$ in this limit, this converges to $\left(4N\mu \sum_i \gamma_i^2\right)\beta - 2\mu \langle \bar{z} \rangle$. The exact solution, Eq. 3.28, converges to the same limit with weak selection. This is as expected, since when selection is weak, the population approaches a mutation-drift equilibrium, with genetic variance $\left(4N\mu \sum_i \gamma_i^2\right)$; the trait mean then changes at a rate $\left(4N\mu \sum_i \gamma_i^2\right)\beta$ due to selection, and $-2\mu \langle \bar{z} \rangle$ due to mutation.

Figure 3.13 shows an example where selection is strong, changing abruptly from $N\beta$ = -4 to +4. The effects of 10 loci are

drawn from a gamma distribution, as in Eq. 3.25. The predictions for the mean are indistinguishable (Fig. 3.13, left). There are substantial errors in the predictions for the underlying allele frequencies, with the rate of change of alleles of small effect being greatly overestimated (Fig. 3.13, lower curve at right), and that of alleles of large effect, slightly underestimated. However, these errors almost precisely cancel in their effects on the mean.

3.5 DISCUSSION

The maximum entropy approximation: A fundamental aim of quantitative genetics is to understand the evolution of the phenotype, without knowing the underlying distribution of all possible gene combinations. Assuming linkage equilibrium simplifies the problem, which then depends only on the allele frequencies, rather than on the full distribution of genotypes. However, if we include random drift, as well as selection and mutation, a full description of the stochastic dynamics requires the distribution of allele frequencies - a formidable task. We know that in general, we cannot predict phenotypic evolution without knowing the frequencies of all the relevant alleles: the future response to selection may depend on the frequencies of alleles that are currently so rare that they have negligible effect on the phenotype, and so are essentially unpredictable. To avoid this difficulty, we make the key assumption that selection and mutation act only on observable quantities. Then, the distribution of allele frequencies tends towards a stationary state that depends only on those forces. If selection could instead act on individual alleles, it could send the population into arbitrary states by picking out particular alleles (e.g. Fig. 3.1). Selection on individual alleles would be analogous to Maxwell's Demon,

which perturbs individual gas molecules so as to generate improbable states that violate the laws of classical thermodynamics (Leff and Rex, 2003).

Populations tend towards a stationary state that maximizes entropy - that is, the distribution of allele frequencies spreads out as widely as possible, conditional on the average values of the quantities that are acted on by selection and mutation. The maximum entropy approximation to the dynamics amounts to assuming that the allele frequency distribution always maximizes entropy, given the current values of the observed variables, even though those variables may be changing. This approximation converges to the exact solution when changes in mutation and selection ($\vec{\alpha}$) are slow. However, we find that even if selection abruptly changes in direction, predictions for the trait mean are remarkably accurate.

The analogy between the population genetics of quantitative traits, and statistical mechanics, is intriguing. As well as suggesting methods for approximating phenotypic evolution, it also helps us to better understand the scope of statistical mechanics, by showing that it does not depend on physical principles such as conservation of energy (Ao, 2008). Selection can be seen as generating information, by picking out the best-adapted genotypes from the vast number of possibilities, despite the randomizing effect of genetic drift. This is analogous to the way that a physical system does useful work, despite the tendency for entropy to increase. Such issues are discussed by Barton and Coe (2009) and in Ch. 5. Here, we concentrate on the use of maximum entropy as an approximation procedure.

Provided that the number of mutations, $4N\mu$, is constant, and not too small, the method accurately predicts the evolution of the trait mean - even when allelic effects vary across loci, and even when selection changes abruptly. This accuracy even when parameters change rapidly is surprising, because the un-

derlying allele frequencies may not be well-predicted (e.g. Fig. 3.13). Indeed, Prügel-Bennett, Rattray and Shapiro make accurate predictions even though they use an arbitrary entropy measure that does not ensure convergence to the correct stationary distribution. Although we believe that our entropy measure is the most natural for quantitative genetic problems, and it does guarantee convergence to the stationary state, it may be that the maximum entropy approximation is an efficient method for reducing the dimensionality of a dynamical system, even when an unnatural measure is used.

The maximum entropy method predicts the full allele frequency distribution from just a few quantities, such as the expected trait mean, $\langle \bar{z} \rangle$. We do still need to know the genetic basis of the trait - for an additive trait, we must know the allelic effects. We could hardly expect to predict the evolution of phenotype without knowing anything about its genetic basis. However, we could apply the method knowing just the distribution of allelic effects, which could be estimated in a number of ways: by detection of QTL, from evolutionary arguments about plausible distributions (e.g. Orr, 2003) or from the distribution of allele frequencies at synonymous and non-synonymous sites (e.g. Loewe et al., 2006).

Extension to dominance and epistasis: We have only analyzed the simplest case, of directional selection on an additive trait. An extension to allow dominance is straightforward, since the loci still fluctuate independently of each other, and the generating function, $\mathbb{Z}$, can still be written as a product of integrals across loci. Extension to more than two alleles is also possible, but only under the restrictive condition that mutation rates allow for detailed balance, and hence for an explicit potential function. Similarly, alleles at different loci may interact in their effect on the trait. If such epistasis involves non-overlapping

pairs of loci, then calculations can still be made, but require integrals over pairs of allele frequencies. Though it is beyond the scope of this paper to present such calculations, it is important to point out that, despite the technical difficulties, the method itself is general and as such it does not depend on the selective scheme, epistatic model, number of alleles, etc.

It is relatively straightforward to allow for stabilizing selection on an additive trait. In this case, allele frequency distributions at different loci are no longer independent. However, they are only coupled via a single variable, the trait mean: if this lies above the optimum, then all loci experience selection for lower $\bar{z}$, and vice versa. This simple coupling allows explicit solutions for the stationary distribution, and for the rate of jumps between metastable states. These calculations are given in Barton (1989) and Coyne et al. (1997, appendix). We outline the maximum entropy approximation to the dynamics of stabilizing selection in Chapter 7.

For complex models, involving epistasis between large numbers of genes, calculation of the maximum entropy approximation (i.e., of the matrices B^*, C^*) by numerical integration would not be feasible. They could still be calculated by a Monte Carlo method: one would fix the parameters $\vec{\alpha}^*$, and simulate the distribution to determine the expectations $\left\langle \vec{A} \right\rangle$. The matrix C could be found from the covariance of fluctuations, and the matrix B from Eq. 3.13. The two matrices, $B\left(\vec{\alpha}^*\right)$ and$C\left(\vec{\alpha}^*\right)$, would then give the dynamics on the reduced space of $\vec{\alpha}^*$; this would be feasible numerically for two or three variables. Of course, this approach involves the same kind of computation as a direct simulation. Our claim is that the reduced dynamics will be approached, regardless of the initial allele frequency distribution: the system is expected to move close to the lower-dimensional space defined by the maximum entropy approximation. The implication is that we could predict the evolution of the expec-

tations $\left\langle \vec{A} \right\rangle$ by a closed set of equations, without knowing the actual allele frequency distribution. This will require that we know the genetic basis and mutability of the trait, and that selection acts only on that trait.

Low mutation rates ($4N\mu < 1$): We describe mutation and selection by using a potential function $\mu U + \log\left(\bar{W}\right)$, where U = $2\sum_i \log\left(p_i q_i\right)$, and include the variable $\langle U \rangle$ together with selected variables such as the expectation of the trait mean, $\langle \bar{z} \rangle$. However, this approach fails to describe the effects of changes in mutation rate when $4N\mu < 1$, because then, populations are likely to be close to fixation, in which case U diverges. The fundamental problem is that the distribution at the boundaries changes rapidly as mutation rate changes, whilst the bulk of the distribution does not. We can, however, extend the method to the case where $4N\mu$ is very small, because then, populations jump between fixation for one or other genotype, through the substitution of single mutations. This limit is in fact more general, in that it applies even with linkage or with asexual reproduction. It could be extended to give a more accurate approximation for appreciable $4N\mu$, by calculating the probability of a jump between states of near fixation, taking into account the polymorphism at other loci (see Barton and Turelli, 1989).

Long-term response to selection

A basic and long-standing puzzle in quantitative genetics is the success of artificial selection: in moderately large populations, traits respond steadily to selection for a hundred generations or more, with little change in additive genetic variance, and often, with concordance between replicates (Barton and Keightley, 2002). This is surprising, because the genetic variance is expected to change as alleles sweep through the population.

However, if the distribution of allele frequencies is proportional to $(pq)^{4N\mu-1}$, as we assume, and if $4N\mu$ is small, then the additive genetic variance is expected to stay constant for long periods under directional selection. This is because the baseline distribution $\phi(p) = (pq)^{-1}$ is uniform when transformed to a logit scale (i.e., $\phi(z) =$ constantfor $z = \log(p/q)$). Since $\log(p/q)$ increases linearly with time under directional selection, that implies that the increase in genetic variance due to rare alleles increasing to become common is precisely balanced by the decrease due to common alleles approaching fixation. Thus, the response to standing variation is expected to continue steadily at a rate $d\,\langle\bar{z}\rangle/\,dt\, = 4N\mu\beta\sum_i \gamma_i^2$ for $\sim (1/s)\log(2N)$ generations, whereas if alleles were typically polymorphic (as would be the case if $4N\mu > 1$), it would continue for only $\sim (1/s)$ generations. Of course, the response will continue indefinitely as a result of new mutation, at just the same rate. This is because variation is initially maintained in a balance between mutation and drift; the genetic variance is not affected by directional selection, and so the rate of response stays the same even as it shifts from alleles that were originally present, to new mutations.

The stationary density under mutation, selection and drift has been exploited before to help understand the evolution of quantitative traits (e.g. Keightley and Hill, 1987; Keightley, 1991). In this paper, we have shown that the dynamics of polygenic traits can be accurately approximated by assuming that the underlying distribution of allele frequencies always takes this stationary form. We are now starting to get detailed estimates of the distribution of allele frequencies and of allelic effects on traits and on fitness (e.g. Loewe et al., 2006; Boyko et al., 2008): it may be that we will soon be able to use such data to apply the methods developed here to natural and artificial populations.

Manuscript in preparation: H.P. de Vladar and I. Pen. – *G spotted in Rana temporaria!*

Chapter 4

$\mathcal{G}$ spotted in *Rana temporaria*!

> The standard selection equations have been taken too literally; and genetic assumptions with little empirical support have gained undue credibility.
>
> Nick Barton and Michael Turelli

Abstract

Plenty of work has focused in understanding the evolution of the genetic covariances ($\mathcal{G}$ matrix). Yet this empirical and theoretical knowledge has not fully merged. Thus we lack the big picture about $\mathcal{G}$'s evolution. We present a model that considers how $\mathcal{G}$ relates to allele frequencies and pleiotropic structure. Averaging over these gives estimators of $\mathcal{G}$ that are independent of these variables, but which depend only on mutation rate, selection differentials, population size and allelic effects. The latter may be approximated by average values. The model thus integrates the mechanisms of population with quantitative genetics, but requires only phenotypic (quantitative) information. This is already a significant achievement. However we apply our ideas to previous results in experimental evolution of Rana temporaria, addressing a classical question: which factors affect the diversification of $\mathcal{G}$? We give concrete answers on the role of selection, mutation, and drift in the observed experimental patterns.

4.1 INTRODUCTION

Since the pioneering work of Lande (1979) much research has been done to understand the evolution of genetic covariances (Steppan et al., 2002; Blows, 2007; Arnold et al., 2008). These are essential for understanding the evolution of metric traits (Barton and Turelli, 1987; Bürger, 1991). Throughout the evolutionary process, for few generations genetic variation remains unchanged (Turelli, 1988). As with the breeder's equation, the formula $\Delta\bar{z} = G.P^{-1}.\vec{\beta}$ would allow prediction of the mean of multivariate traits of a population, $\vec{z}$ (Lande, 1979, 1980). $\mathcal{G}$ is the genetic covariances (of the traits) matrix , P is the matrix of phenotypic covariances, and $\vec{\beta}$ is a vector of selection differentials. Under certain conditions $\mathcal{G}$ can be stable across generations (Brodie, 1993; Roff, 2000; Begin and Roff, 2001, 2003; Nosil et al., 2006; Renaud et al., 2006). However, other observations show that many factors affect $\mathcal{G}$'s constancy (Wagner, 1984; Shaw et al., 1995; Roff, 2000; Phillips et al., 2001; Widen et al., 2002; Cano et al., 2004; Kotiaho, 2007; Doroszuk et al., 2008), which are also supported by theoretical understandings (Turelli, 1988; Reeve, 2000; Jones et al., 2003, 2004, 2007). But the theories on the evolution of $\mathcal{G}$ are incomplete (Arnold et al., 2008), hence employing measurements of the $\mathcal{G}$ at one given time, might not be enough to explain or predict phenotypic variation and diversification in ecological or evolutionary times (Steppan et al., 2002; Blows, 2007; Kotiaho, 2007). Selection aligns $\mathcal{G}$ to evolve in a particular direction (Reeve, 2000; Roff, 2000; Steppan et al., 2002; Jones et al., 2004), but random drift make it wobble unpredictably generation after generation (Roff, 2000; Jones et al., 2003; Arnold et al., 2008). Mutation, depending on the degree of pleiotropy and linkage, and migration will act like a torque inducing or reducing the correlations among the traits (Jones et al., 2003,

2004; Guillaume and Whitlock, 2007; Arnold et al., 2008). Despite this knowledge on how different factors affect $\mathcal{G}$, the quantitative predictions are raw (Arnold et al., 2008). Hence, empirical quantifications of $\mathcal{G}$ are hard to relate to the theoretical knowledge.

Our first brass ring is to develop a theory that comprehends population and quantitative genetics, that is applicable for multivariate response to selection, mutation, and drift (SMD). A previous approach predicted the evolution of a quantitative characters subject to SMD, considering the influence of the genetic states (allele frequencies), but without making direct reference to them (Barton and de Vladar, 2009; Barton and Coe, 2009). We extend these methods to the multivariate case, with pleiotropic effects. A given trait is affected by a set of genetic variables (e.g. allele frequencies), whose distribution can be described by the Wright-Fisher SMD equilibrium distribution (Crow and Kimura, 1970, pp. 442-445). From this distribution, we can calculate a generating function, which considers all genetic states and averages over them. Thus it is implicitly dependent on the genetic variables, but depends explicitly only on the selective gradients over each trait $\vec{\beta}$, mutation rate μ, population size N and the additive effects of each locus over the trait, γ. The expectancies of the mean traits $\vec{z}$, genetic co-variances matrix $G = \{\nu_{ij}\}$, phenotypic covariances, etc. can be calculated from the generating function. Ergo, we offer a method to calculate quantitative aspects of a population's traits in such a way that genetics is not disregarded, but the knowledge of its details is dispensable for the quantitative description of the population. This merging of the genetic with the quantitative variables has been a riddle for decades, and its failures imbued the understanding of the $\mathcal{G}$'s stability. Although we have by no means solved all questions regarding the evolution of $\mathcal{G}$, our results are opportune to address some of the relevant aspects

about quantitative evolution.

Much is to be done in the theorety of $\mathcal{G}$, but at this point we are encouraged to formulate our questions by empirical motivations. We seek a marriage between the practical needs and the theoretical capabilities. Specifically, we chose to re-evaluate the experimental results of Cano et al. (2004) see also Laurila et al. (2002); Palo et al. (2003); Ovaskainen et al. (2008) from which the non - constancy of $\mathcal{G}$ has been verified for four correlated traits in *Rana temporaria*: development time, mass, body length, and tail length. Employing suitable experimental design and statistical analyses the authors verified that the $\mathcal{G}$ matrices of two populations were statistically different (Cano et al., 2004; Ovaskainen et al., 2008). However Jones et al. (2003), identified that drift is a major source of fluctuations of $\mathcal{G}$. Whilst selection effectively affects $\mathcal{G}$, it would have lamer and predictable repercussions than drift (Roff, 2000). Our second goal is to appraise the roles of selection and drift from the data of Cano et al. (2004). Employing the proposed theoretical construct we will characterize from the trait data the conditions maintained at two distinct selection-mutation-drift (SMD) equilibria. Then we will predict the corresponding $\mathcal{G}$-matrices, which we compare to the empirical estimations. Randomly sampling the distribution of allele frequencies and computing $\mathcal{G}$ for these, illustrates the variability that the genetic convariances can show, and whether it is (or not) a plausible explanation for the observed diversification in $\mathcal{G}$, instead or along with selection.

4.2 THEORETICAL BACKGROUND

Throughout this paper, we will assume that the populations are in Hardy-Weinberg equilibrium. Consider M autosomal traits, affected by N independent loci (i.e. in linkage equilibrium), each

traitof them is determined by the contribution of the diploid set at each locus $x_\ell^{\female}$ and $x_\ell^{\male}$, $z_m = \sum_{\ell=1}^{n} \gamma_{m\ell} \left(x_\ell^{\female} + x_\ell^{\male} - 1 \right)$, where $x_\ell^{\female}$ and $x_\ell^{\male}$ are either 0 or 1 (unfavorable or favorable copies of the alleles). Averaging x_ℓ over the population, and calling p (q) the frequency of $x = 1$ ($x = 0$) in a population, the mean trait results in

$$\bar{z}_m = \sum_{\ell=1}^{n} \gamma_{m\ell} \left(p_\ell - q_\ell \right) \ , m = 1, 2, \ldots, M \tag{4.1}$$

where $\gamma_{m\ell}$ is the effect of locus l over the trait m. We consider only additive on all traits (there is neither epistasis nor dominance), but pleiotropic effects are present (unless $\gamma_{m\ell} = 0$). We consider selection over all traits to be directional and of exponential nature (Kingsolver et al., 2001; Hoekstra et al., 2001): $\bar{W} = \exp\left[\vec{\beta}.\vec{z}\right]$, $\vec{\beta}.\vec{z} = \beta_1 \bar{z}_1 + \beta_2 \bar{z}_2 + \ldots \beta_M \bar{z}_M$. The gradient of log-mean fitness is $\frac{\partial}{\partial p_\ell} \log\left[\bar{W}\right] = \beta_m \gamma_{m\ell}$, where $\beta_m$is the intensity of selection over the trait *m*. The rate of change of the frequency p at every locus, including SMD is given by the Wright-Fisher model (Wright, 1938; Kimura, 1955):

$$\frac{\partial \psi}{\partial t} = M_{\delta p} \frac{\partial \psi}{\partial t} + \frac{1}{2} V_{\delta p} \frac{\partial^2 \psi}{\partial t^2} \tag{4.2}$$

$$M_{\delta p} = \underbrace{pq\beta}_{\text{selection}} + \underbrace{\mu(2p-1)}_{\text{mutation}} \tag{4.3}$$

$$V_{\delta p} = \underbrace{\sqrt{\frac{pq}{2N}}}_{\text{drift}} \tag{4.4}$$

where $q = 1 - p$, μ is the mutation rate, and ζ represents the drift, as a normal distribution with variance $\frac{pq}{2N}$, and N the size of the population. This leads to the classical equilibrium distribution of joint allele frequencies (Wright, 1938; Crow and

Kimura, 1970, pp. 442-445)

$$\psi = \frac{\mathbb{Z}^{-1}}{V_{\delta p}} \exp\left[\int \frac{M_{\delta_p}}{V_\delta p} dp\right] \tag{4.5}$$

Here $\mathbb{Z}$ is the normalizing constant:

$$\mathbb{Z} = \int \exp\left[2N\vec{\beta}.\vec{z} + 2N\mu U\right] / \prod_{\ell=1}^{n} p_\ell q_\ell d\vec{p}. \tag{4.6}$$

where with $U = 2\sum_{\ell=1}^{n} \log(p_\ell q_\ell)$, the contribution by mutation of all loci to the quantiative evolutionary potential. Notice that beyond just normalizing, it is a generating function; taking derivatives of Log($\mathbb{Z}$) with respect to β_m and μ, leads to the expected values of the trait, and of the mutation effects U:

$$\begin{aligned} \frac{\partial \log \mathbb{Z}}{2N\partial\beta_m} &\equiv \langle \bar{z}_m \rangle & (4.7)\\ &= \frac{1}{\mathbb{Z}} \int \bar{z}_m \exp\left[2N\vec{\beta}.\vec{z} + 2N\mu U\right] / \prod_{\ell=1}^{n} p_\ell q_\ell d\vec{p} \\ \frac{\partial \log \mathbb{Z}}{2N\partial\mu} &\equiv \langle U \rangle & (4.8)\\ &= \frac{1}{\mathbb{Z}} \int U \exp\left[2N\vec{\beta}.\vec{z} + 2N\mu U\right] / \prod_{\ell=1}^{n} p_\ell q_\ell d\vec{p} \end{aligned}$$

The angle brackets $\langle \ldots \rangle$ indicate statistical expectation over drift. The reader can check that the second derivatives correspond to variances and covariances of the population means. The interesting issue is that if there is an algebraic expression for $\mathbb{Z}$, the expectations can be calculated explicitly. Indeed, $\mathbb{Z}$ is:

$$\begin{aligned} \mathbb{Z} &= \prod_{\ell=1}^{n} \mathbb{Z}_\ell \left(\mu, \vec{\beta}.\vec{\gamma}_\ell\right), & (4.9)\\ \mathbb{Z}_\ell &= \sqrt{\pi} 2^{1-8N\mu} \Gamma(4N\mu)\, {}_0\tilde{F}_1 \left(4N\mu + 1/2; N\vec{\beta}.\vec{\gamma}_\ell\right) \end{aligned}$$

where $\vec{\gamma}_\ell = (\gamma_{1\ell}, \gamma_{2\ell}, \ldots \gamma_{M\ell})$ is the vector of effects of locus ℓ over each trait. In the last expression, Γ is the Gamma function, and ${}_0\tilde{F}_1$ is the regularized hypergeometric of order (0,1)

(this can also be written as Bessel functions, see Barton and de Vladar (2009)). Explicit formulas for the values of the mean traits follow from the derivatives:

$$\langle \bar{z}_m \rangle = \sum_{\ell=1}^{n} \gamma_{ml} \frac{I_{4N\mu+1/2}\left(2N\vec{\beta}.\vec{\gamma}_\ell\right)}{I_{4N\mu-1/2}\left(2N\vec{\beta}.\vec{\gamma}_\ell\right)}, \tag{4.10}$$

The elements of $\mathcal{G}$ are,

$$\nu_{mr} = 2\sum_{\ell=i}^{n} \gamma_{m\ell}\gamma_{r\ell}p_\ell q_\ell \ , \tag{4.11}$$

whose expectations are can also be given explicitly:

$$\langle \nu_{mr} \rangle = 2N\mu \sum_{\ell=1}^{n} \frac{\gamma_{m\ell}\gamma_{r\ell}}{N\vec{\beta}.\vec{\gamma}_\ell} \frac{I_{4N\mu+1/2}\left(2N\vec{\beta}.\vec{\gamma}_\ell\right)}{I_{4N\mu-1/2}\left(2N\vec{\beta}.\vec{\gamma}_\ell\right)}. \tag{4.12}$$

We point out for the reader, that beyond the mathematics, the relevance of the expressions (4.10 and 4.12) is that they consider the genetic states by construction, but the expressions themselves are not explicitly dependent on the allele frequencies. Thus the expectations embed the genetic variables with the quantitative traits, merging both levels of description.

If the algebraic equations above are not of much insight, the reader may still notice that they can be used for making estimations from the data. We presented only those formulas of immediate interest for this study, but any other statistic can be calculated from direct integration (at worse numerically), or by derivatives of $\log(\mathbb{Z})$ (hints: covariance, higher moments of the trait, like skewness or kurtosis, etc. see Barton and de Vladar (2009)).

The dynamics of the expectations of the trait, can be calculated substituting the formulas of $\bar{z}_m$ and U in Eqns. 4.7 and

4.8 and using the rule of chain with Eq. 4.2 to calculate the rates of change. Details are presented in Barton and de Vladar (2009). In short:

$$\frac{d}{dt}\left\langle \begin{array}{c} \vec{z} \\ U \end{array} \right\rangle = \left\langle \begin{array}{cc} \mathcal{G} & \vec{z} \\ \vec{z} & H \end{array} \right\rangle \cdot \begin{pmatrix} \vec{\beta} \\ \mu \end{pmatrix} \tag{4.13}$$

U is included because together with log-mean fitness (in this case, the traits) couple the eeffects of mutation to the change of $\mathcal{G}$ at all time points (Eq. 4.2), and H is the genetic variance of the mutation effects U. This treatment of the effects of mutations is somehow different to the mutation matrix M (e.g. Jones et al. (2007)). We are for the moments uncertain about the relationship between M and U and H. However we know that Eq. 4.13 faithfully leads to long-term predictions Barton and de Vladar (2009). We approximate the genetic variances $\mathcal{G}$ and H, by the corresponding expectancies (e.g. Eqns. 4.10 - 4.12). Notice that the parameters $\vec{\beta}$ and μ (Barton and de Vladar, 2009; Barton and Coe, 2009), are not bound to be constant, but are allowed to change in order to keep the distribution of allele frequencies coupled to the evolutionary dynamics. Details on this method can be found in (Barton and de Vladar, 2009).

4.3 MATERIALS AND METHODS

Summary of the Data

In the original study, Laurila et al. (2002), collected female and male frogs from two Swedish populations of *Rana temporaria*. The Kiruna 'Northern' population lives in a stream that rarely (if ever) experiences desiccation. The Lund, or 'Southern' population, is situated in a pond that dries up frequently (Laurila et al., 2002). For each location, eggs of four females were ar-

tificially fertilized with sperm of five males for total of 45 full-sib families. Once the tadpoles reached certain developmental stage, 18 of them from each cross were individually dispensed to vials with 0.75 L of water and allowed them to develop until metamorphosis. In the meantime, each tadpole was exposed to one of the three desiccation treatments: control (constant water level), slow (reduction of water level by 15% at each water change), and fast (reduction of water level by 30% at each water change). At metamorphosis the tadpoles were weighed and the body and tail length of the individuals were measured. Their development time (days elapsed from the start of the experiment until metamorphosis) was also measured. To avoid scaling effects and to homogenize variances, the natural logarithm of the trait values was used in the analyses (Further details can be found in Laurila et al., 2002; Cano et al., 2004).

To estimate $\mathcal{G}$ the authors fitted a linear animal model (Lynch and Walsh, 1998, pp. 755-758) that considered the additive genetic effects of the pedigree structure, additive genetic effects, maternal identities in the pedigree, and nonadditive genetic effects (i.e., dominance and epistasis). T-tests were employed to verify whether the estimations of the heritabilies and genetic correlations were significantly different from zero. The $\mathcal{G}$ -matrix values (and other estimations) are reported in Cano et al. (2004) study (but see also Ovaskainen et al. (2008)).

Quantitative Estimations

Typically inbreeding experiments are used to estimate the number of loci and their effect over the traits (Wright, 1968; Lande, 1981; Ollivier and Janss, 1993). But we proceeded by a different method, since these experiments are at the moments not available. We calculate the minimal number of loci and their

average effect over a trait to be respectively:

$$\underset{\sim}{n} = \frac{(\tilde{z})^2}{2\tilde{\nu}},\ \bar{\gamma} = \frac{2\tilde{\nu}}{\tilde{z}}.$$

Where $\tilde{z}$ and $\tilde{\nu}$ are the maximal meant trait and genetic variance. We pooled all the data of the populations, and performed a bootstrap analysis to estimate $\tilde{z}$. Also assuming the pooling of data, the maximal genetic variance was calculated as $\mathrm{Var}_{\mathrm{tot}} = \mathrm{Mean}(\nu) + \mathrm{Var}\,(\bar{z})$. Both quantities were compared to the actual occurring maxima in the individual samples. See Supplementary information for further details on these estimations.

There can be many possible patterns of pleiotropic interactions affecting the traits for a given number of loci and their effects. We performed a random Monte-Carlo generator to sample the space of pleiotropic architectures (see Supplementary Material). For each of these architectures, we numerically computed the solution to $\langle \vec{z}_m \rangle_{(\vec{\beta}|N,\mu)} = \hat{\vec{z}}_m$, to obtain the variables $\vec{\beta}$ (one selection differential for each measured mean trait). For this we assumed a population size N and mutation rate μ. The left side of the equation is derived from the generating function, Eq. 4.10, and the right hand side are empirical estimations from Laurila et al. (2002).

To asses the effects of drift, we resampled the distribution of allele frequencies Eq. 4.5 to generate a hypothetical populations for each of the estimated scenarios. For each of these samples of the allele frequencies, the $\mathcal{G}$ matrix was calculated and plotted. Each element of $\mathcal{G}$ is computed from the definition of ν, Eq. 4.11.

A similar procedure was employed to estimate the effects of sampling within a population (simulating a field sampling procedure), but instead of randomly choosing values with the distribution 4.5, we resampled one particular $\hat{p}$ (which we assumed

as the expectancy $\langle\vec{p}\rangle = \frac{1}{2}(\langle\vec{z}\rangle + 1)$). Each iteration simulated the effects of sampling a particular population, in which we end up with an array of 'measurements' for particular hypothetical individuals. For each of these populations we re-estimate the allele frequencies (i.e. $\hat{p}$), and computed and plotted $\mathcal{G}$.

4.4 RESULTS

Table 4.1 reports the results for the maximum values of the mean traits and genetic variances, effective number of loci, and average effect of the alleles. The reader is deferred to the Supplementary Material for details on the results of the bootstrap analyses analyses, and Monte Carlo search in patterns of pleiotropic interactions. For each of the 444 resulting pleiotropic architectures, we calculated the values of $\vec{\beta}$ that match the empirical mean traits with Eq. 4.10. Pleiotropic architectures that were not compatible with the observed values, at a mutation rate of 10^{-3} and population size N=300 (following Palo et al. (2003)), were discarded. The distribution of $\vec{\beta}$ was different for different treatments and specially for different locations (see Supplementary Material). Incidentally, not all pleiotropic structures allowed solutions for the given empirical values. Thus solving Eq. 4.13 not only resulted in the identification of the SMD conditions, but also discriminated the possible pleiotropic structures which that are consistent with the data. We found 57 pleiotropic overlaps that are consistent with data (see Supplementary Material), and which in turn happened to be common for the estimations at both locations and all treatments.

Then we forecasted the $\mathcal{G}$ -matrices for each estimations of $\vec{\beta}$ and their pleiotropic structures, and averaged over the latter. The eigenstructure of these averages are in good agreement with those of the empirical $\mathcal{G}$'s (Table 4.1), specially for

Table 4.1: Quantitative Estimations from the data, and theoretical predictions of $\mathcal{G}$.

Trait[a]					North		South	
	$\tilde{z}_{max}$	ν_{max}	n	$\bar{\gamma}$	$\hat{\nu}$	$\langle\nu\rangle$	$\hat{\nu}$	$\langle\nu\rangle$
1	2.51	0.9	4	0.627	0.64	0.94	0.9	0.94
2	0.52	$1.6\ 10^{-3}$	85	$6.1\ 10^{-3}$	$1.3\ 10^{-3}$	$1.8\ 10^{-3}$	$5\ 10^{-3}$	$1.8\ 10^{-3}$
3	1.99	0.11	19	0.105	0.064	0.13	0.026	0.13
4	2.47	1.35	3	0.822	0.92	1.21	0.32	1.21

a)(1) Development time, (2) mass, (3) body length, (4) tail length.

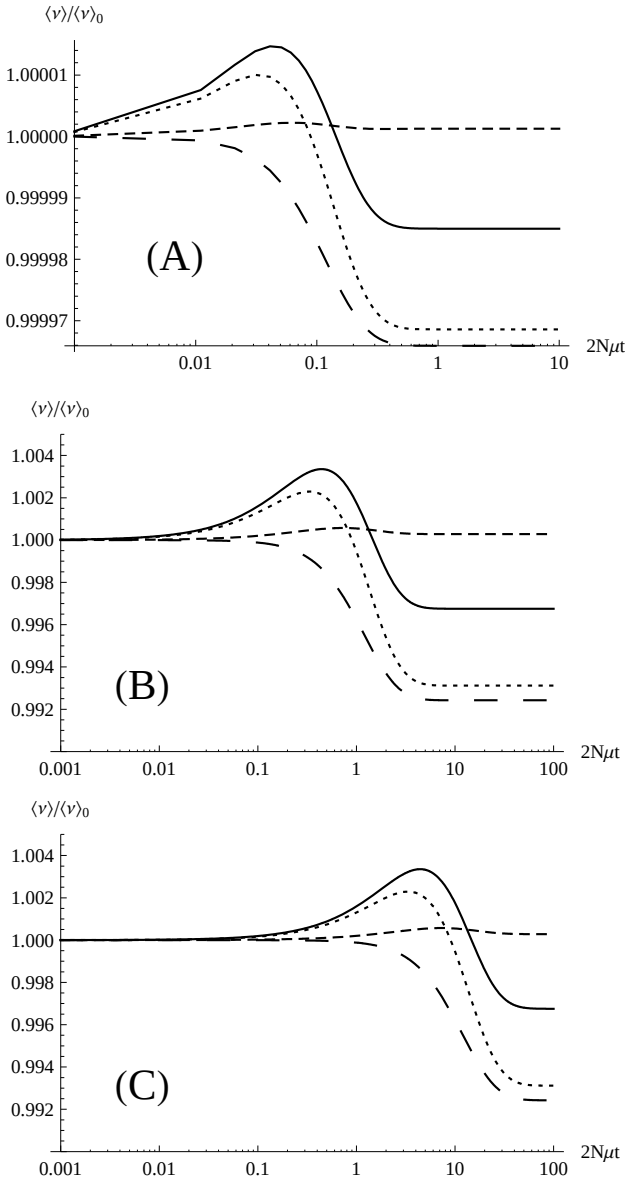
the leading eigenvalues, although some deviations are obvious, specially in the third eigenvalue (see Supplementary Material).

The expectancies of the genetic covariances, according to our theory, show little difference between the two populations, at most 0.8% in one of the genetic variances (Fig. 4.1, middle panel), although the expectancies for the traits are actually different (by construction, since they were fixed in the estimation)

Then, for each estimation, we randomly sampled the distribution of allele frequencies generating 10 populations per pleiotropic structure (4440 choices in total). We assumed a sample size of 300, as in the experimental design (and coincidentally, the population size). The resulting $\mathcal{G}$ matrices are shown in Fig. 4.2.

In this way it was revealed that the differences in $\mathcal{G}$ across the populations might be attributable to sampling effects and drift. The following sections dissect this conclusion according to our logic.

Figure 4.1: (Opposite page) Evolution of genetic variances with distinct mutation rates. (A) $\mu = 10^{-2}$ (B) ($\mu = 10^{-3}$) (C) ($\mu = 10^{-4}$). In all cases, time is scaled as $t = 2N\mu t_g$ where t_g is the time in generations. Solid lines: developmental time; large dashes: mass; short dashes: body length; dots: tail length. Selection is weak, starting from the conditions estimated for the Northern population (Kiruna), $N\vec{\beta} \simeq (-0.11, 0.99, 0.07, 0.004)$, and evolved towards an equilibrium defined by the conditions estimated for the Southern population (Lund) $N\vec{\beta} \simeq (0.025, 14.1, -0.73, 0.037)$. Population size is 300. The dynamics are based on Eq. 4.13 and Barton and de Vladar's (2008) method.


⟨ν⟩/⟨ν⟩0
1.00001
1.00000
0.99999
0.99998
0.99997
(A)
0.01
0.1
1
10
2Nμt
⟨ν⟩/⟨ν⟩0
1.004
1.002
1.000
0.998
0.996
0.994
0.992
(B)
0.001
0.01
0.1
1
10
100
2Nμt
⟨ν⟩/⟨ν⟩0
1.004
1.002
1.000
0.998
0.996
0.994
0.992
(C)
0.001
0.01
0.1
1
10
100
2Nμt

4.5 DISCUSSION

Mutational Variance

Based on neutral microsatellite diversity F_{ST}, Palo et al. (2003) estimated that the mutation rate in *R. temporaria* is $6 \cdot 10^{-3}$ with a population size (estimated form capture-recapture field studies) of 141 individuals, but they also point out that this number might be biased by local migration (Ellegren, 2000). Thus another possibility they discuss is that keeping the same value of F_{ST}, assuming that $\mu = 10^{-3}$ and absence of migration, the size of the populations is of 300 individuals; this is the scenario we have used, since we still did not develop the theory to include migration effects. In any case at the moments it seems that there is no decisive argument to precise neither N nor μ. Estimates for the other mutational scenario awaits for an extension of the theory to migration factors. Nevertheless, since the effects of mutation is an open question, we can give a brief theoretical account on its effects over $\mathcal{G}$.

The expected time to achieve changes in $\mathcal{G}$ increases when diminishing the mutation rates. Assume for the moments that mutation rate is 10^{-4}. One generation in *R. temporaria* takes about 4 yrs, which is equivalent to $t = 0.0015$ (time is scaled as $2N\mu t$). At this rate, about 1700 generations (6.7 millennia) would be needed to reach MSD balance in $\mathcal{G}$ (Fig. 4.1, top pannel). Contrast this result at mutation rates of 10^{-3}, when we should about 17 generations (less than 70 years) of continuous selection are needed to reach a SMD equilibrium in $\mathcal{G}$ (Fig. 4.1, middle pannel). Actually, these mutation rate scenarios (Ellegren, 2000) fit very well to the range of time that Laurila et al. (2002) suggest for the divergence of the two populations, which had to happen during the last 10 millenia.

However, notice that for low mutation rates, the overall change in the genetic covariances is very low. The predictions of $\langle G \rangle$ in both populations differ in at most 0.8%, which is a difference too small to be detected. Nevertheless, phenotypic changes are conspicuous between the traits of both populations of *R. temporaria*. But even when these phenotypic changes have indeed taken place, they are not accompanied by a big change in the covariances. This is what raises our doubts in that selection is the responsible factor for the observed differences in $\mathcal{G}$.

Action of Selection

The original study about the evolution of the $\mathcal{G}$ -matrix in *R. temporaria* reported that the index of quantitative variation (Q_{ST}) deviated significantly from F_{ST}, the index of differentiation at neutral genetic markers (estimated by microsatellite analyses Palo et al., 2003). This shows that selection is acting, and inducing phenotypic diversification. The empirical estimations of the $\mathcal{G}$ -matrices for different treatments and locations were analyzed statistically (Cano et al., 2004) to reveal that the additive genetic co-variances, for most traits, are non-equal between the two locations. This lead to optimism that selection is the cause for such variation.

The distribution of selective gradients $\vec{\beta}$ (estimated from the data, for each pleiotropic structure) between both populations is different, supporting that the diversification between the two population was driven by selection. Yet the expectancies of $\langle G \rangle$ do not show such contrasting differences as the reported empirical $\mathcal{G}$ -matrices. The eigenstructures of $\langle G \rangle$ in both populations are highly similar (Supplementary Material; Table 4.1 and Fig. 4.2). Even though our theory supports that selection has acted to shift the phenotypic values, it reveals has not acted strongly enough to shift the genetic covariances.

If selection for a character proceeds in the opposite direction as in the natural equilibrium conditions that maintain SMD, there can be a transitory increase in the genetic covariances, due to pleiotropic effects. This seems to have happened in the recent diversification of the two populations (Fig. 4.1, middle panel). We presume that this happened during the first two or three centuries after the populations separated, provided that the ecological conditions were such that selection and population size remained, in average, constant.

Effects of Genetic Drift and Random Sampling

The estimations based on our theory indicates that the observed changes in $\mathcal{G}$ are most likely attributable to drift, rather than to mutation or selection. The amount of individuals employed in the experiments allow for significant deviations by genetic drift. Each locus contributed to the variance of drift σ^2 by an amount of $pq(2p-1)^2/2N$, which has a maximum value of $(32N)^{-1}1.210^{-4}$. If we account for all loci, and for a population of size 300, we can have a range of percentile standard error ($\sigma_z/\bar{z}$) from 8% to 220%. Thus the power to discern selection from drift can be rather low.

This should not be confused with the power of the statistical analyses in Cano et al. (2004). Their analyses have enough power to discern differences in the $\mathcal{G}$ -matrix structures, supported by a good experimental design (Lynch and Walsh, 1998). Our argument is that the cause of $\mathcal{G}$'s differentiation is genetic drift.

To assess this possibility, we sampled the distribution of allele frequencies. For each sample of allele frequencies, say $\hat{p}$, the covariance matrix, $\hat{G}$ was calculated. Figure 4.2 (top panel) illustrates that the variation pattern on $\mathcal{G}$. Notice how $\mathcal{G}$ is distributed in a bimodal fashion. Furthermore, the empirical $\mathcal{G}$ for both populations seem to fit well in these distributions.

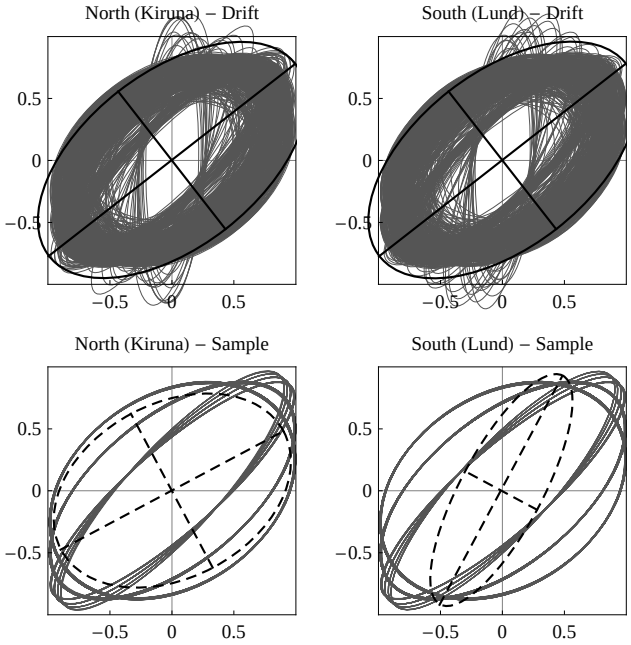


Figure 4.2: Drift and sampling on $\mathcal{G}$ (two principal components: development time and tail length) in two Swedish populations of *Rana temporaria*. Top row: effects of genetic drift on $\mathcal{G}$ evaluated by randomly sampling the allele frequencies (gray ellipses). Black ellipses: expectancies $\langle\mathcal{G}\rangle$. Bottom row: effects of sampling individuals from a particular population (gray ellipses). Dotted black ellipses: 'empirical' $\mathcal{G}$s (Cano et al., 2004). Samples include the 10 realizations of each pleiotropic combination of genetic effects, along with their respective β. Mutation rate is $\mu = 10^{-3}$. Population and sample sizes are 300 for the expectations, and 72 for the drift samples. The representations of $\mathcal{G}$ follow the conventions by Arnold et al. (2008): the semi-axes of the ellipse are eigenvectors of the components, with length $1.96\sqrt{\lambda}$, where λ is the eigenvalue of the component.

But on top of this stochasticity, there is also randomness due to experimental sampling. One population is in itself a sample , $\hat{p}$, of the distribution of allele frequencies. Thus the genetic states of an individual follows a Bernoulli distribution (assuming two alleles at each locus) with certain probability $\hat{p}$. Hence the $\mathcal{G}$ matrix associated to a population, is restricted to a particular realization of the distribution of allele frequencies. Figure 4.2 (lower panel) shows the sampled genetic covariances along with the empirical $\mathcal{G}$'s. It is clear to our eyes, that the differences in $\mathcal{G}$ from both populations are attributable first to drift, and second to sampling, rather than to selection.

How far are we?

The details that affect the evolution of $\mathcal{G}$ are vast. The quantitative trait loci, and the effect that each of these can have over the traits, are in general a real puzzle. We are able to access so little information about the genetic and epigenetic effects, that we are unable to predict how $\mathcal{G}$ will respond in their presence. We have assumed very restrictive conditions, like Hardy-Weinberg equilibrium, which thanks to the appropriate experimental design of the data herein used, can safely be assumed. Also, selection was assumed to be directional and the estimations show it is fairly weak. This allows the possibility that, to some extent, the effects of linkage can be disregarded (Barton and Turelli, 1991; Kirkpatrick et al., 2002). But other factors with potential consequences in our predictions were left out. Namely recombination, epistasis, and dominance, to mention popular ones. Still, the specific model we have herein introduced allowed us to blur much of the information that is experimentally tedious to obtain. The technical details of the method are discussed by Barton and de Vladar (2009), some of which are subtle. But essentially we have shown that we can dispense of many de-

grees of freedom, which were thought to be required to forecast evolution.

The theories on the evolution of $\mathcal{G}$, and its empirical studies have remained separated. We have merged some facets of population and quantitative genetics. We still have some degree of uncertainty, with respect of biological factors. If, for example, we were able to precise N and μ, our estimations would lead to predictions of the $\mathcal{G}$ matrix, which are testable. Thus appropriate experimental design in line to the assumptions of our calculations can properly help to discern what precisely is affecting $\mathcal{G}$'s evolution.

This work is an exercise illustrating that the line of modeling that we are following, that is to study the evolution of the expectancies of the quantities of interest, can be rewarding. This is still to be done for realistic eco-evolutionary scenarios; but so far so good.

Manuscript submitted: H.P. de Vladar and I. Pen – *Evolution of Polygenic Traits: Adaptation at Maximal Entropy.*

Chapter 5

Evolution of Polygenic Traits: Adaptation at Maximal Entropy.

> It is a very sad thing that nowadays there is so little useless information.
>
> Oscar Wilde

Abstract

We review models that aim to predict quantitative evolution in a bottom-up derivation from population genetics theory when the allele frequencies are under the effects of mutation, selection and drift. Up to now the problem remains largely open, since the existing approaches require restrictive assumptions and many approximations to be successful. However, recent works have addressed the problem from a top-down approach. Based on the fact that maximizing an appropriate measure of entropy — constrained to quantitative measurable quantities — recovers the microscopic distribution of allele frequencies, it is possible to predict evolution in a deterministic way, at least for the expectations of quantitative characters. We point out what the simplifications with respect to the other approaches are, and give a comprehensive view of the possible predictions. We explain both aspects of entropy maximization: the technical advantages, as well as their interpretation in the evolutionary process. We highlight some key aspects of this approach, and its relation to fitness landscapes, Fisher's Fundamental Theorem of Natural Selection, and the evolution of the $\mathcal{G}$-matrix of correlated quantitative characters. We also point out some examples, showing the potential of this approach.

5.1 INTRODUCTION

Is genetic variability maintained by selection-mutation balance, or by neutral mutations? Perhaps by pleiotropy? How does drift affects variability? These are typical questions that population genetics (PG) and quantitative genetics (QG) ask. These disciplines describe the evolution at the levels of allele / genotype frequencies and phenotype frequencies, respectively. The mathematical foundation of each is solid (Ewens, 1979; Lynch and Walsh, 1998), and the relationship or equivalence between these two exists, but relies on labyrinthine complicated mathematics. Thus understanding the mechanisms that maintain or induce genetic polymorphisms can give insights to predict and understand phenotypic quantitative evolution. The resulting traits after selection at every time point depend on the current values of genetic variability, whose change cannot be predicted from quantitative measurements alone (Turelli, 1988; Barton and Turelli, 1989).

If infinitely many loci and/or infinitely many alleles with differential effects segregate, genetic variance remains unchanging in time, which is enough to account for short term predictions even in multivariate traits (Lande, 1979), but not for long term predictions (Turelli, 1984; Barton and Turelli, 1989). Alternatively, allowing for notable effects of each mutation induces asymmetry (e.g. non-normality) in the distribution of allele frequencies, which in turn induces a change in genetic variance (Kingman, 1978; Turelli, 1984; Barton and Turelli, 1987). Provided that new alleles remain rare in the population good approximations for the change in genetic variance can be found (Kingman, 1978; Bulmer, 1980; Turelli, 1984; Barton, 1986; Barton and Turelli, 1987).

The general theory to predict quantitative evolution solely in terms of measurable metric characters has been relying on the mapping of the allele frequencies to *moments* (Barton and Turelli, 1987; Frank and Slatkin, 1990; Bürger, 1991) or *cumulants* (Bürger, 1991, 1993; Rattray and Shapiro, 2001). Although elegant mathematically, the applicability of the results is highly shadowed by the fact that the space of the trait involves infinite moments or cumulants. Even if we have criteria to choose some of them, in general the equations still depend on the allele frequencies (Barton and Turelli, 1987; Bürger, 2000).

For practical reasons, this is too cumbersome to be useful, because of technical limitations on genetic measurements. DNA can be screened to identify which regions, genes, or generally speaking which alleles can have effects on different traits. But these polygenic states are rarely screened in the whole population or through time, at least with enough accuracy to comply with predictions from PG. We typically do (or can) not know in detail how many loci contribute to an evolving trait (quantitative trait loci, QTL), except for the few that contribute with a large effect (Barton and Keightley, 2002; Roff, 2007).

Other lines of work have studied different theoretical aspects of the distribution of allele frequencies. Namely, Iwasa (1988) found that measuring entropy with respect to the equilibrium frequency distribution, leads to a non-negative increase of this quantity. Yet other work (Prugel-Bennett and Shapiro, 1994, 1997; Rogers and Prugel-Bennett, 2000) truncated the system of trait moments, and maximized an entropy measure to account for the remaining information not included in the truncated dynamics. This had notable success. Other articles have recently reported an analogous structure of the coupling of PG and QG with statistical mechanics in the physical sciences (Sella and Hirsh, 2005; Ao, 2005, 2008; Saakian et al., 2008; Barton and de Vladar, 2009; Barton and Coe, 2009). Par-

ticularly, this approach has been successful in eliminating the explicit dependence on the allele frequencies, and the need to truncate arbitrarily the *macroscopic* space (mean trait, genetic variance, etc, thus we could say that it is successful in explaining and complementing some aspects of QG theory. This raises the hope that we can actually make long term predictions of the evolutionary dynamics, still considering the 'microscopic' factors (allele frequencies and their effects over the trait), but making minimal use of this information.

But how much evolutionary chance can we predict without directly addressing microscopic information? This is the main question with which we will deal in this article, and to which entropy maximization gives some solutions. Although previous works on this subject were mostly of a theoretical nature, to show the potential of the approach we apply it to selected examples, such as the Fundamental Theorem of Natural Selection (Fisher, 1930, 1958), Wright's adaptive Landscapes (Wright, 1967, 1988), Quasispecies, and the evolution of the covariance among characters (the $\mathcal{G}$-matrix).

5.2 MECHANISMS OF QUANTITATIVE EVOLUTION: BOTTOM-UP THEORIES

If the polygenic basis of a trait would consist of infinite alleles of infinitesimal Gaussian effect (Fisher, 1918; Kimura, 1965a) the genetic variance would remain constant under the action of selection, and at equilibrium the quantitative trait would be normally distributed (Kimura, 1965a). This is the basis for Lande; Lande's (1979; 1980) extension to the evolution of multivariate traits, and a common assumption in quantitative genetics. In general, selection induces asymmetry in the distribution of allele frequencies, followed by a change in genetic variance, which

is widened by frequent mutations of finite effect (the "house of cards model", Kingman, 1978; Turelli, 1984). But mutations are typically infrequent (rate of about 10^{-3} per locus per generation) thus we expect that the fittest alleles fixate at equilibrium, depleting all standing genetic variation (Fisher, 1930; Bulmer, 1971; Kingman, 1978; Turelli, 1984). Interestingly, the predictions for finite numbers of alleles, say two (Wright, 1935; Bulmer, 1972), three or more (Turelli, 1984) are the same as for infinite alleles; hence the genetic variance is independent of the number of segregating alleles, as long as they remain rare in the population (Turelli, 1984; Barton, 1986).

Despite the fact that the above models are valid only close to fixation, they are not suitable for a quantitative approach, since only the allele frequencies at every locus, denoted by the vector $\vec{p}$, evolve. A possible solution would be to track the genotypes, which is actually convenient for a multi-allelic situation, since the problem can be reduced to a 2-allele equivalent (Szathmary, 1993). The problem is that counting the number of genotypes for many loci involves even more variables than the allele frequencies. The desirable alternative is to make a change of variables $\vec{p} \rightarrow \vec{A}$ where $\vec{A}$ is a vector of (only) macroscopic variables. It turns out that $\vec{A}$ consists of the moments of the distribution of the trait. This change of variables has been characterized to describe the evolution of the macroscopics (Barton and Turelli, 1987):

$$\frac{d\vec{A}}{dt} = \mathbf{B}.\tilde{\beta} \tag{5.1}$$

where **B** is the covariance matrix (or Jacobian matrix) of $\vec{A}$ and $\vec{p}$ (typically non-trivial moments of the distribution of $\vec{p}$ and allelic effects), and $\vec{\beta} = \partial_{\vec{A}} \log(\bar{W})$, the (multivariate) gradient of the log-mean fitness. But there are two unfortunate hindrances with Eq. 5.1. First, a standard result of statistics (Karlin and Taylor, 1975, Ch. 1) is that to represent a probability distribu-

tion in terms of moments or cumulants, we need to specify all of them, that except for special cases the number is infinite. Thus $\vec{A}$(and hence **B** and $\vec{\beta}$) are of infinite dimension. The second hindrance is that **B** still depends on the genetic frequencies. To approach Eq. 5.1 in a tractable way two approximations can be made. For the first, we can decide which of the predictors in $\vec{A}$ are the most important, and neglect the rest (Barton and Turelli, 1987; Bürger, 1993; Rattray and Shapiro, 2001). The second is to assume an underlying distribution of $\vec{p}$ to explicitly calculate the terms in **B**. Then the dimensionality of the system is reduced and the allele frequencies are averaged-out.

As a simple, first example approximate $\vec{A} = \bar{z}$, then **B**= ν, the additive genetic variance. This is basically results in a version of the breeder's equation. Note that the breeder's equation in tis canonical form, $\Delta\bar{z} = h^2 s$, employs heritability $h^2 = \nu/\sigma^2$, and the difference Δz measures changes from the parental traits, resulting from directional selection with intensity s. The difference that we are speaking about is $\Delta\bar{z} = \nu\beta$ measures the change on the mean trait between generations, and the selective gradient β is the correlation between fitness and the trait. Both differences, Δz and $\Delta\bar{z}$ measure the average response of the traits to selection, the meaning is slightly different.

Likewise, assuming that $\vec{A}$ = ($\bar{z}$,ν) we obtain in addition that $\Delta\nu = \nu/n_{\text{eff}}$ where n_{eff} is the effective number of loci (Barton and Turelli, 1987). Essentially the last equation is Bulmer's (1971) equation for a Gaussian model, which unfortunately underestimates the change in ν (Barton and Turelli, 1987; Bürger, 1993), and requires an estimate of n_{eff} that is both experimentally unavailable and changing in time.

Keeping the first three terms, $\vec{A} = (\bar{z}, \nu, m_3)$ (where m_3 stands for the third central moment of $\vec{p}$), *and* assuming the house of cards model, the evolutionary dynamics are to some extent well predicted when many alleles are included, approximating the

situation that each mutation leads to a rare allele.

Instead of employing the moments of the trait as macroscopics, the cumulants have been also proposed (Bürger, 1991, 1993; Rattray and Shapiro, 2001) as descriptors. But this description suffers from the same pathologies. For directional selection, a house of cards model with an initial distribution of allele frequencies, using four cumulants (mean, variance, skewness and kurtosis of z) provide a fairly good agreement on long term predictions when mutation rate is low enough ($< 10^{-2}$), the allelic effects are small and thus the initial variance is also small. Otherwise, the long term predictions are not reliable (Bürger, 1991, 1993).

Perturbation analysis (analysis of a system in terms of another one that is closely related) over neutrality can lead to good approximations for weak directional selection. Furthermore, low mutation rates allow approximate closed solutions, if on top it is assumed that the higher moments reach equilibrium much faster than the lower ones (Rattray and Shapiro, 2001). Whether this is also true for other selection schemes like stabilizing, unequal effects, epistasis and linkage disequilibrium, we still don't know (Bürger, 2000).

These mechanistic theories have given some insight into how the micro-macroscopic levels are related, but the theory as a whole has remained wide open in predicting, or fully explaining QG theory. In particular, how the additive genetic variance is changing in time.

5.3 ENTROPY MAXIMIZING THEORY: TOP-DOWN APPROACH

Iwasa (1988) studied some population and quantitative genetical problems from a rather uncommon approach for the field. He calculated the rate of change of the entropy of the distribution of allele frequencies, ψ, motivated and justified only by its use in statistical physics as the H-Theorem (see Reif, 1965, Appendix A.12, pp. 624-625). As presented by Iwasa, entropy measures the proximity of $\psi(\vec{p})$ with respect to its equilibrium state. At maximal entropy, $\psi(\vec{p})$ would correspond to the one dictated by equilibrium between selection, mutation, and drift (SMD) balance.

Aita and collaborators (Aita and Husimi, 1998, 2003; Aita et al., 2004, 2005) chose a similar way of studying evolution under directional selection and mutation. They also have shown how entropy increases during a fitness hill-climbing process as evolution leads to its maximum, thus employing it to monitor the optimization process rather than directly the microscopic variables.

In an analogy to this kind optimization procedures Prügel-Bennet and Shapiro (1994; see also the further works of Prugel-Bennett and Shapiro, 1997; Rogers and Prugel-Bennett, 2000) studied the evolution of the mean fitness of a population with polygenic basis, under the influence of directional selection and mutation. Naturally they were confronted with the problem of infinite recursion of the moments of mean fitness.

Thus they decided to track only mean fitness and fitness correlation (proportional to the additive genetic variance), and assumed that all remaining moments would follow a distribution that maximizes an entropy measure. The method proved successful in the sense of describing the quantitative evolution (mean fitness and its correlation in his case) without directly addressing the microscopic variables. Related works applied this methodology for variants of the problem (although focused on genetic algorithm dynamics Rogers and Prugel-Bennett, 2000) also including the effects of recombination (Prugel-Bennett and Shapiro, 1994).

These works introduced a very interesting approach, but most of them neither considered central biological questions, nor paid too much attention to the biological assumptions. The price for this is actually that most of these models are inconsistent with population genetics theory, and therefore the predictions are not entirely reliable for the purpose of quantitative geneticists.

SELECTION-MUTATION-DRIFT EQUILIBRIUM

Given the evidence that entropy, rather than fitness (though see Metz et al., 2008, and also Appendix A) is maximized at equilibrium, and that it allows to close the moments and average over the microscopic variables, Barton and de Vladar (2009) reconsidered these ideas in a more biologically-consistent framework. The entropy measure S is (Barton and de Vladar, 2009; Barton and Coe, 2009):

$$S[\psi_{(\vec{p})}] := -\frac{1}{2N} \int \psi_{(\vec{p})} \log \left(\frac{\psi_{(\vec{p})}}{\phi_{(\vec{p})}} \right) d^n \vec{p} \tag{5.2}$$

What makes this entropy measure relevant for population genetics is the prior distribution ϕ. S would be maximum only

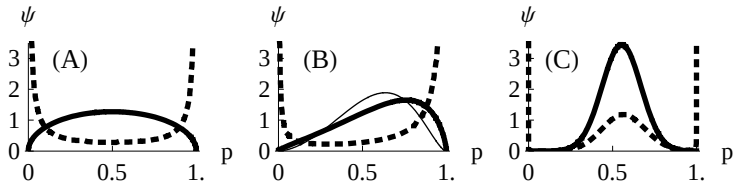


Figure 5.1: Distribution of the favorable allele under: (A) No selection, (B) directional selection, (C) stabilizing selection. Dotted lines: distribution under the effects of drift; solid curves: distribution under the effects of drift and mutation. Notice that in the absence of mutation, some of the alleles fixates, irrespective of the action and type of selection. Therefore the dotted curve on the left panel is the 'base' distribution of the drift component. These distributions are the same if we calculate by maximum entropy, or by diffusion approximation. The thin line in the middle panel, corresponds to Prügel-Bennett's (1997) assumptions (uniform prior in the max-entropic measure).

when$\psi_{(\vec{p})} = \phi_{(\vec{p})}$; in physics, as well as in Prügel-Bennet's work, ϕ is a constant and thus any state is equally probable. We know from population genetics that this is not the case, since in the absence of other evolutionary factors, drift drives the alleles to fixed states $p_\ell = \{0, 1\}$ at all of the n loci (Crow and Kimura, 1970, pp. 327-329). This distribution, following Kimura (1955, pp. 147 Eq. 9) is $\phi_{(\vec{p})} = \prod_\ell^n \left[p(1-p)\right]^{-1}$ (Fig. 5.1a).

Other evolutionary effects can be included if in the maximization of S we include certain variables $\vec{A}$ (termed observables, in the same sense as Prugel-Bennett (1997) see also (Barton and de Vladar, 2009; Barton and Coe, 2009). Then Eq. 5.2 leads to

$$\psi_{(\vec{p})} = \frac{\phi_{(\vec{p})}}{\mathbb{Z}} e^{\vec{\alpha}.\vec{A}} \,. \tag{5.3}$$

The quantity $\mathbb{Z} = \int \phi e^{\vec{\alpha}.\vec{A}} d^n\vec{p}$ normalizes the distribution, and the vector of parameters $\vec{\alpha}$ determines the evolutionary processes maintaining equilibrium between the macroscopics $\vec{A}$;

if all these constants are zero, then the distribution is dictated by genetic drift, as we discussed above.

Barton and Coe (2009), setting $\alpha_0 A_0 = \log(\bar{W})$, recovered Sella and Hirsh's (2005) results for low mutation rates and drift. If selection is directional over an additive mean trait (aka Sella & Hirsch's additive fitness), then α_0 would be the selective gradient, and $A_0 = \bar{z}$ (see also Ao, 2005, 2008). This reduces the distribution of allele frequencies to the distribution of fixated genotypes, and the conditions of selection-drift equilibrium reveal that mutations need not to be nearly neutral in order to maintain a constant substitution rate. Sella and Hirsh (2005) rediscovered one of the special cases of Iwasa's (1988) *free fitness*, that is a function which balances "the evolutionary tendencies in finite populations to increase both fitness and entropy" (quoting from Sella and Hirsh, 2005). Barton and Coe (2009) complemented the theory finding that the macroscopic that describes processes at high mutation rates and which dovetails with the Wright-Fisher model would be $U = 2\sum_\ell \log[p_\ell(1-p_\ell)]$, that is a log-measure of heterozygosity, and its conjugate process is mutation, quantified by the rate μ. In this case, the free fitness would be exactly that of Iwasa (1988). We will return to the free fitness when discussing Fisher's fundamental theorem of natural selection.

Maximizing a genuine entropy measure avoids an arbitrary moments truncation, but how does it 'hide' the microscopic variables? The quantity $\mathbb{Z}$averages over the allele frequencies $\vec{p}$, and depends explicitly on population size N and $\vec{\alpha}$, which define the factors maintaining SMD equilibrium. In fact derivatives of the form $(\partial/\partial\alpha_i)\log(\mathbb{Z}) = 2N\langle A_i\rangle$ yield the statistical expectations of the macroscopic variables, and the cross derivatives the covariances among the macroscopics, $(\partial^2/\partial\alpha_i\partial\alpha_j)\log(\mathbb{Z}) = 4N^2\mathrm{cov}(A_i, A_j)$.

These expectations do not depend on the allele frequencies,

but like $\mathbb{Z}$ they depend on N and α which are in principle experimentally measurable.

Defining different macroscopics will describe different selection schemes. Consider three important examples. (i) If the only observable is $\vec{A} = U$, we will recover neutrality: mutation-drift equilibrium. (ii) If the only observable is $\vec{A} = \bar{z}$ then the resulting distribution is that of directional selection (Sella and Hirsh, 2005; Barton and Coe, 2009), and if we define $\vec{A}$ = $\{\bar{z}, U\}$ then it is directional SMD (Iwasa, 1988; Barton and de Vladar, 2009). (iii) If we include $\vec{A}$ = $\{\bar{z}$, var($\bar{z}$), $\nu\}$ then we will obtain stabilizing selection, and further inclusion of U will add mutation to the evolutionary scheme (Fig. 5.1).

Furthermore, to have concrete numbers for the estimators, we can assume a genetic architecture, e.g. the way in which the genotype maps to the phenotype. Works on diallelic loci have, for simplicity, often assumed equal additive effects (Bulmer, 1972; Bürger, 1991, 1993; Prugel-Bennett and Shapiro, 1994, 1997; Rattray and Shapiro, 2001; Saakian et al., 2008). Barton and de Vladar (2009) relaxed the equal effects assumption and provided explicit solutions for directional and stabilizing SMD of an additive trait, as well as the framework for dealing with certain classes of epistasis. Although the predictors do not depend on the allelic frequencies, they still depend on the number of loci and their effects. But it is striking that the system is not only well defined by just a few macroscopics, $\vec{A}$, but also that they define the whole microscopic distribution, even when the microscopic degrees of freedom is much bigger than the macroscopic degrees of freedom.

It is notable that we can recover and predict results for SMD by direct calculation from the function $\mathbb{Z}$. A pertinent example is the maintained equilibrium genetic variance. At high mutation rates, the expected genetic variance maintained by MSD would be

$$\langle \nu \rangle = \langle \bar{z} \rangle \frac{2N\mu}{N\beta} + \mathrm{cov}(\nu, \bar{z}) \ ,$$

under stabilizing selection (if selection is directional, the covariance vanishes). This is the same prediction that arises form the Wright-Fisher model. The assumptions behind, are not too restrictive; it allows for arbitrary number of loci of distinct effects, as long as there is no pleiotropy, or linkage disequilibrium.

EVOLUTIONARY DYNAMICS

Since the descriptors that are needed to track evolution are well defined, the evolution equations would be resumed by the rate of change of these variables, instead of infinite moments or cumulants, thus eliminating the need to arbitrarily truncate the macroscopic space. But there is more to learn from the previous approaches. If the distribution of allele frequencies is initially concentrated (like a Gaussian, Gamma, or at an equilibrium maintained by and initial MSD balance that later changes), even under random drift, its path towards the new equilibrium is not arbitrary, but evolves as a travelling wave (Bürger, 1993; Rattray and Shapiro, 2001; Rouzine et al., 2003, 2007), smoothly morphing from the initial distribution to the final one, that is dictated by the max-entropic MSD equilibrium. Actually, a good approximation for the dynamics results from assuming "local equilibrium" which means that out of equilibrium the entropy is still maximized, constrained to the same observables as in equilibrium, but with virtual selective value and mutation rate such that they match the expectancies. These virtual variables are not necessarily the actual selective gradient or mutation rate, but they are forged and changing in time in order to keep (a) entropy maximized, and (b) the observables at their genuine macroscopic values at all times (Barton and de Vladar,

2009). Only at equilibrium these virtual variables match with the real selective gradient and mutation rate.

This would work as long as selection is stabilizing, or is directional with mutation rates that are high or very low. The distribution of allele frequencies will remain far from the borders of fixation (p_l = 0, 1) when $4N\mu > 1$, and the changes towards the new equilibrium are smooth. Similarly, when $4N\mu \sim 0$, selective sweeps induce changes in the proportion of fixed states in a smooth way. In both cases the max-entropic distribution remains accurate. However in the middle regime, close to the critical rate $\mu_c = 1/4N$, both effects are present, but the time scale for selective sweeps to occur is lower than the time for a change due to standing genetic variation. The max-entropic approach fails to recover the evolutionary paths either of the microscopic or the macroscopic variables, since local equilibrium is disturbed (Barton and de Vladar, 2009). Away from μ_c every given locus and its copy evolve independently, but close to μ_c these microscopic changes are correlated. Although the equilibria will be well described by the two observables $\bar{z}$ and U, the dynamics will require an extra degree of freedom that accounts for that correlation, $\bar{k}$. This correlation is proportional to the excess of the genetic variance $\bar{k} = 2(\nu_{\max} - \nu)$, where $\nu_{\max}$ is the maximal possible genetic variance. The evolution of this quantity happens to follow from the dominance of the allele that is being selected, with strength η over fitness. The dominance, set initially to $\eta = 0$ will transiently evolve, and vanish again when MSD is consummated. This means that at these intermediate mutation rates, dominance will effectively appear and affect the relative fitness of the alleles, even when the effect is not present *ad hoc*.

5.3.1 Adaptive landscapes and adaptive potential

(Fisher, 1941) never accepted Wright's ideas about fitness landscapes (Wright, 1967, 1988). He found the surfaces of selective value an artificial construct appearing only in very specific cases. Artificial as it might be (Provine, 1986), it has been a paradigm to think about evolution, and yet as a generally applicable concept it remains unjustified. But it is a good aid to understand the connection between the distribution of microscopic with the macroscopic descriptors (Arnold et al., 2001).

The max-entropic approach can contribute with a concept: a landscape is induced rather than assumed. It is based on quantities that are measurable that lead to an adaptive landscape, the golden child of Wright's conceptualism of the evolutionary process.

Contrary to the fitness landscape, the max-entropic induced landscape is the actual potential for evolutionary outcomes in the presence of mutation and drift effects, and not just naïve expectations deduced only from fitness arguments. As seen in Fig. 5.2a, a landscape in the allele frequencies indicates that the optimal region in equilibrium is not fixation of the fittest alleles (since there is also mutation and drift). Furthermore, we are able to reconstruct the landscape in terms of the macroscopic variables. Examples with several numbers of loci of either equal or unequal effects are given. Notice in particular that with some distributions of effects only those that contribute the most can be enough to reconstruct the phenotypic landscape in good approximation (Fig. 5.2b).

Furthermore, we can predict for a given value of the mean traits (or for any other quantitative measurement) the distributions of allele frequencies at a locus. This is a posterior distribution, and it differs from that of the one-locus distri-

Figure 5.2: Different representations of the adaptive landscape (in the max-entropic sense, see text). (A) Adaptive landscape parameterized in terms of allele frequencies (for two loci of different effect). Lighter shades represent higher fitness, and darker shades low fitness. The curves are fitness isoclines. (B) Fitness landscape parameterized in terms of the quantitative character, for distinct number of loci n (solid lines) and (dashed line) for 15 loci of exponentially distributed (mean=2) random (but fixed) allelic effect. (C) Marginal distribution of allele frequency for the major contributing locus, subject to a given value of the trait (solid line) and the prior expectation for the same allele (dashed line) [not entirely clear]. All plots constructed with $N\beta = 0.5$, $N\mu = 1$.

bution (Fig. 5.2c) that results when no further information is defined (in analogy to Bayesian estimation of distributions, see e.g. Shoemaker et al., 1999). This difference appears because the allele frequencies are not entirely free to wander in the genotype space. In analogy with Wright's shifting balance process (Wright, 1932) where selection would lead to maximal fitness, if drift introduces a deviation in the population's allele frequencies, selection and mutation would induce a net response of the allele frequencies in the direction that maximises entropy. Thus the allele frequencies are all coupled and obliged to fit the observables that maximize entropy. This reasoning in terms of entropy might sound abstract but actually it is very consistent, because maximal entropy is a macroscopic measure of the equilibrium between SMD.

5.3.2 Fisher's fundamental theorem of natural selection in a max-entropic context

Fisher's Fundamental Theorem of Natural Selection gives the rate of change of mean fitness that is due to selection (Fisher, 1930). Although the conditions under which the theorem holds are broad, Fisher's derivation (1930) of his theorem is obscured by his cryptic explanations, although later authors have elucidated what the author meant (Frank and Slatkin, 1992; Edwards, 1994; Fisher, 1999; Edwards, 2002; Grafen, 2003). The FTNS states that the change in mean fitness of a population due to selection is proportional to the variance in fitness. Mathematically, we write this as $\Delta\bar{W} = \mathrm{var}(W)/\bar{W}$; since the variance is always positive, then mean fitness always increases.

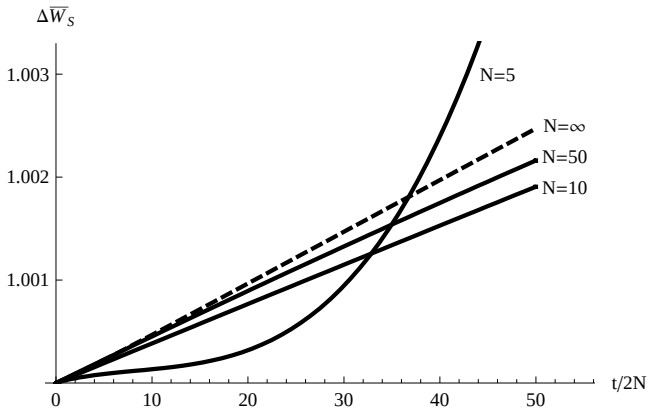


Figure 5.3: The change in the expectance of mean fitness, as predicted by Fisher's fundamental theorem (dashed line, red on-line), the predictions including finite size-effects for distinct population sizes N only showing the component due to selection (solid lines).

This equation is very general and the scope of its applications reaches more than PG and QG (for some generalizations, applications and extensions see Ewens, 1969, 1976; León and Charlesworth, 1976, 1978; Frank, 1997; Vlad et al., 2005).

In finite populations drift will also influence the genetic variance, producing an indirect change in $\bar{W}$, such that in a particular stochastic realization it may actually decrease (Nagylaki, 1993). But we can compute Fisher's principle as statistical expectation on the universe of possible realizations. A general equation for the change in mean fitness is possible, but we consider now only directional selection over the fitness. We find, applying the max-entropic calculations, that:

$$\frac{d_s}{dt}\langle\bar{W}\rangle = \beta^2\left(1+\frac{1}{2N}\right)\langle\nu\bar{W}\rangle \geq 0 \tag{5.4}$$

where we denoted by d_s the change in fitness due only to the se-

lection component, as in Fisher's interpretation. The term $1/2N$ arises from the effect of drift, and can be obviated if we want to be sharper with Fisher's statement. (We choose to include it since it always appears, extending the theorem to the effect drift: the change in expected mean fitness due to selection and drift, will always increase.) The expected mean fitness changes proportionally to the expected covariant change between the genetic variance and the mean fitness, but this expectation is always positive, contrary to the individual realizations. (It should not to be confused with Price's (1970) equation, whose covariance term would be in level of averaging within a population; the expectancy that we speak here is with respect to the universe of possibilities of population realizations). The evolutionary components that don't act directly on a trait (or fitness) still deform the adaptive landscape (or in general, potential), and have an indirect effect over the genetic variance.

We evaluate Eq. 5.4 as well as the deterministic version (Fig. 5.3). The max-entropic approach preserves the FTNS in a statistical sense, but the quantitative trajectories are actually different: they are initially delayed, and accelerate at latter times. Still as $N \to \infty$ we would recover the deterministic version. This is because the fluctuations by drift have variance proportional to $(2N)^{-1}$, which vanishes for big population sizes, and the evolutionary trajectories become entirely deterministic.

5.3.3 The $\mathcal{G}$ Matrix and the Evolution of Correlated Characters

As a last example, we consider a more practical and and useful model: the correlated response to multivariate traits. A classical way to handle this problem, is assuming that the evolution

of a vector of trait means $\vec{z}$, can be described by $d\vec{z}/dt = \mathcal{G}.\mathbf{P}^{-1}.\vec{\beta}$, where $\mathcal{G}$ is the genetic variance-covariance matrix, and $\mathbf{P}$ is the matrix of environmental variances (Lande, 1979). Consequently quantitative genetics pays huge attention to the evolution of the $\mathcal{G}$ matrix (Steppan et al., 2002; Blows, 2007; Kotiaho, 2007). If it were constant, there would be no major predictive problem (Lande, 1979). However, we know from theoretical studies (Lande, 1980; Turelli, 1988; Lynch and Walsh, 1998), and from empirical estimates (Wilkinson et al., 1990; Paulsen, 1996; Phillips et al., 2001; Cano et al., 2004; Ovaskainen et al., 2008), that $\mathcal{G}$ changes. Still, there is no consensus: the change in $\mathcal{G}$ might be slow, and effectively constant (Björklund, 2004), or fast (Cano et al., 2004; Doroszuk et al., 2008). Depending on the selection intensity and the genetic structure of the traits (e.g. allelic effects), pleiotropic effects (Lande, 1980; Barton, 1990; Keightley and Hill, 1990; Slatkin and Frank, 1990; Kondrashov and Turelli, 1992; Turelli and Barton, 2004; Tanaka, 2005; Jones et al., 2007; Albert et al., 2008), genetic drift (Roff, 2000; Jones et al., 2003, 2004), migration (Guillaume and Whitlock, 2007), etc, $\mathcal{G}$ will effectively change across generations in an unpredictable way. Even after identifying experimentally that the genetic covariances have changed, estimating the rate of these changes is still an experimental and theoretical challenge.

Since $\mathcal{G}$ is basically a set of macroscopic descriptors, and is influenced by the microscopic hidden variables, we can estimate the max-entropic expectation for $\langle\mathcal{G}\rangle$. We will only highlight some results on the subject, and an application. The complete analysis is a subject on itself and will be published elsewhere.

When mutation is small and compensated by drift ($4N\mu \ll 1$) most loci will have an allele very near to fixation. A change in selection will proceed only by standing genetic vari-

ation, that is proportional to mutation rate. This predicts that there are many generations of latency before a change in genetic variance is observed. After selection acts, the eigenvalues would be severely shrunk. Conversely, the change of $\mathcal{G}$ when mutation rate is big ($4N\mu > 1$) will proceed by both, standing genetic variation and newly produced mutational variation. But the observable that quantifies mutation, U , also needs to be considered to take into account how $\mathcal{G}$ responds. This means that the rate of change of genetic variability U, needs to be included in and extended version of $\mathcal{G}$. In doing so, the latency periods are much shorter, but $\mathcal{G}$'s eigenvalues suffer much less change. Thus the max-entropic theory advises how to include the effects of mutation.

As a practical example, in the previous chapter we compared the predicted values of the $\mathcal{G}$ matrix with the experimental data for *Rana temporaria* reported in Cano et al. (2004). We made rough approximations of average effect, effective number of loci and pleiotropic structure, summarized in Table 4.1. The mutation rate μ is estimated to be on the order of 10^{-3} (Palo et al., 2003). Thus we proceeded to fit the selective values that explain the measurements of the traits, according to the theoretical predictions (Cano et al., 2004, with n=300, that is the sample size from the real population, but at the same time, the size of the experimental population). From these fits the $\mathcal{G}$-matrices were predicted before and after selection (see chapters 4 and 6). The deviations from the empirical $\mathcal{G}$ are not big (Table 4.1) and entirely attributable to drift. The orientations (eigenvectors) of the theoretical expectancies are in good agreement, before an after selection (Fig. 4.2).

We are assuming that factors like epistasis and dominance are negligible. Yet the estimations are still satisfactory. Much can be improved, but for our purpose -showing the advantages of the our approximation- the example is just adequate.

5.4 CONCLUDING REMARKS

To what extent can we predict evolution? The answer to this question is hidden in both the degrees of freedom at different evolutionary levels, and on the mathematical complexity that relates these levels.

While PG considers loci as unit of selection, QG considers the quantitative characters as units of selection. Neither of them is more fundamental; they conform to two different coordinate systems to model evolution (Barton and Turelli, 1987), each one is blazing trails to different aspects of the evolutionary process (Orr, 2005). The max-entropic tool clarifies certain issues, showing how distinct aspects of PG and QG that were considered dead ends match together.

The conceptual convergence by different groups to the max-entropic and statistical mechanics-like approaches (Iwasa, 1988; Prugel-Bennett and Shapiro, 1997; Sella and Hirsh, 2005; Ao, 2005, 2008; Saakian et al., 2008; Barton and de Vladar, 2009; Barton and Coe, 2009, as well as several other works) indicates that it can play an important role in understanding evolution.

We have explained and exemplified for evolution under SMD that we can drastically reduce the amount of information that we need to infer microscopic aspects of an evolving polygenic trait from quantitative measurements. But also the other way around: we may use minimal information of the microscopic variables to make quantitative predictions. This raises the optimism with respect to empirical counterparts, as for example, the QTL which are able to give only a rough idea of factors responsible for genetic variation. Indeed this information might actually be enough to have a richer predictive power, contrary to what is expected by the bottom-up models, which suggest the need for detailed genetic properties. This is great news, since we don't really need to know the whole degrees of free-

dom of the hidden variables, as required by other bottom-up approaches and which are subject to huge limitations.

We have dealt however, with very specific and convenient conditions. Among others, linkage equilibrium, absence of epistasis and diallelic sub-dominant multiple loci. Indeed, including these other factors would change the predictions that can be drawn from the model. But the approach in itself does not change.

The effects of recombination have been studied in a simple model of additive and equal effects (Prugel-Bennett and Shapiro, 1994; Saakian et al., 2008), from which we can also extend the analyses. We know that although selection might induce linkage among recombining loci (Bulmer, 1971) weak selection justifies a quasi-linkage equilibrium at all times, partially uncoupling of the loci in a tractable way (Kirkpatrick et al., 2002). The way to approach epistasis, has also been sketched by Barton and de Vladar (2009), although explicit solutions depend on specific architectures of the genotype-phenotype map, about which we currently know little (Hansen, 2006). The extensions to multiple alleles is actually straight-forward, only introducing a higher dimension at each locus in the genotype space. Intuitively, the consequences are not expected to be big. Under selection and drift with low mutation rates, only one allele is expected to be maintained at each locus, if selection is directional, and two if it is stabilizing (Turelli, 1984; Barton, 1986; Bürger and Gimelfarb, 2004; Schneider, 2006), a result that of course is not necessarily valid at high mutation rates. But for practical purposes there might be little need to introduce more alleles at every locus. Besides, under low mutation rates, the predictions for maintenance of genetic variance are insensitive to the number of alleles that are segregating at each locus. Also, extension to dominant effects can easily be included when considering separate effects of each copy of a locus in the diploid

individual from which we can also extend the analyses.

Notably, all these changes will affect the results and predictions, but do not interfere with the philosophy of the top-down approach. These other factors, define the way in which fitness affect the microscopic space, and evidently it affect the way allele frequencies evolve. Although these are functional constraints that determine the patterns of evolution of the phenotypes, they do not change the fact how selection and drift are acting over the macroscopic space. On these lines, we might point that several examples that we brought assume weak directional selection. Still there is consensus -mainly theoretical arguments- that stabilizing selection is among the main forces maintaining diversity in nature (Charlesworth et al., 1982). For example extreme phenotypes are more likely to produce deleterious mutations. Stabilizing selection has also been formulated and worked out for single polygenic traits with unequal effects (Barton and de Vladar, 2009). However, recent meta-analyses over empirical estimations have found that most quantitative traits are maintained by weak directional selection (Hoekstra et al., 2001; Kingsolver et al., 2001). This adds to our raised optimism about the possibility of predicting evolution, and allows waiving many complications. The max-entropic theory does not solve them all, but is a step further in our ability to predict evolutionary change in the long term.

Chapter 6

Perspectives: Pleiotropic effects on the $\mathcal{G}$-matrix

6.1 STANDING AND MUTATIONAL VARIATION

The amount of variation maintained in populations is the source of evolutionary change as a response to natural or artificial selection. When per-locus mutation rates are low - as is typically the case- the predicted amount of genetic variance that is maintained is also low, at least to account for the levels observed in natural populations (Turelli, 1984; Barton and Turelli, 1989). However increased genetic variance can be maintained even at low mutation rates if the effects of selection is pleiotropic over various traits (Keightley and Hill, 1990; Barton, 1990; Kondrashov and Turelli, 1992; Zhang and Hill, 2005). In such a case, the pattern and rates of evolution of the characters is compromised to the joint effects of multivariate selection at every given locus. At which rate are the genetic (co)variances, the G matrix, changing? Naturally that depends on (i) the amount of loci, (ii) their effects, and (iii) the pleiotropic constitution of the traits, or in general, the genetic architecture including epistatic effects. If we were to account for directional selection mutation drift (SMD) balance under pleiotropy, even with a simple bi-allelic system of non-interacting loci (i.e. linkage and Hardy-Weinberg equilibrium), at least we need to recur to a distribution of effects that weights few loci of big effect and many of small effect, consistent with QTL observations (Otto and Jones, 2000), since the genetic attributes are not entirelly negligable.

Fisher proposed a model, known as the geometric model, to argue that most mutations affecting the phenotypic space have a small effects (Fisher, 1930), and analyzed the consequences of the shifting of the phenotype -a vector of traits- away from the optimum state. From that perturbed state small mutations will have nearly 50-50 chance to bring the phenotype closer to the optimum, whereas mutations with big effects will increase

the chances to drive the phenotype away from the optimum. Actually, these calculations are naïve and oversimplified. As discussed by Orr (2005), this model only explains how a trait responds by mutational variance, and is a flawed picture when evolution proceeds by standing genetic variation.

In average mutation will decrease the trait value proportional to $\mu \langle \bar{z} \rangle$ (Eq 3.24), restoring the equilibrium value exponentially, in line with Fisher's predictions. Even if the mutation rate is high this would be true, although not obvious from the geometric model. To picture it in quantitative terms, consider as an example the number of ovarioles and thorax length, two traits of *Drosophila melanogaster* (Bergland et al., 2008). Suppose that these traits are displaced 10% above from their optimal value, for example (which are taken to be about 25 ovarioles and 1 mm. resp.) and assume 5 and 7 loci of exponentially distributed effects (expectancies averaged over 1000 realizations of effects). With a mutation rate of $N\mu = 10^{-3}$, it will take about 1150 generations to restore at least the .99$\%$ of the optimum values. In contrast, if $N\mu$ were increased to 1.0, it would only take one generation. In the former case, the trait means will show standard errors of about $\pm\{0.02, 1.7\}$, whilst in the second case it will be of $\pm\{0.16, 12.5\}$. For instance higher fluctuation steps will keep the traits scattered from the optimum value. Even when the time to attain equilibrium is much lower at high mutation rates, the standard error is enormous, so an accurate equilibrium at the optimum will not be typically attained. This can be seen as the impossibility to reach the optimal state in selected realizations of stochastic trajectories.

These neutrality calculations are only one part of the story, showing how evolution would proceed by the variability generated only by mutation, and where at each generation the new alleles might be fixated by drift, as in Fisher's model. The standing variability, over which selection acts, will give very different

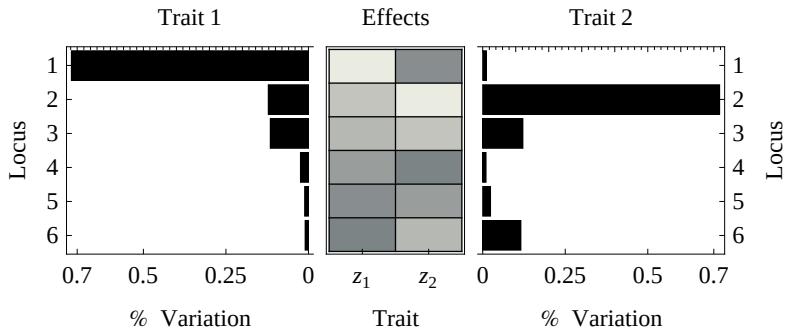


Figure 6.1: Genetic effects. In general, we don't know the distribution that the effects of each locus has over a trait. Thus for simplicity an exponential distribution with mean 1 is assumed, from which the effects of the first trait are randomly selected. For the second trait, the effects are the same, but the positions shuffled. Thus but the total variation they contribute over each trait is the same. These effects will be considered to remain constant during the process of evolution. Numerically random choices are: $\vec{\gamma}_1 = (1.79, 0.74, 0.72, 0.33, 0.23, 0.21)$ and $\vec{\gamma}_2 = (0.23, 1.79, 0.74, 0.21, 0.33, 0.72)$

outcomes.

Assume a converse scenario as before, where selection in the absence of mutation, is applyed. The initial genetic variation will be depleted, and drift will favour fixation, the optimal state (in potential terms, *not* in fitness terms) in the absence of mutation (or at very low mutation rates).

But the response is more difficult to evaluate if mutation generates variability at the same time as selection uses it. We can study this situation comparing the response to selection at different mutation rates. First, variability is generated proportionally to the mutation rate.

A straightforward conclusion is then that the time for changes to be noticeable are larger for lower rates (Figs. 6.2-6.3). We might further assume that before selection effectively acts, the population is at an initial equilibrium (as discussed in Ch. 3), thus the initial standing genetic variation will be low under low mutation rates, for which the speed of change of the traits would be also slow. Since selection acts stronger on the alleles of bigger effect, change in the frequency of alleles at each locus will proceed in a hierarchical fashion. Since most loci have small effects, $\mathcal{G}$ will remain in a period of stasis and its change is delayed to future generations. This is because the genetic variance is more sensitive to rarer (i.e. infrequent) alleles. Only after the major-effecting loci approach close to an equilibrium, a cascade of loci with minor effects flows, inducing the change in the $\mathcal{G}$ matrix. At this point, the impact over the traits is low because -contrary to $\mathcal{G}$- they are mainly influenced by loci with big effects. In conclusion, we expect to observe the changes in genetic (co)variances only at late generations, when the mildest loci begin to change.

Although taking much less time, the situation is similar for high mutation rates. This accelerated differences are mainly due to mutational variation. Even when we would start from low genetic variances, mutation will generate enough of it every generation, thus response to selection is fast. Notwithstanding, the changes in $\mathcal{G}$ are much more dramatic at low mutation rates. The orientation and specially the eigenvalues change considerably at low $N\mu$. The mutational load is larger at large $N\mu$, which is reflected in that the eigenstructure of $\mathcal{G}$ (Fig. 6.2-6.3), in an ellipse graphical representation) are much more similar before and after selection, whilst at low $N\mu$ they differ considerably.

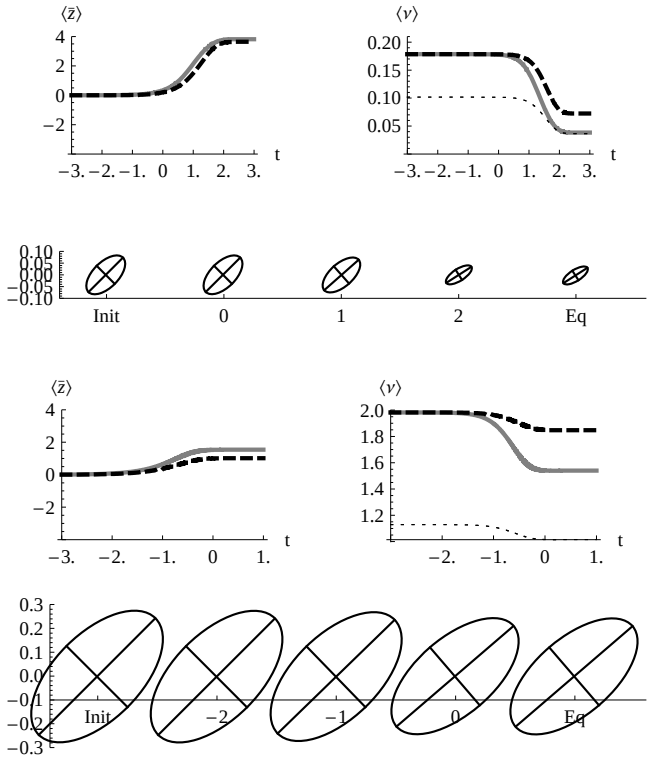


Figure 6.2: Indirect Selection. Response of the traits and genetic (co)variances under pleiotropic effects, subject to selection over the first trait ($N\beta_1 = 2$, black dashed lines) while the second trait is not selected ($N\beta_2 = 0$, gray solid lines). The initial state is neutral for both traits ($N\beta_1 = N\beta_2 = 0$ with constant low mutation rates (leftmost panels, $N\mu = 0.01$) and high mutation rates (rightmost panels, $N\mu = 1.0$). The thin dotted lines in the genetic variance plots indicate the genetic covariances. The lower row show a graphical representation of snapshots of the $\mathcal{G}$ matrices. The length of the semiaxes are given by the $1.96\lambda^{1/2}$, where λ are the eigenvalues of $\mathcal{G}$. The orientation of the ellipses are given by $\mathcal{G}$'s eigenvectors. This representation follows the convention of Arnold et al. (2008). The genetic effects are given in Fig 6.1. The time axes are in logit scale.

6.2 SOME CONSEQUENCES OF PLEIOTROPY

Even in the absence of epistasis and in linkage equilibrium, the effects of a locus over different traits can account for more excentric versions of the above, and to that of univariate traits. Beware that the core of the distribution of allele frequencies of a mutivariate trait is equivalent to that of a single trait evaluated -at every locus- by the product $\vec{\beta}.\vec{\gamma_l}$ (Eq. 4.5, and Appendix section D.1.2). Thus the selective strength experienced at a locus with univariate effect over a trait, can be equivalent to that when the a locus has an effect over arbitrary many (lit. infinite) traits in the multivariate case. Despite how equivalent the distribution of the allele frequencies might result, the patterns of evolution are severely constrained (as exemplified above). Two major sequelae are indirect selection and apparent stabilizing selection, which I will shortly address now.

Indirect selection. When selection is shifted for only one character (say, the first), other characters (the second) will indirectly also experience selection. The shifted allele frequencies will pleiotropically have an effect over other traits. In the example of Figs. 6.2 the values of the trait are very similar (we chose the total effects to be the same 6.1). The pairwise product of the effects will determine the values of the covariance, thus also the speed of change of the trait that is indirectly selected.

Delayed responses. Notice in Fig. 6.2, that genetic variance increases first for the trait under selection (and hence the trait under selection also increases more rapidly than the other). Since loci with high effects respond first, the appreciable increase on the second trait will be registered only when its major contributing loci change, because these loci will not typically have the highest effects over first. In particular, as in Fig. 6.1,

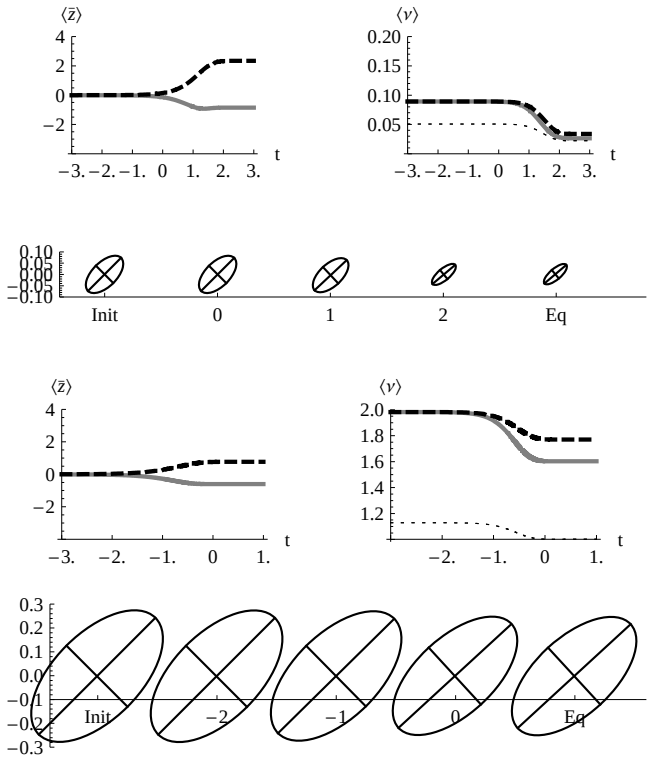


Figure 6.3: Aparent Stabilizing Selection. Response of the traits and genetic variances under pleiotropic effects, subject to selection in opposite directions ($N\beta_1 = 2$, black dashed lines, and $N\beta_2 = -2$, gray solid lines). Otherwise as Fig. 6.2.

locus 1 would be the first to change when selection acts over trait 1, but the effect of this locus over trait 2 is actually minimal, thus any change in the latter would be practically undetectable (as happens in Fig. 6.2)

Table 6.1: Maintenance of genetic variance by pleiotropic effets. Under co-directional selection over two traits, mutation, and drift, the variance maintained is lower than when selection is acting in opposing directions, thus creating the effect of apparent stabilizing selection. The effect is dimmed as the number of pleiotropic loci diminish. Low mutation rates ($4N\mu < 1$) are assumed.

Selection:	Co-directional[a]	Apparent stabilizing[b]		
Pleiotropy:	6[c]	6[c]	3[d]	0[e]
$\nu_1/2\mu$	1.245	2.647	1.954	1.911
$\nu_2/2\mu$	1.245	3.392	2.116	1.911
$\nu_{12}/2\mu$	0.751	2.220	0.411	0.000

[a] $N\beta_1 = N\beta_2 = 2$
[b] $N\beta_1 = 2, N\beta_2 = -2$
[c] Full pleiotropy. Effects as in Fig. 6.1.
[d] The first three loci of both traits are uncoupled.
[e] All loci are uncoupled.

Notice finally that the genetic variance is more critically depleted for the trait under selection, contrasted to traits indirectly selected, which have a milder depletion.

Apparent stabilizing selection. If pleiotropic loci have opposing effects over two traits, then patterns that remind those of stabilizing selection are retrieved (Barton, 1990). Selection of opposed directions over each trait will result in low fitness for the extreme traits. Increased metrics for a trait induce a reduction of the metrics of the other. As a response to selection the second trait tends to restore its value. Conversely, this restoring force induces reduction of the first trait, also restoring its value. Hence, a compromise among both traits is a stable situation. The outcome is similar as in stabilizing selection, where

the extremes are the lowest in fitness. An interesting feature, is that the amount of variability that is maintained by aparent stabilizing selection is higher than when selection acts only over one character, and even higher than when it favours both characters in a common direction (see Table 6.1 and Fig. 6.3).

Degree of pleiotropy. In the examples above I assumed that all loci have pleiotropic response to selection and mutation. But this is not necesarily the case. QTL experiments have shown that the number of loci conveying an additive effect over the traits is not the same, and not even necessarily a superset, of those with pleiotropic effects (Mackay, 2001, 2004; Kalisz and Krishnamurthy, 2007; Albert et al., 2008; Kenney-Hunt et al., 2008; Ma et al., 2008). Therefore my curiosity was awakened by how much variation can be induced by indirect selection under distinct degrees of pleiotropy.

To investigate these effects, I proceeded to systematically uncouple the amount of loci which convey pleiotropy. I randomized the effects of the loci with a exponential distribution (λ=1) and averaged over 10^4 replicas. The results show that the amount of genetic variation in MSD equilibrium linearly decreases with the amount of pleiotropic loci. This trend is particularly strong at low mutation rates because mutational load is -in average- lower. At high mutation rates, the trend is so soft that it would hardly noticeable in experimental observations (Fig. 6.4)

I proceeded then to see the effects of the uncoupling on the evolutionary rates. Under aparent stabilizing selection, pleiotropic responses does not seriously compromise the velocity of change of neither the traits nor the $\mathcal{G}$ matrix. The structure of $\mathcal{G}$, however, as well as the trait values, is affected by selection. Under high levels of pleiotropy, $\mathcal{G}$ has higher eigenvalues, and is more 'bulky' than at low levels, as explained above (Fig. 6.5).

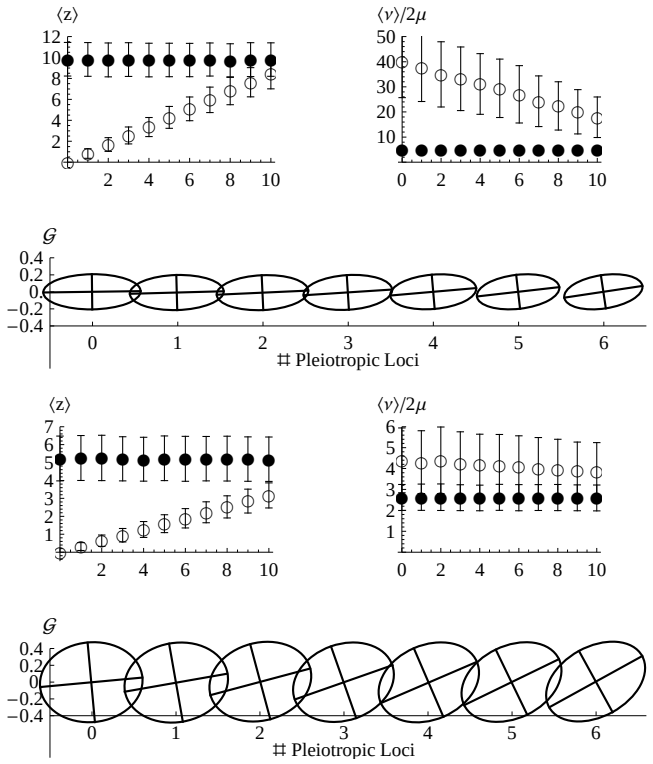


Figure 6.4: Effects of pleiotropy over selection-mutation-drift balance. The equilibrium is maintained by selection only over the first trait (black dots, $N\beta = 2$), while the second trait (open dots) is not under selection. The number of pleiotropic loci is increased systematically (ordinate axes). Each point (and $\mathcal{G}$ matrix) is an average of 10^4 values of the expectations of the mean trait and genetic variances with distinct genetic effects. Each trait consists of 10 loci, whose effect was sampled form an exponential distribution of mean 1. The bars at each point represent half the standard deviation of the samples of the genetic effects. The leftmost panels assume low mutation rates ($N\mu = 0.01$), and the rightmost panels high mutation rates ($N\mu = 1.0$). The lower plots are ellipse representation of the averaged $\mathcal{G}$ matrices (see Fig. 6.2).

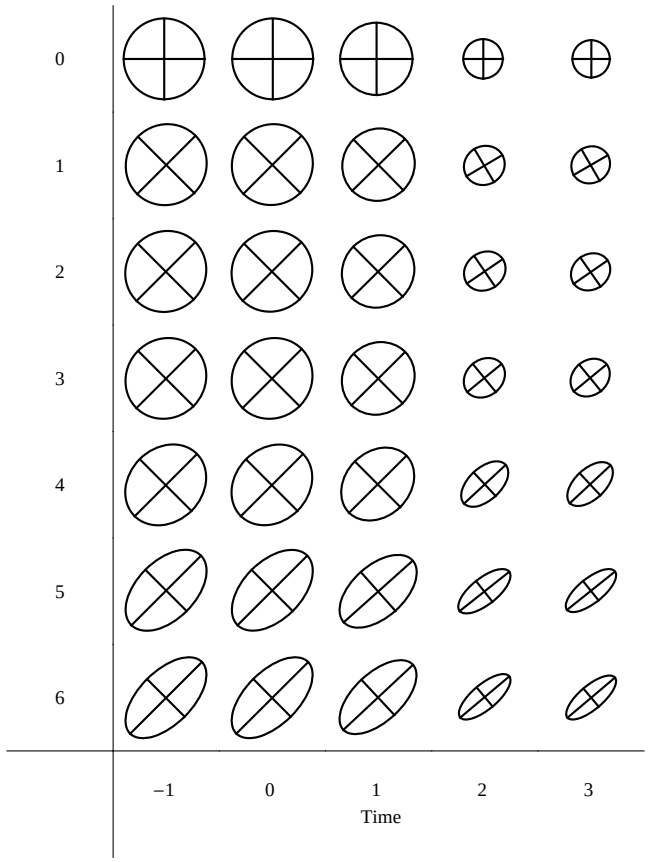


Figure 6.5: Effects of pleiotroy over $\mathcal{G}$'s evolutionary dynamics. Response of the $\mathcal{G}$ matrix of two traits subject to selection in opposite directions, with varying degrees of pleiotropic effects. From top to bottom the amount of pleiotropic loci increase systematically from 0 to 6. Selection proceeds from neutral states ($N\beta_1 = N\beta_2 = 0$) to $N\beta_1 = 2$ and $N\beta_2 = -1$, and $N\mu = 0.01$. The horizontal direction is the time (in log-scale). $\mathcal{G}$ representations as in Fig. 6.2.

6.3 MOVING ON WITH THE $\mathcal{G}$ MATRIX

Pigliucci (2006) has criticized the quantitative genetics approach, and the overall efforts to understand quantitative variation using controled experimental designs on the basis that estimations of $\mathcal{G}$ are *local measures* and change along with the change of allele frequencies, and thus does "not provide a useful measure of the *long-term capability of traits to respond to selection*" (his emphasis). This in a sense is true. But, from my perspective, these critiques are to some extent as narrow as the measures of the heritabilities, or of $\mathcal{G}$. Naturally, a measure is just a (set of) number(s), and the predictive capabilities of any measurement are only meaningful when supported by a theory. If a theory is limited, that might just be the state of the art, and it does not imply that the theory cannot be extended. In fact, the theory behind the $\mathcal{G}$ matrix is still at an early stage, and the predictions that we can make are limited precisely because of that. But in this chapter and the previous, I provided some calculations that allowed long-term predictions of $\mathcal{G}$. We still require refinements in the theory, and maybe a better experimental support in order to really make predictions from measurements of $\mathcal{G}$. I have shown that it is possible to predict the course not of $\mathcal{G}$ itself, but of its expectancy. Yet, these are two different things: an empirical measurement of $\mathcal{G}$ will in general be deduced from a set of offsprings, that is just a realization of the process (this is one of Pigliucci's arguments against the use of $\mathcal{G}$). As we saw in the previous chapter (Fig. 4.2), the deviations induced by genetic drift can be very big. This would redirect the course of the populations away from the theoretical expectancies. But how much? We can calculate the variance of the estimators of $\mathcal{G}$ (Appendix D.1.2, Eqns.D.51), from which we can compute the standard deviation, an expectancy of the error. In Fig. 6.6 an example is presented, along with a real-

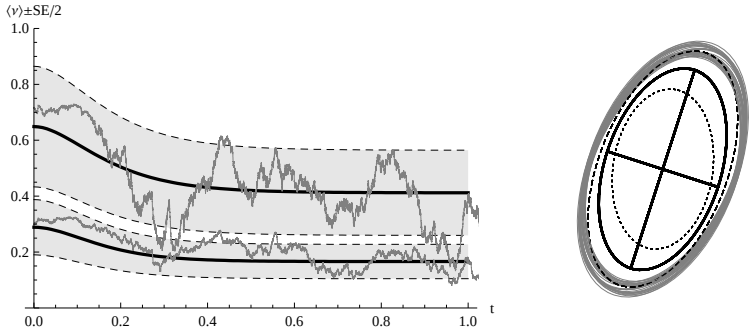


Figure 6.6: Error in the $\mathcal{G}$-matrix's evolution. Left panel: apart from the expectancy of $\mathcal{G}$'s component (genetic variances, solid black lines), we can estimate the expected standard deviation (shaded region between the dashed lines). This measurement is consistent with particular realizations (gray line, from a Wright-Fisher model). The dynamics of $\mathcal{G}$ correspond to Fig. 6.2. Right: the solid ellipse is the equilibrium $\mathcal{G}$, and the dotted ellipses are $\mathcal{G}\pm$ half of the standard deviation. The gray ellipses are distinct realizations.

ization. We can clearly see that the expectancies agree with the Wright-Fisher model. Naturally, an accurate measurement of $\mathcal{G}$ requires sufficient replicas in order to determine confidently the range of variation. Fig 6.7 shows that averages over realizations coincide accurately with the expectations. I have done this for the whole trajectories, but it naturally can be done for a single point in time.

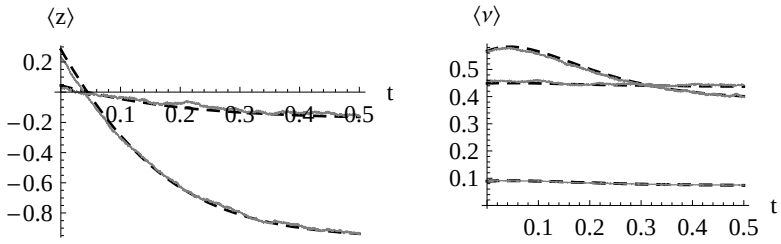


Figure 6.7: Averages and expectancies of $\mathcal{G}$. The left panel shows the evolution of the expectancies of the traits (black dashed lines), over which directional selection is acting, along with averages over 200 stochastic realizations (gray lines). Right: idem, but for the genetic co-variances.

6.4 CONCLUDING REMARKS

The question whether $\mathcal{G}$ evolves or not, is nowadays obsolete. The modern debates are related to the velocity and direction of change in $\mathcal{G}$'s evolutionary course. The answer varies depending on the biological nature of the traits. Still, the formal aspects have remained dim.

I have introduced an application of the statistical mechanics formulation to address several of these questions, at least for the limited interpretation of purely additive traits, under ideal biological conditions. Specific questions remain to be answered, but the general ideas herein introduced, along with those of the previous chapter, indicate that there are good prospects to understand the subject of the evolution of $\mathcal{G}$, and its evolutionary consequences.

But there is another possible way of predicting the course of $\mathcal{G}$. In fact in the calculations of Figs 6.2-6.7 I have not computed the evolution of $\mathcal{G}$ itself, but of the local variables determining the max-entropic distribution. Thus it might be possible to turn the problem around: to determine with accuracy the observables, $\langle \bar{z} \rangle$ and $\langle U \rangle$. Then $\mathcal{G}$ can be just evaluated. But would this work? Actually, it is what I have been doing. The observables ensure the coupling of the micro and macro-states, and knowing the local variables other quantities are computed straightforward. Notice however, that computing $\mathcal{G}$ is necessary but not enough, because the evolution of the local variables – or of any other macriscopic– requires computing the matrix of correlations between the observables and genetic effects, which contains $\mathcal{G}$ but also other entries (Eq. 4.13). Nevertheless, if we have some measurements of the genetic effects, we can, in principle forecast evolutionary response.

Although this is in principle a valid approach, how much the theory can be literally applied to natural populations is obviously a sensitive question. There are many assumptions about the genetics. But although these assumptions can be relaxed, with some technical effort, the applicability of the statistical mechanical theoretical framework can be misleading. In this sense, my main concern is about the fitness landscape: we are assuming very convenient representations. Although we could quantitatively identify if selection is (or not) acting directionally (Hoekstra et al., 2001; Kingsolver et al., 2001), the net effect might not need to be described by a multiplicative effect as we have done; or if its stabilizing, by a Gaussian landscape. Furthermore, even if we could make the simplifications to these simple functions, it is hard to know *a priori* what the selective gradient is, and whether it will remain constant. Long term evolution might comprise very long times, and is unlikely that we can safely assume that selection will act in an homogeneous

way. In this sense even knowing the genetic details of an organism, and the observables, the direction of evolution will be only dictated by the action of selection. Thus the uncertanties over how selection might change direction under ecological varying circumstances would impair a safe forecasting.

On the other hand, we can employ the method for comparative analyses. This is a direction over which I have not developed, and that remains open and promising. Comparing populations, and using the information of the observables, we could perform hypotheses tests to determine and identify the nature of selective processes that have drive the divergence between two populations, a bit on the direction of the calculations of the previous chapter on the *Rana temporaria* populations. This could be a way of falsifying certain approaches, as that of the local equilibrium as an evolutionary forcast method.

Chapter 7

Perspectives: The evolution of quantitative characters under stabilizing selection, mutation, and drift.

7.1 INTRODUCTION

In getting to understand the nature of selection, some contradictions arise, at least from the classical theory of population genetics with respect to maintenance of genetic variability. Not only that we know, empirically and theoretically, that selection depletes genetic variance, and that mutation, with the aid of recombination induce it, but also that the levels that we expect from theory are in most cases lower than those experimentally observed (Maynard-Smith, 1983; Charlesworth et al., 1982, and references therein). But the question is not only one of the matter of scale, but also of logical thought. A standard argument is that extreme phenotypes tend to be less viable than intermediate ones. Examples (a) giraffes with too short necks are dissadvantaged in that they cannot reach the food composing their main diet, and the blood pressure might be too high in their head's vascular system when they bend to drink. Conversely, giraffes with too long necks may not have enough blood pressure in their heads and brains at normal postures; thus intermediate sizes are of higher viability. (b) The amount of chlorophyll in higher plants when too low, cannot account for enough energy transduced for the individual's vital processes; too much chlorophil requires an even bigger amount of energy investment for creating cellular, anatomical and physiological structures that are neither compensated by the energetic gain, nor useful because it would imply an increased metabolic rate that might not be sustainable by size and CO_2 intake rates, and inability to dissipate heat. We can build this kind of argument untill the (z) with most phenotypic characters of any species. For instance, the stabilizing selection hypothesis seems a widespread possibility in natural populations.

Where is then, the contradiction? Although we would expect this kind of selection to be observed frequently, studies on patterns of selection in the wild have revealed that it is scarcely present, and that directional selection is the conspicuous choice (Kingsolver et al., 2001; Hoekstra et al., 2001). If this is the case, there is excess of genetic variability attributable to mutation and linkage. But mutation and recombination rates are not high enough as to account for such levels of variability.

A sound hypothesis is that most characters are under pleiotropic effects, and that these characters have opposing effects to the traits under selection. This alternative, apparent stabilizing selection, was explored in chapter 6.

Another reasonable alternative is that indeed weak stabilizing selection is acting. If the mean phenotype is displaced from the optimal state, directional selection effectivelly acts towards the new optimum. In this case genetic variability can be more easily argued to be maintained at higher levels, as explained above. Notice that the response of genetic variance to selection over a trait tends to be delayed in prolonged application of selection, in particular if there are pleiotropic effects. Unless deviations would be very far from the optimum in one specific (genetic) direction, for example if drift would introduce an irrationally large deviation, genetic variance is unlikely to be (statistically speaking) changed. The return to the optimum would proceed with a linear rate, with variance essentially unchanging at noticeable scales, and experiencing directional selection.

This second alternative, is compatible with the experimental observations (Hoekstra et al., 2001, although with secondary importance after directional selection following meta-analyses)) and with the logical argumentations about the lowered fitness of the extreme phenotypes.

The evidence and arguments for the stabilizing nature of selection, thus demands a dedicate analysis. The approach of the

statistical mechanics theory developed in the previous chapters, will be extended to this situation. First to be able to quantify the course of evolution, and second to make available the quantitative tools for a comprehensive comparative evaluation of the possible evolutionary forces in action.

But the stabilizing selection situation, in inherently complicated. Any given equilibrium between stabilizing selection, mutation, and drift (SSMD) would have many possible microscopic equilibria (Barton and Shpak, 2000; Turelli and Barton, 2004). The dynamic of two or more loci coupled through a trait under stabilizing selection leads to a range of possible dynamics, that are far from entirely characterized (Barton and Shpak, 2000; Willensdorfer and Bürger, 2003; Gavrilets and Hastings, 1993). Although at the moment we are not directly concerned with these dynamics, and the characterization of the local equilibria, it is clear that perturbations to the allele frequencies (e.g. by drift) can induce metastability in these states Barton and Rouhani (1987); Rouhani and Barton (1987). Thus continously perturbating the equilibrium states, leads to an ever changing microscopic dynamics that show increased genetic variance from quantitative measurements approach.

The situation is not as trivial as in directional selection. There are non-linearities in the dynamics, because selection occurs over the squared mean trait, which among other consequences, it fully couples the loci. Hence a decomposition of a polygenic trait as a many independent one-locus problems is not possible, as it was in the case of directional selection. Yet there are stratagems to be victorious in averaging out the microscopic variables. Maximum entropy could work if we find appropriate macroscopics. In short, the question is whether the local equilibrium approximation holds. If it does, we would be free to track evolutionary dynamics, like moving optima, enhaced (or relaxed) strength of selection to the extremes, etc.

7.2 MAX-ENTROPIC APPROACH

In the chapter 3, the methodology inspired by the analogy with statistical mechanics in physics was derived and applied in detailed analyses to directional selection, to study the evolution of quantitative characters in univariate traits. In chapters 4-6 this methodology was extended for multivariate polygenic traits with unequal effects.

We employed maximal entropy as a starting point, constrained by the macroscopic variables that are maintained by the evolutionary processes. That is the trait -or fitness- (maintained by selection), and genetic variability (maintained by mutation), if the selective scenario is directional over an additive trait.

Stabilizing selection over a quantitative character removes from the population those genotypes whose traits are far from an optimum. This is equivalent to have directional selection against the genetic variance. But this is not enough, since we would get for equilibrium a single point at trait zero, without any variance. So we must also include genetic variability (to account for mutations, if the rate is $\mu > 1/4N$, and directional selection over the trait, towards an optimum. Intuitively, this should be enough. But since selection is assumed over the mean trait, then fitness of the mean trait is not the same as mean fitness, so a second order term would appear (variance of the mean trait). Mathematically, we typically choose to model stabilizing selection as a gaussian landscape of fitness:

$$W_z := \exp\left[-\beta(z - z_{op})^2\right] \tag{7.1}$$

We first average over the frequencies of the traits ($P(z)$ to get the mean fitness, $\bar{W} = \int W_z dP(z)$. Now if β is sufficiently small (selection over the trait is weak) we can expand to get

$$\bar{W} \simeq \exp\left[-\beta\nu_z - \beta(\bar{z} - z_{op})^2\right] \ . \tag{7.2}$$

Notice that when expanding the square in the parenthesis, three terms appear: a constant (βz_{op}^2), $\beta z_{op}\bar{z}$ that is directional selection towards the optimum, and $-\beta\bar{z}^2$ that is selection against the squared trait. The stationary distribution of the allele frequencies is recovered from entropy maximization (Eq. 3.6, Ch. 3) constraining the expectations of the above quantities, i.e.:

- Normalization of the distribution.

$$\int_{(0,1)^n} \psi(\mathbf{p}) d^n\mathbf{p} = 1 \rightsquigarrow \lambda \rightsquigarrow \mathbb{Z} \tag{7.3}$$

- Selection of the mean trait towards the optimum

$$\int_{(0,1)^n} \bar{z}\psi(\mathbf{p}) d^n\mathbf{p} = \langle \bar{z} \rangle \rightsquigarrow 2N\beta \tag{7.4}$$

- Selection against genetic variance.

$$\int_{(0,1)^n} \nu_z\psi(\mathbf{p}) d^n\mathbf{p} = \langle \nu_z \rangle \rightsquigarrow 2N\sigma \tag{7.5}$$

- Selection against the variance of the mean trait

$$\int_{(0,1)^n} \bar{z}^2\psi(\mathbf{p}) d^n\mathbf{p} = \langle \bar{z}^2 \rangle \rightsquigarrow 2N\alpha \tag{7.6}$$

We know that maximizing entropy leads to the distribution 3.6, that in this case of SSMD case it would be

$$\psi(\mathbf{p}) = \frac{\phi}{\mathbb{Z}} \exp\left[2N\beta\bar{z} + 2N\alpha\bar{z}^2 + 2N\sigma\nu_z + 2N\mu U\right] \tag{7.7}$$

with $\phi \equiv \phi(\mathbf{p}) = \prod_{\ell=1}^{n}(p_\ell q_\ell)^{-1}$, as explained before, and if we are able to compute the integral $\mathbb{Z}$, we will have a macroscopic description of the system, that is supported by and consistent

with the microscopic stationary dynamics. But more interesting than finding the closed form expression (if possible at all) is to be able to predict, also from this macroscopic point of view, the dynamics. Microscopically, this is given either in a stochastic version, or probabilistically by the corresponding Wright-Fisher process, and its diffusion equation, respectivelly. Actually those will be our points of comparison.

The partition function can be explicitly written if we define how the trait relates to the genetic variables. As before, we assume an additive trait of n loci, each with constant effect γ_ℓ (constant in the sense that they do not evolve, but each effect is in general different at every locus). As treated in previous chapters (see also appendixes) the mean trait and genetic variance are functions of the allele frequencies. However, the squared trait introduces some tricky properties that complicate the computation of $\mathbb{Z}$, since it cannot longer be expressed as the product of the partition functions of independent loci. So, for reasons that will be obvious later, I will express any power of $\bar{z}$ implicitly. Thus the partition function is

$$\mathbb{Z} = \int_{(0,1)^n} d^n\mathbf{p} \prod_{\ell=1}^{n} (p_\ell q_\ell)^{-1} \times \tag{7.8}$$
$$\times \exp\left[2N\beta\bar{z} + 2N\alpha\bar{z}^2 + 2N\sigma \sum_{\ell=1}^{n} \gamma_\ell^2 p_\ell q_\ell + 4N\mu \sum_{\ell=1}^{n} \log(p_\ell q_\ell)\right] ,$$

which cannot be computed in a closed analytical form. Before dooming the expression 7.8 to numerical computations, we can do something about it. The reason is that as it is, its computation is $n-$dimensional, thus prone to slow convergence, and since the allele frequencies might be clustered, the integrals might be close to zero. Also, there are some simplifications that can be made. The whole point, is that if we are able to compute $\mathbb{Z}$, we can as well compute any statistic. Hence we are

able to test the local equilibrium as a model for the polygenic evolutionary dynamics.

The trick, following Barton (1989, p. 64), is to transform the $n-$dimensional integral into a complex-valued 1-dimensional integral, which if not anallytically solvable, at least simplifies (a) the calculation of the expressions of the traits, and (b) their numerical computations.

I will give only a sketch on how to proceed with the calculations. The intermediary steps should be straightforward. Lets calculate a function $F(\bar{z})$. It can be expressed as an integral of a Dirac delta function as:

$$\begin{aligned} F(\bar{z}) &= \int_{-\infty}^{\infty} F(\zeta)\delta(\zeta - \bar{z})d\zeta \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(\zeta)\exp\left[-i\omega(\zeta + \bar{z})\right]d\zeta d\omega \ . \end{aligned} \tag{7.9}$$

where the second expression employed the (inverse) Fourier transform of an exponential function to express Dirac's delta. If the last expression is not disregarded by the reader, the advantages are clear: the function F is evaluated in a mute variable, ζ and the dependence of $\bar{z}$ are segregated to the exponential factor, which avoids cross terms (like $p_i p_j$) of the allele frequencies. Following this formula 7.8,

$$\begin{aligned} &\exp\left[2N\beta\bar{z} + 2N\alpha\bar{z}^2\right] = \\ &\sqrt{\frac{\pi}{|2N\alpha|}}\int_{-\infty}^{\infty}\exp\left[-i\omega\bar{z}\right]\exp\left[-\frac{(2N\beta - i\omega)^2}{8N\alpha}\right]d\omega \end{aligned}$$

This expression can be substituted into Eq. 7.8, which after some calculus it gives

$$\mathbb{Z} = \sqrt{\frac{\pi}{|2N\alpha|}}\int_{-\infty}^{\infty}\exp\left[-\frac{(2N\beta - i\omega)^2}{8N\alpha}\right]\prod_{\ell=1}^{n}\mathbb{Z}_{\ell\ (\mu;-\frac{i\omega}{2N}\gamma_\ell;\sigma\gamma_\ell^2)}d\omega \tag{7.10}$$

where the per-locus partition functions are given by:

$$\mathbb{Z}_\ell := \int_0^1 \exp[2N\beta(2p_\ell - 1) + 4N\mu \log(p_\ell q_\ell) + 2N\sigma p_\ell q_\ell](p_\ell q_\ell)^{-1} dp_\ell \tag{7.11}$$

Admittedly, the expressions are not too simple, and lack general solutions[1]. Even if each $\mathbb{Z}_\ell$ is to be integrated numerically, these are independent among each other. Thus originally, there were n fully coupled integrals (big problem), and now there are $n+1$ integrals, one of which depends on the other n integrals, but which are independent among each other (small problem).

The general partition function. Eq. 7.10 can be dissected into two terms, generally of the form

$$\int F_{1\ (\omega|\alpha,\beta)} F_{2\ (\omega|\mu,\sigma)} d\omega$$

that is, it is separable into two terms which depend on non-overlapping (intensive) variables. This is a very useful property when calculating the macroscopics.

Take notice that solving a one-locus problem for stabilizing selection might sound as a trivial exercise. But there are three reasons why at the moments it is desirable to do it. First, the polygenic expectations can be represented (exactly) as convolutions of the one-loci corresponding formulas (see appendix D.2); thus it is a necessary move. Second, the Fourier-representation, being equivalent to the n-loci representation of the partition function (and of the expectations), can be checked for one locus problems; this is just a control of the numerical experiments.

[1]The per- locus partition function, however can be expressed in Taylor series over $\sigma = 0$, in which case the formula for directional selection is recovered. The terms for $\sigma^k,\ k > 1$ involve derivatives of this partition function at $\sigma = 0$, and thus are the single-locus-directional-selection expectations.

Third, the one-locus SS problem has implications for the DS case, with respect to the boundary problem that appear near $N\mu = 1/4$, and makes the local equilibrium fail. One locus SS might provide a subterfuge for this collapse.

7.3 SINGLE LOCUS DYNAMICS

In order to solve the general case of the multi-locus dynamics, we need to have a complete characterization of the statistics of single locus model. Among virtues of a mean trait affected by only one locus is that the variance of such mean trait is proportional to the genetic varince, since

$$\begin{aligned} \bar{z}^2 &= \gamma^2[2p-1]^2 \\ &= \gamma^2 - 2\nu_z \ , \end{aligned}$$

Because the effects γ are non-evolving parameters, the statistical mechanics in this case does not require constraints in both $\bar{z}^2$ *and* ν_z; it rather requires constraints in *one* of them. Yet the general locus formula 7.8 applies. For one locus, we will end up with only one of these two quantities, say genetic variance, and the multipliers to the constraints over entropy maximization will be reduced as

$$\begin{aligned} \sigma &\mapsto \sigma - 2\alpha \\ \lambda &\mapsto \lambda - \alpha\gamma^2 \\ \mathbb{Z} &\mapsto e^{-\alpha\gamma^2}\mathbb{Z} \ . \end{aligned} \tag{7.12}$$

Hence, a single-locus model of SS requires only three macroscopics. We could say that this case is a small extension to the directional selection case, where we took a second order approximation to mean fitness (which would result in including genetic variance as a second order correction term in the mean

fitness term). But his extended model gives whole new properties, and is more than a small quantitative correction. First, on the technical side, including selection against the variance does not allow the privilege of having closed form solutions of the partition function (or of the expectances), so we must proceed numerically and/or with some approximations. Notwithstanding, the integration procedures are not too demanding computationally, since most integrands are well-behaved, even near the point $\mu \sim 1/4N$. Second, including selection against (or for) genetic variance breaks the symmetry that exists in directional selection with respect to $\langle \bar{z} \rangle$, where this function is odd with respect to β and $\langle \nu_z \rangle$ is even. Directionally selecting for a favorable allele is -from the point of view of genetic variance- equivalent to selecting for the contrary allele. Also the measure $\langle U \rangle$ would be unaffected. However, if selection over genetic variance is included all these symmetries disappear. Selection will still deplete genetic variation, but on one direction (favoring an allele) will in general be higher than when favoring the contrary allele. The same is true for generic variability $\langle U \rangle$.

The evolutionary dynamics of this one-locus system under SSMD can be computed through the local equilibrium approximation (Section 3.2). That is to calculate the rate of the effective parameters $(\mu^*, \beta^*, \sigma^*)$ that correspond to the quantitative measurements ($\langle U \rangle$, $\langle \bar{z} \rangle$, $\langle \nu_z \rangle$) at every time-point during transient (i.e. non-equilibrium) evolution. Mechanistically, this is determined by the change of the allele frequencies in the population, averaged over the drift realizations. This is described by the diffusion equation (Crow and Kimura, 1970, Ch. 8). The change in the observables (the quantitative variables required for the max-entropic restrictions) follows Eq. 3.13. Thus in addition to the mutational variability U and the mean trait $\langle z \rangle$, we also need to include as observable the expectations of the genetic variance. To proceed in local equilibrium analysis (Eqns. 3.15

and 3.16), the matrices of genetic co-variances B and of fluctuations C are required.The matrices are, respectively:

$$B = \left\langle \begin{array}{ccc} H & -2\bar{z} & 2\bar{z}^2 \\ -2\bar{z} & \nu_z & -\bar{z}\nu \\ 2\bar{z}^2 & -\bar{z}\nu & \bar{z}^2\nu \end{array} \right\rangle \tag{7.13}$$

$$C = \{\text{Cov}(A_i A_j)\}_{i,j\in\{U,\bar{z},\nu_z\}} \ , \tag{7.14}$$

where each of the terms can be found in the appendix D.2, and the dynamics are given by

$$\frac{d\alpha^*}{dt} = C^{-1}.B.(\alpha - \alpha^*) \ . \tag{7.15}$$

Here $\alpha = (\mu, \beta, \sigma)^\top$ ($\top$ stands for transpose, since the formulas require a column vector). Beware that even if this equation seems to 'include' Eq. 3.24 (for directional selection) in that there is an extra column with respect to that, the explicit forms of the statistics are different. Implicitly there is a resemblance, although the specific forms of the expectances would be very much different, because -as can be noticed from Eq. 7.11 and the formulas in the Appendix D.2-, the same statistics will have different quantitative values in directional and stabilizing selection.

Now it is possible to proceed for some case studies. Compare to directional selection, where only two intensive variables existed: mutation rate and selective value of the trait. Now there is another one, selective value of the genetic variance. Thus it is possible to have a bigger scope of possibilities in the evolutionary dynamics. Only a handful will be investigated here.

Evolution towards directional selection, mutation and drift. This single locus model, would reduce to that of directional selection if we let $\sigma = 0$. How would be the response to this situation?

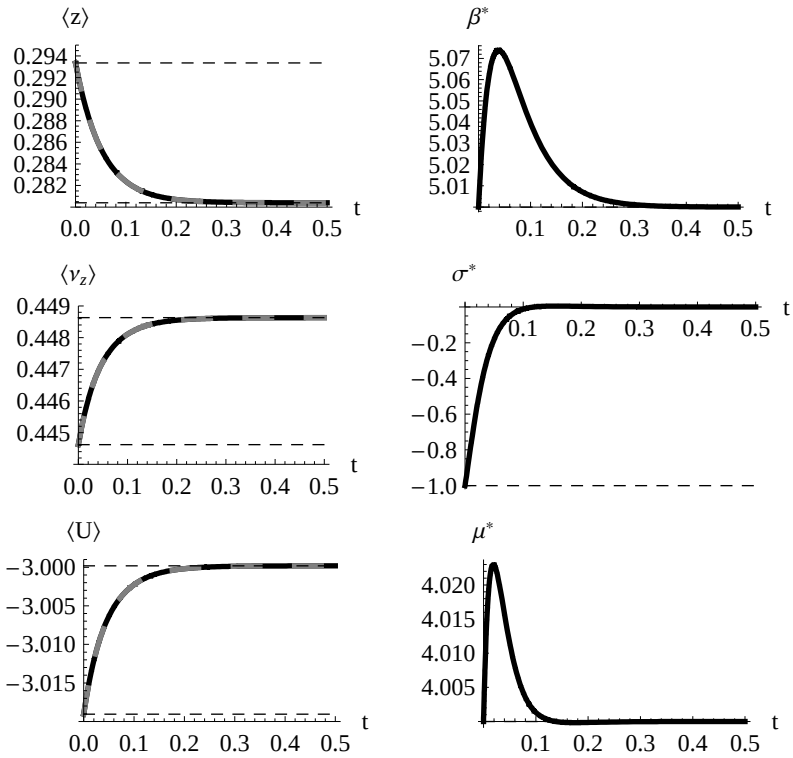


Figure 7.1: Evolutionary dynamics from an initial equilibrium maintained by selection for the trait, selection against the variance, mutation and drift, evolving towards a state without selection against the variance. This is parametrized by $(N\mu, N\beta, N\sigma) = (4, 5, -1) \to (4, 5, 0)$. Black curves: local equilibrium calculations; gray dashed curves: diffusion equation integrations; dashed thin lines: equilibrium values.

Fig. 7.1 shows this experiment. As we would expect, the genetic variance increases (since there is no selection against it anymore). Other consequence is increased mutational variability and reduced mean traits. But notice that even when the

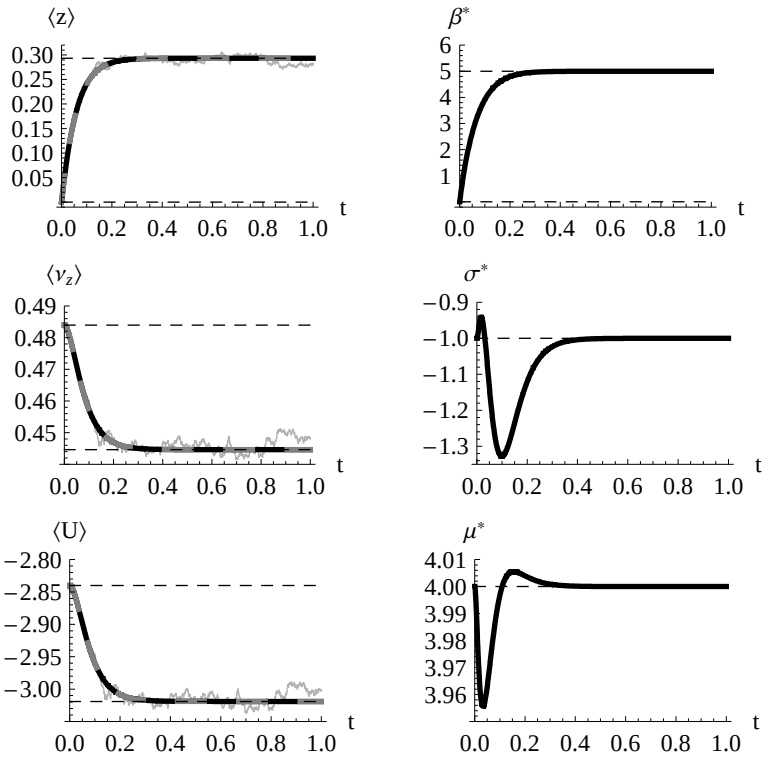


Figure 7.2: Evolutionary dynamics responding to a shift in the optimum. Starting at an equilibrium state where the optimum value is at the origin, the system evolves towards an equilibrium at a higher optimal value. The system is parametrized by $(N\mu, N\beta, N\sigma) = (4, 0, -1) \to (4, 5 - 1)$. The thin gray lines, are averages over 600 realizations of solutions to the Wright-Fisher model, with the above parameters with a population size of $N = 10$. Otherwise as in Fig. 7.1.

traits are (in average) reduced, they are more widely spread, so we would expect to find bigger and lower extremes.

Shifted Optimum. Many problems in stabilizing selection assume that there is a shifting optimum that motors the evolutionary dynamics. This situation can be modeled by a sudden change in this optimum, to which the response will smoothly described by the local equilibrium. The trait will experience directional selection. As regarded in Fig. 7.2, the response is comparable to that of directional selection (Figs. 6.2 -6.3): increase in the trait, and depletion of genetic variance and mutational variability.

Although at first sight no major qualitative difference is noticed with respect to the DSMD, recall the problem near the border $\mu \sim 1/4N$ (Fig. 3.11, and section 3.4) where allele fixation suddenly becomes abruptly likely and makes local equilibrium inapplicable. If genetic variance is 'controlled' (i.e. prone to selection), the problem seems to disappear. Figure 7.3 shows that shifting the optimum at different mutation rate, closer to $1/4N$ works out well. The statistical mechanical approach does not fail like in DSMD. (Actually, although only a matter of numerical methods, the differential equations of the local equilibrium approach behaves better than the partial differential equations of the diffusion equation which has, under many methods, a leak of probability density mass. This was corroborated through averaging 600 realizations of the corresponding Wright-Fisher process; Figs. 7.2 and 7.3.)

Deaccelerating the mutation rate. The '$4N\mu$-boundary-problem' in the DSMS formulation with statistical mechanics limits the case studies that we can address with this method. That is unfortunate because the effects of lowering (or increasing) the rates of mutations or studying bottlenecks, for example, cannot be analyzed. However, this problem seems to be absent under SS. Fig. 7.4 shows how evolution, using the top-down approach, is accurately described. Mutation $N\mu$ was switched

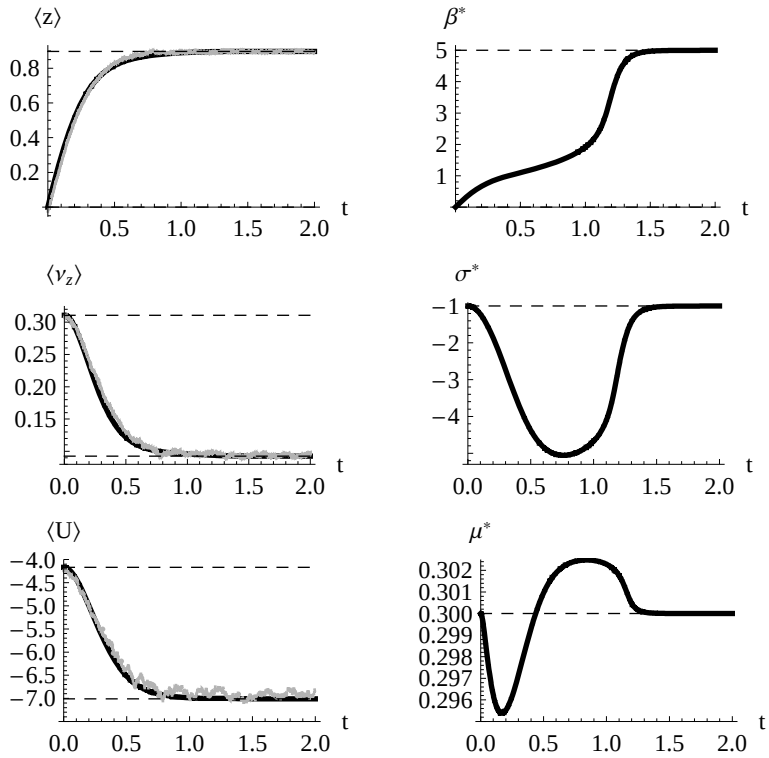


Figure 7.3: Evolutionary dynamics responding to a shift in the optimum at low mutation rates, with $N\mu = 0.3$. Otherwise as in Fig. 7.2

down to 0.3 without apparent discordance in any of the macroscopics. Notice that the local variables show curious paths as $N\mu \to 1/4$.

Revisiting directional selection. The paragraphs above show that as $N\mu \to 1/4$, where statistical mechanics fails under DSMD, evolution is authentically traced. Moreover, when

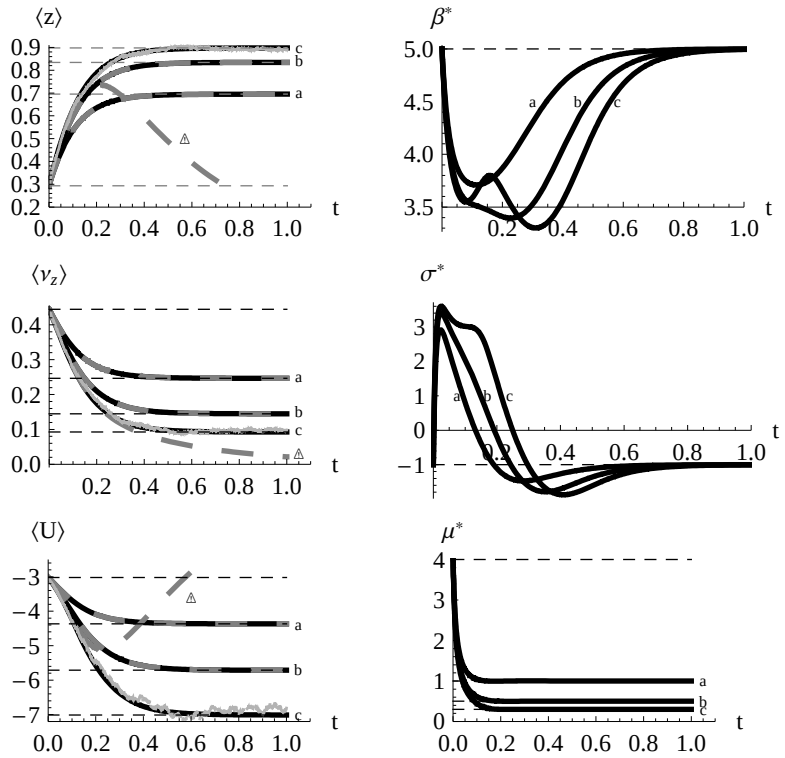


Figure 7.4: Evolutionary dynamics when mutation is slowed down (a) $N\mu = 1$, (b) $N\mu = 0.5$ and (c) $N\mu = 0.3$. Legends as in Fig. 7.2. Regard that for (c) integrations using the diffusion equation have a leak of probability density (⚠).

$\sigma \to 0$ the statistics of DSMD are recovered. Thus there is a new prospect: if at equilibrium we constrain $\sigma = 0$, but we allow it to evolve, we could predict the dynamics of DS near the critical mutation rates. I performed this experiment considering the SSMD statistics at $N\mu = 0.3$. Naturally at equilibrium DS and SS produce the same results. But in the course of evolution,

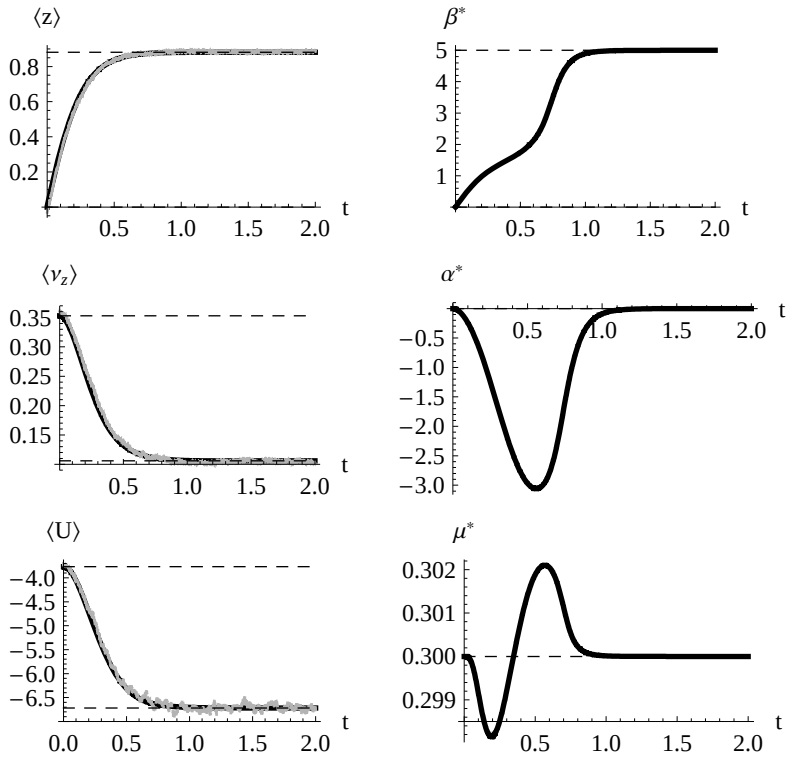


Figure 7.5: Evolutionary dynamics after shifting the optimum at low mutation rates. The solutions are compared to the Wright-Fisher process. Parameters of the system: $(N\mu, N\beta, N\sigma) = (0.3, 0, -1) \rightarrow (0.3, 5, -1)$. Otherwise, as in Fig. 7.2

as shown in Fig. 7.5, the predictions match that of the Wright-Fisher model. Surprisingly, the method is robust: calculation at the critical point (setting $N\mu = 1/4$) still give propitious predictions, Fig. 7.6 (curiously, the computing time for the numerical solution is considerably higher than in other cases, roughly an hour, at least two orders of magnitude higher than for the

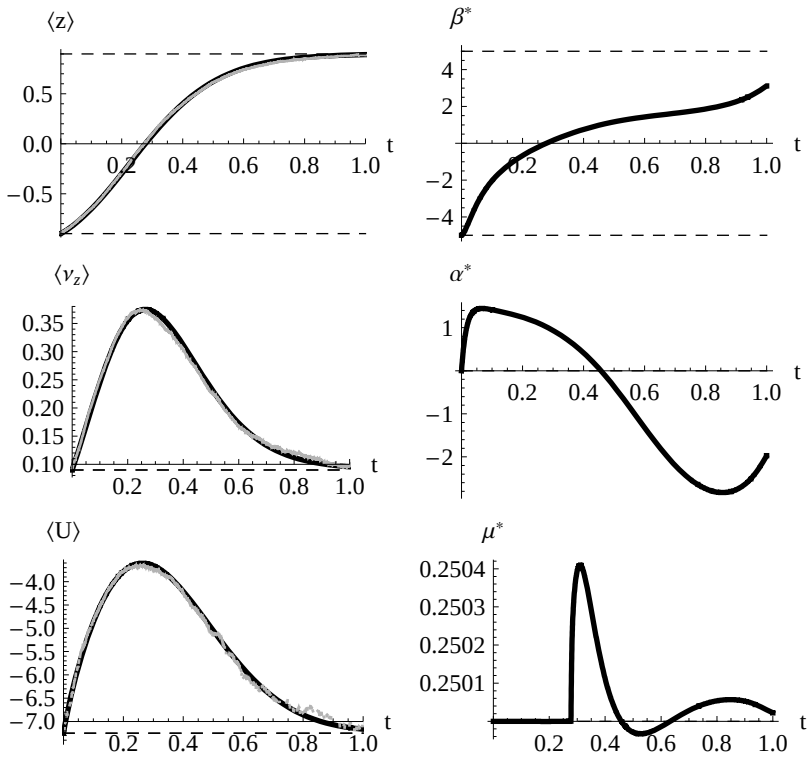


Figure 7.6: Evolutionary dynamics after shifting the optimum at the critical mutation rate. The solutions are compared to the Wright-Fisher process. Parameters of the system: $(N\mu, N\beta, N\sigma) = (1/4, 0, 0) \rightarrow (1/4, 5, 0)$. Otherwise, as in Fig. 7.2

previous ones, between one and two min.)

Summary. The one-locus model has shown that (i) the SM method is not limited to simple directional selection descriptions, (ii) the point at $4N\mu = 1$ is not a limitation for the macroscopic dynamics when including genetic variance as constrained macroscopic, (iii) there is a smooth limit from stabilizing selection to directional selection when $\sigma \to 0$ and (iv) that evolution under directional selection might have stages where the effective forces out of equilibrium are of stabilizing nature.

7.4 FORMULATING POLYGENIC DYNAMICS

Much of the work for polygenic systems under stabilizing selection has been achieved using a 2-locus model or a haploid approximation (Bürger, 2000, Ch.VI). Weak (Gaussian) stabilizing selection is unable to maintain variability, provided that the contribution of all loci over the trait is the same (Wright, 1935, unless linkage is strong Karlin and Feldman, 1970). However, if the allelic effects at each locus are different, then it is possible to keep elevated genetic variance Nagylaki (1989); Gavrilets and Hastings (1993). In any case, there are many microscopic equilibria. For a system as simple at two loci, there are at least 18 equilibrium points (Willensdorfer and Bürger, 2003). All of these are able to maintain distinct degrees of variability. Yet they do not account for the levels observed in quantitative traits, for typical allelic mutation rates.

The other extreme, which is also a common approach, is the infinite alleles model. As we mentioned before, the House of Cards (Kingman, 1978; Turelli, 1984) and the Gaussian approximations (Kimura, 1965a; Lande, 1976) can give an idea on the amount of quantitative variation that is maintained by mutation-selection drift. But as Slatkin and Frank (1990) point, "neither model can be regarded as being typical". The amount

of variation that is predicted curiously depends very slightly on the amount of loci (Turelli, 1984). This is a consequence that the distinct loci have alleles are close to fixation, and variability is maintained in only one of them.

The exact model (hypergeometric), on the other hand, consists of many genes of equal effects, and allows to identify many possible combinations of microscopic equilibria (Barton and Shpak, 2000). Essentially, all genotypes with the same amount of favorable alleles have the same fitness (although not all of them are stable).

We thus see, that to analyze the composition of the alleles in the population requires a thorough characterization. Yet the macroscopic states to which these states correspond are much more simpler. As mentioned above, and seen from the partition function 7.10, the polygenic statistics are not simple 'superposition' of the effects of each locus. Yet the polygenic framework for SSMD relies on the properties of single loci, although in in a non-linear way. Since we need four macroscopics to define the SSMD equilibrium, then the matrices B and C require also an extra dimension, the statistics for $\bar{z}^2$. The problem is more than just calculating the necessary parameters in these matrices. Not only that we need to calculate for each locus these quantities, but we need to convolve them with the trait distribution in a complex space. To my big regret there is little hope that analytic expressions are possible; although perhaps approximations will be workable, which certainly would enlighten our understanding of evolutionary quantitative genetics. For the moments the goal is to set up the problem. Extensive investigation of the macroscopic solutions is needed, since the microscopic dynamics have a wide range of solutions whose consequence over the macroscopics' we don't know.

For this purpose we can apply the theory stated above, which accounts to extend the matrices 7.13 to include the effects of

the variable $\langle \bar{z}^2 \rangle$. This leads to

$$B = \left\langle \begin{array}{cccc} H & -2\bar{z} & 4(\nu_{\max} - \nu_z) & -4\bar{z}^2 \\ -2\bar{z} & \nu_z & \mu_{z\,3} & 2\nu_z\bar{z} \\ 4(\nu_{\max} - \nu_z) & \mu_{z\,3} & \mu_{4\,z} & 2\bar{z}\mu_{z\,3} \\ -4\bar{z}^2 & 2\nu_z\bar{z} & 2\bar{z}\mu_{z\,3} & 4\nu_z\bar{z}^2 \end{array} \right\rangle \tag{7.16}$$

$$C = \{\mathrm{Cov}(A_i A_j)\}_{i,j\in\{U,\bar{z},\nu_z,\bar{z}^2\}} \tag{7.17}$$

where each of the terms can be found in the appendix D.2.

The reader may notice the following difference with respect to 7.13. Besides the above mentioned extension, the column and row corresponding to ν_z have been written in different way. The reason is that some identities do not apply for multilocus formulas. For example, the term $2\gamma^2 pq\gamma(1-2p)$ for one locus corresponds to $-\nu_z\bar{z}$ however the term $\sum_\ell 2\gamma_\ell^2 p_\ell q_\ell \gamma_\ell (1-2p_\ell)$ is *not* the same as $-\nu_z\bar{z}$, but rather the third moment, $\mu_{3\,z}$ of the trait within a population. For one locus is then true that $\mu_{3\,z} = -\nu_z\bar{z}$ in the same way that it is true that, as we saw, $\bar{z}^2 = 1-2\nu_z$. But these identities are not extendable to the polygenic formulas. In other words, the statistics for the mean trait are not necessarily a lumping of the statistics of the individual loci, as in the case of the mean trait or of DS.

We need to calculate these macroscopics numerically. There are some tricks to calculate them, from the Fourier-space integrals, as given in Appendix D.2. In short, separating the partition function as indicated above leads to some ways of expressing the polygenic statistics as a function of the single-locus statistics. These calculations are much simpler than those in the space of genetic frequencies, essentially because these are 1-dimensional calculations. Still, the amount of time they take to compute is enormous , making it impractical to compute for many loci (see below). Thus further work is needed to advance

in this direction. But for the moments, we are able to make some equilibrium predictions.

If we set $d\langle A_j\rangle/dt = 0$ then we obtain the conditions for mutation-selection-drift equilibrium. For the sake of simplicity, lets assume that the trait distribution is normal. In that case $\mu_{z\ 3} = 0$ and $\mu_{z\ 4} \propto \nu^2$ (as in Barton and Turelli, 1987, following Lande, 1976). We can then obtain some expressions which we are able to interpret. For the genetic variance we have that

$$\langle \nu_z \rangle = \frac{2\mu}{\beta}\langle z \rangle - \frac{2\alpha}{\beta}\langle \nu_z \bar{z} \rangle \ . \tag{7.18}$$

The last term is absent in the case of directional selection. Notice that even if $\mu \to 0$ the variance is still maintained by selection. Thus it is possible to increase the genetic variance without increasing the rate of mutation. This effect was identified by Gavrilets and Hastings (1993) for a two locus model. But we find here for arbitrary number of loci. The relation between the expectancies presented here and previous estimates of genetic variance, like those mentioned above, is that the expectancies discussed here is comprised of all those microscopic equilibria. Since drift is present, there will be shifts between those microscopic equilibria, and we are averaging over those.

Figure 7.7 shows that *without changing the mutation rate*, genetic variance can be maintained, and even increased by other factors. The observables are intrinsically pleiotropic, thus the change of a given factor results in the change of all observables. This is what is show in Fig. 7.7: when we select for the trait (A,B), for the genetic variance (C,D) and for the variance of the trait (E,F) we still get a response of the genetic variance, in the last two cases, an increase. These changes however, are lamer when more loci are present (even if most of them are of small effect), also indicating a stronger pleiotropic effect.

Although at the moments the efficiency of the algorithms are

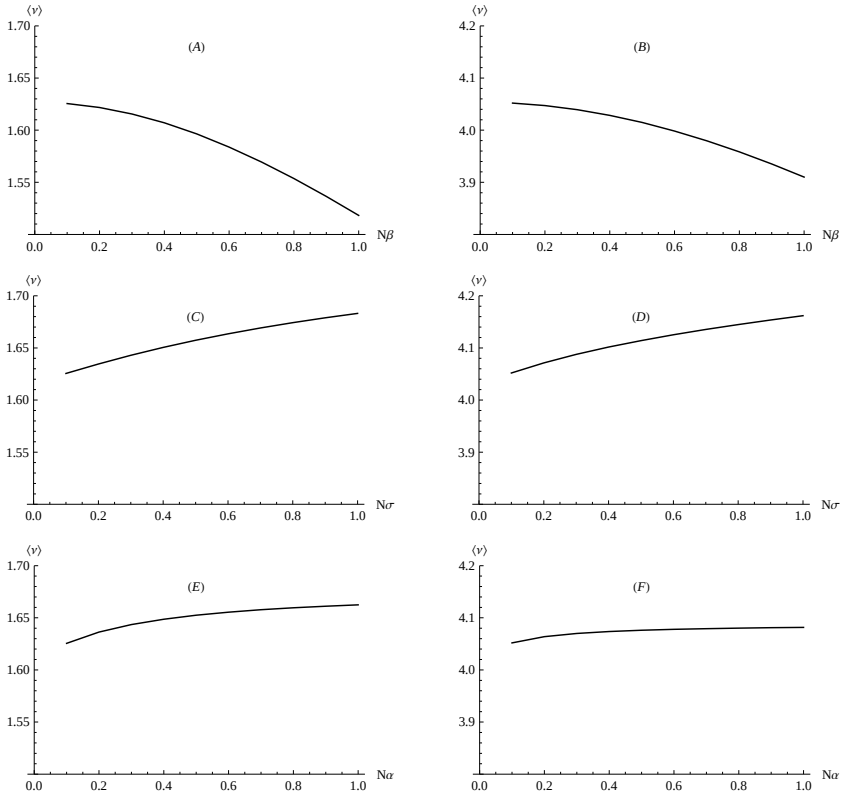


Figure 7.7: Expectancies of the genetic variance of polygenic traits under stabilizing selection as a function of (A,B) $N\beta$, (C,D) $N\sigma$, (E,F) $N\alpha$. Left column: traits composed of 4 loci of effects $(0.15, 0.76, 1.09, 1.73)$. Right column: traits composed of 10 loci of effects: $(0.03, 0.12, 0.36, 0.50, 0.68, 0.80, 0.83, 1.01, 1.38, 2.6)$. Unless the parameters are changed as indicated in the axes, these are $N\beta = N\sigma = -N\alpha = 0.1$ and $N\mu = 0.5$.

limiting the numerical analyses, I calculated the evolutionary response for a two locus system after shifting the optimum 7.8. Here we recover the classical result that selection depletes ge-

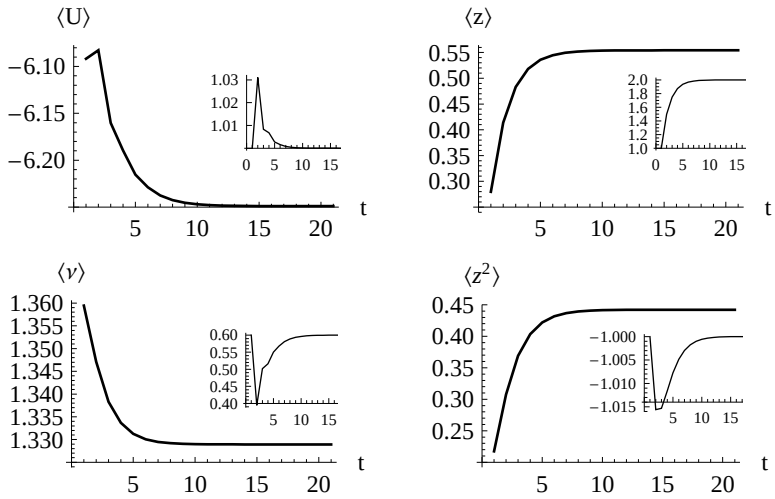


Figure 7.8: Evolution of a tri-locus trait under stabilizing selection. Evolutionary response of (A) genetic variability (inset: local mutation rate $N\mu^*$), (B) mean trait (inset: local selection $N\beta$), (C) genetic variance (inset: local selection $N\sigma$), (D) square of the mean trait (inset:local selection $N\alpha$). The initial state is given by $N\mu = 1.0, N\beta = -1.0, N\sigma = 0.6, N\alpha = -1$, (equivalent to the optimum at $z_o = -1/2$), and $N\beta$ is changed to 2.0 (equivalent to a shift in the optimum to $z_o = -1$). The loci have equal unit effects.

netic variance. However, we re not imposing any restriction on whether specific combinations of the loci should match the optimum, or that the mean trait itself is at that point. In general, $\bar{z} \neq z_o$, and also $\langle \bar{z} \rangle \neq z_o$.

At the moments, I have found hope to be able to predict the evolution of quantitative characters. The pieces seem to be coming together to answer the fundamental question in quantitative genetics: how is genetic variability maintained? and its related dynamics counterpart: how does genetic variability evolves? I have given only partial answers but the trend is set,

and hopefully sooner than later conclusive answers will come, for which only some details have to be worked out, as explained in the following sub-section.

7.4.1 Overcoming numerical limitations

Unfortunately for the problem dealt with, the numerics are difficult. The main reason for this is that the integrals that we need to perform to compute the observables, involve the product of the per-locus partition functions. These products are often falling in the limits of numerical zero. Of course, since all macroscopics involve the ratio of an integral with the partition function, although each term in itself is small, their ration converges to a finite number. But numerically this is problematic. Most of the time, even such calculations can actually be performed, but the amount of time that the integrator takes to compute them, is enormous. And this grows with the amount of loci. Second, the calculations involve the inversion of the matrix of covariances, which even numerically is very time costly. One stem on the computation can take (depending on the required precision and on the precise values of the parameters) more than an hour. Thus computing a whole trajectory is for practical terms, impossible. Third, if the step for integrating the trajectory are not small enough, then the computations simply diverge.

This can (and will) be solved. There are two methods that can be combined at this point. On the one hand, we can perform a Monte Carlo simulation to perform the integration, using a variant of the Metorpolis-Hastings algorithm (Metropolis et al., 1953; Hastings, 1970). (Notice that a Monte Carlo simulation with the Metropolis-Hastings method *is not* the same as the Monte Carlo method for integrating a function, Press et al., 2007, Ch. 7 p. 397-402) The virtue of this method is that we

do not need to compute the normalization constant of the distribution (i.e. the partition function) so many of the numerical issues are avoided. Notice here two hindrances. First, the integrals are complex, so the methods needs to be adapted for this situation. For this we can expand the macroscopics in their real and complex part and evaluate them separately. But this separation involves further algebra, which at the moments I have not explored.

The second way to solve it, is representing the integrals in terms of a series expansion with respect to $N\sigma$. Then the resulting terms are statistics of directional selection with dominance effects. These statistics are easier to compute, as shown in the previous section. The problem there is that we must include multinomial terms, whose sums are also hard to compute. A possible solution is to sample randomly the multinomial distributions (which is actually faster!) and evaluate the macroscopics at these points.

I have advanced with both method, although there is still some tailoring to be done and implement working versions. For the moments, further insights are disguised in the complexities of the analytic results and in the hindrances to unveil them from the respective numerical computations.

7.5 POSTSCRIPT ON STABILIZING SELECTION

In the mean time between the culmination of this thesis and a day before of sending to print, there has been substantial advance in the statistical mechanics theory regarding stabilizing selection. AS mentioned above, the main limitations are with respect to the computing times. There is a way to approximate the integrands of the partition function in terms of a Gaussian distribution. This approximation is based on the fact that

the partition function of directional selection, is the product of several independent per-locus partition functions (Eq. 7.10). Since these partition functions are in essence characteristic functions, their product is in the limit of large n a Gaussian function, as a consequence of the central limit theorem. The problem is complicated enough, and we are proceeding step by step. So far we have developed the method for arbitrary many loci of equal effects. Following this reasoning, we obtain that the partition function for the polygenic system is

$$\mathbb{Z} = \frac{\left(\tilde{\mathbb{Z}}^0\right)^n}{\sqrt{1+4N\alpha f_0}} \exp\left[\frac{2f_0 N b^2}{1+4N\alpha f_0}\right] . \tag{7.19}$$

From this formula, all the pertinent statistics follow. That is, the observables $\{\langle U\rangle, \langle \bar{z}\rangle, \langle \bar{z}^2\rangle, \langle \nu\rangle\}$, and the other macroscopics of the matrices C and B (Eq. 7.16). All the formulas are expressed in terms of simpler statistics of single locus of a trait without selection, but for which the genetic variance is selected for, namely the statistics generated by the partition function:

$$\mathbb{Z}^0 = \int_0^1 \exp\left[4N\mu\gamma^2 pq\right] (pq)^{4n\mu-1} dp . \tag{7.20}$$

and for which $f_0 = \langle \bar{z}^2\rangle_0 = 2n\gamma(1-\langle\nu\rangle_0)$.

Figure 7.9 shows how well the approximation is even for as few as for three loci, when compared to the exact integrations of the Fourier method.

The most radical test for our approximation is that when selection changes abruptly. For example a sudden shift of the optimum would trigger a quick response of the trait, and a radical reconfiguration of the genetic states. The prediction of the change of the trait mean and of the genetic variance is thus not a trivial task. In turn, our approximation allows to estimate

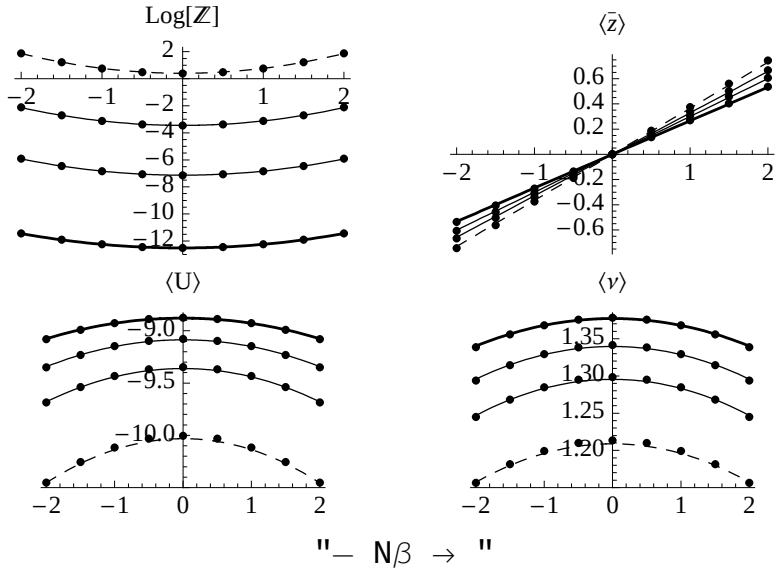


Figure 7.9: Comparing the exact Fourier integration (dots) with the Gaussian approximations for different mutation rates ranging from $N\mu = 0.3$ (dashed lines), 0.5, 0.7 (solid lines), 1.0 (thick solid line). $N\alpha$= -1, $N\sigma = 1$. The trait consists of 3 loci of equal effects = 1

the change of their expectancies, which give robust predictions of their evolutionary course. Figure 7.10 presents a comprehensive analysis of this situation. We compare this response with intensive calculations from the Wright-Fisher model, for distinct numbers of loci. Naturally, the response is quicker for more loci. Unlike the numerical effort required to compute the dynamics, it is reassuring that the precission of the approximation does not seem to depend critically on this number, except for very low number of loci (n between 3 and 5), where the covariances in the matrix C have significant deviations resulting from the Gaussian approximation (results not shown). These

deviations are insignificant for higher number of loci ($n > 10$). Yet the predictions of the macroscopics are in very good agreement with the numerical expectations fromt he Wright-Fisher Model, even for n as low as 4.

However, the change in the genetic variance is very low. In most cases, this change would be so tiny that it would pass inadverted in any practical situation. As it is shown in the previous figure, even when there are conspicuous changes in the mean trait, the changes in the genetic variance are minimal, less than 1% in all cases. (This by the way makes it not only hard but to some extenct pointless to attempt to have an accurate averaging from numerical realizations, and more critically from experiments). This should be compared with the variance from genetic drift fluctuations. Thus it is safer to compare the variance measures of the genetic variance (Var(ν) in an ensemble of populations (realizations) because these are more robust, and is a much more clear cut predition from the SM approach, and thus a way for falsification. After all, there it is meaningless to aim to predict such small changes in genetic variance when we would need an unrealistic number of populations to observe it. For the examples of Fig. 7.10, the fluctuations by drift are so big compared to the range of change of genetic variance, that (a) they completely mask these changes even when averaged for 10^4 realizations, a number of population replicas that is not only unrealistic to achieve even in experiments with micro-organisms, but barely enough to reveal that there is a change, with still an elevated degree of uncertainly, most specially for many loci (Table 7.1).

We can think also of examples where the strenght of selection changes, but without shifting the optimal phenotype. That is, deviations from the optimum become more critical. We can think for example of populations finely adapting to exploiting a particular resource, where a phenotype deviating from the opti-

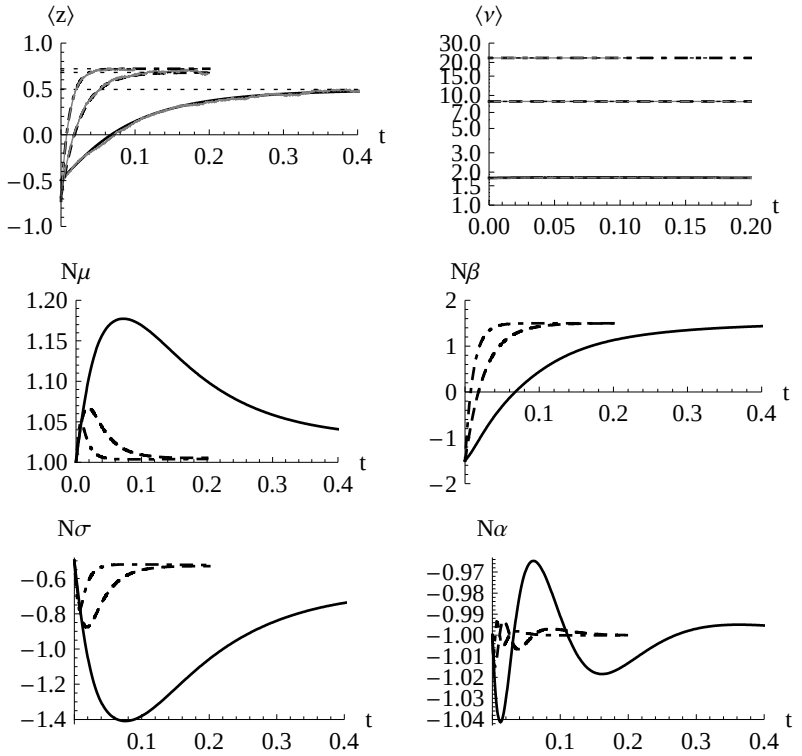


Figure 7.10: Evolution of a polygenic trait under stabilizing selection. Evolutionary response of the trait and genetic variance (top row), and of the local variables (mid and bottom row). Approximations for 4 (solid lines), 20 (dashed lines) and 50 (dot-dashed lines) are shown. The thin lines are averages of the Wright-Fisher model employing 500, 1000, and 5000 replicas respectively. The initial state is given by $N\mu = 1.0, N\beta = -1.5, N\sigma = -1.0, N\alpha = -1.0$, (equivalent to the optimum at $z_o = -1/2$), and $N\beta$ is changed to 1.5 (equivalent to a shift in the optimum to $z_o = -1$). The loci have equal unit effects.

mum (e.g. beak size in the Darwin finches) has less success in exploiting their main resources. The strength of selection would

Table 7.1: Fluctuations in the genetic variance. Traits with distinct numbers of loci (n, first column) show considerable fluctuations in the genetic variance due to drift. Second column, V_ν: statistical mechanical variance of the genetic variance. Third column, $\hat{V}_\nu$: numerical variance form the genetic variance (as in Fig. 7.10). These fluctuations are typically higher than the range of change of the genetic variance (fourth column).

n	V_ν	$\hat{V}_\nu$	Range
2	2.210^{-2}	1.910^{-2} (a)	210^{-2}
20	1.010^{-1}	1.010^{-1} (b)	610^{-3}
50	2.610^{-1}	2.510^{-1} (c)	310^{-3}

a) Variance from 500 replicas.
b) Variance from 2,000 replicas.
c) Variance from 5,000 replicas.

be mediated by a long number of factors, competition, predators, availability of the resource, time allocated to harvest, etc. Any of these factors coul alter the strength of selection without modifying the optimum in a significant way. In such cases, genetic variance changes radically. Some situations that we have successfully tried are when selection becomes disruptive, when selection ceases, or when it is intensified, in all cases without affecting the optimum. The range of change of $\langle\nu\rangle$ is beyond fluctuation by drift and hence the statistical mechanical method predicts accurately the evolutionary trend (data not shown).

As a last point, I would like to briefly comment on the relevance of the previous results. Although it might seem that dealing with the problem of equal effects lacks realism, it provides on the other hand a good tool to understand evolutionary mechanisms. A first example is to understand the dynamics of the microscopic reconfigurations (e.g. following a moving op-

timum; Bürger (2000, pp. 324–331);Jones et al. (2004); Kopp and Hermisson (2007). The allele combinations that are best fit to a given optimum are very sensitive and not linear with the value of such optimum. A small change in that optimal value may involve very different genetic states. The process that allow these changes are the jumps from peaks to peaks in the allele frequencies space (often confused with the fitness landscape, see Ch. 5). The properties of these jumps are in itself a complicated research subject (Barton, 1989; Nagylaki, 1989; Gavrilets and de Jong, 1993; Gavrilets and Hastings, 1993; Coyne et al., 1997; Rogers, 2003; Willensdorfer and Bürger, 2003). Treating the case of equal allelic effects allows to simplify very much the microscopic configurations and disentangle the details of the mechanisms. Because the statistical mechanics allows to have a relatively simple description of the evolutionary dynamics, we can then analyze these situations. A second example that is benefited from assuming equal effects is the contribution of genetic drift to the quantitative variation. Genetic variance evolves erratically due to (a) the rough path in the allele frequencies space, and (b) very unfrequent alleles that sweep in the population. Under unequal effects, these paths are smoothed, thus the contribution due to drift is entangled. Hence assuming equal effects allows to focus on the effects introduced only by drift (Barton et al., 2004; de Brito et al., 2005). A third example is epistatic effects. Again, the complications that epistasis introduce in the response to selection are on the one hand obscure with respect to their contribution to (cryptic) genetic variance (Kondrashov and Turelli, 1992; Gavrilets and de Jong, 1993; Carter et al., 2005; Beerenwinkel et al., 2007; Yukilevich et al., 2008), and on the other hand present a strong non-linear component that is amplified by the presence of alleles evolving due to distinct effective selective strength (Wang et al., 1998; Barton et al., 2004; Roff et al., 2006).

There are of course more examples to be listed, and these are a handful of relevant problems that we are willing to tackle on the framework of statistical mechanics. It is, however not an argument to forget the unequal effects situation. It is not as hard as it might seem on the first look. The Gaussian approximation leading to Eq. 7.19 is just using the central limit theorem. This *does not* require that the variables (allele frequencies) are identically distributed, only independent. Thus it is only a minor complication to extend the analyses for this more realistic situation, which -as that of equal effects- depends on the fact that linkage is not strong. That is the last point which needs to be considered in order to overcome the most stringent limitations between theoretical approaches, and biological reality.

Part III

Synthesis

Chapter 8

Synthesis

My specific goal is to revolutionize the future of the species. Mathematics is just another way of predicting the future.

Ralph Abraham

8.1 AN OVERVIEW OF THE CONCLUSIONS

8.1.1 Population dynamics

- An integration of population dynamics models was achieved. The per capita growth rate does not depend on the population size. Still it recovers determinate and indeterminate growth patterns such as: exponential, potential (hyperbolic and parabolic), logistic, θ-logistic, and Gompertzian. It also includes models of ontogenetic growth. It has other regimes consistent with several ecological and evolutionary models (although this needs further work to make a formal statement).

- The carrying capacity is not invoked as as a fundamental concept for density dependence. It arises from certain initial conditions and combination of parameters. The conditions not leading to carrying capacity (i.e. models of indeterminate growth) are still biologically plausible. Furthermore, assuming a carrying capacity might limit our understanding of the ecological processes.

- General patterns of scaling laws were found, which concerns not only population dynamics but also ontogenetic growth laws. This broadens our view on the question and usefulness of scaling as a tool.

- There are further ways in which populations can be randomly affected, more than just environmentally or demographically: perturbations affecting the inter-specific interaction parameter θ, depending on the "noise to signal" relation, will result in populations that resemble either (a) logistic growth, or (b) exponential growth. On the one hand, this provides a further explanation for the ubiquity

of these growth laws. On the other hand it compromises mechanistic explanations for these patterns. Mechanistic explanations would have little use, in contrast with an understanding of the sources of the fluctuations. In the logistic cases, the stable size is uncorrelated with the deterministic carrying capacity. This is further evidence that threatens the idea of carrying capacity.

8.1.2 Population Genetics

- The idea of *entropy* was formalized for population genetics. It is a measure that is maximized at equilibrium, and accounts for the expected contribution to the evolutionary potential by selection, mutation and drift.

- Based on the previous, a coupling between population and quantitative genetics was achieved through an analogy with statistical mechanics. The methodology is general, but explicit results were performed for additive traits (including dominance effects) under directional or stabilizing selection, and for multivariate traits with pleiotropic effects, subject to directional selection.

  The coupling is not restrictive on the effects and number of the loci, dominance, pleiotropy, or epistasis. It depends however on Hardy Weinberg and linkage equilibrium, diallelic loci, and constant population sizes. It is essentially frequency-independent.

- This method avoids the arbitrary choice of the quantitative variables that are needed to track evolution. It gives a neat way to choose the variables needed to track evolution.

- Knowledge of the allele frequencies is not required to make predictions. Alternatively the predictions are for the ex-

pectancies of the quantitative variables, which only depend on macroscopic quantities.

- It is possible to make long-term predictions of evolution, provided that we know the breeding values, the mutation rates, the size of the population, and the strength of selection. But we do not require to know the allele frequencies at any locus. This is particularly true for the values of the traits in an 'average' population, but also for the genetic variance, the $\mathcal{G}$-matrix, and any other quantity that depends on the allelic effects.

- Specific results:

  - For directional selection, high mutation rates ($4N\mu > 1$) and drift, the quasi equilibrium dynamics assumption and the statistical mechanical approximation are accurate in predicting the evolutionary course of polygenic traits.
  - The statistical mechanical method has to be modified for very low mutation rates ($4Nm \ll 1$). For intermediate mutation rates ($4Nm \simeq 1$) and directional selection , the inclusion of dominance effects allow a correct coupling between the micro and macro states.
  - The extension to multivariate traits was achieved by including several traits as observables. This allows pleiotropic effects to be included.
  - The change in the $\mathcal{G}$-matrix can be computed following the statistical mechanics methodology, and it includes the action of selection, and mutations, averaged over drift.
  - $\mathcal{G}$ is much more sensitive to drift effects than to selection. Mutation does not change the direction of evolution, although it affects the rates of change of $\mathcal{G}$. The

pleiotropic effects over $\mathcal{G}$'s eigenstructure are more noticeable at high mutation rates; strong pleiotropism has only a minute effect over $\mathcal{G}$ if mutation rates are low.

8.2 FURTHER EVOLUTIONARY IMPLICATIONS

The most pervasive trait of populations –in the broadest sense– is that their sizes dynamically change and adapt to specific ecological conditions. In itself population growth is a cornerstone of the evolutionary theory. It is so conspicuous that the evolutionary mechanisms responsible for fixating a given strategy of growth pass inadvertent, and typically these patterns are often assumed as an intrinsic property of an organism. For example, the concept of Malthusian rate of growth is typically used in several evolutionary theories in order to evaluate whether a mutant will invade a population or not. Typically, this approach goes in hand with the assumption that populations remain in their carrying capacity along the evolutionary process. In this sense, the question of quantity is deferred, focusing on the question of quality, that is whether the mutants perform better than the residents. This is the view from the game theoretical and adaptive dynamics theories. A second field that employs equivalent assumptions is population genetics. The difference can be on the recursive nature of the mutants. But when these mutants are rare, the evolutionary analysis is equivalent to that of adaptive dynamics. Life history theory on the other hand, provides explanations on what determines this Malthusian rate of growth, $r_{\max}$, interpreted as the maximal rate of increase of a rarified population on ideal ecological conditions (Fisher, 1930; MacArthur and Wilson, 1967; Stearns, 2004; Charnov, 1993). Thus these three theories -adaptive dynamics, population ge-

netics, and life history theory- seems to build a fairly rounded-up picture of evolution.

The mathematical models of population dynamics often employ carrying capacity as a mechanism to regulate populations, modeling the ecological constrains on growth. The canonical model, the logistic equation, predicts that the population will attain equilibrium at the carrying capacity. When growth is determined, "equilibrium" and "carrying capacity" become one and the same. However, the notion of carrying capacity, defined in terms of the equilibrium of a population can often be misleading and ill-defined. With ill-defined I mean that carrying capacity ambiguously takes as equivalent (i) the equilibrium size of the population and (ii) the maximum number of individuals sustained in the environment, determined by ecological factors. This, although a common practical equivalence, can be regarded as non-scientific. Defined as above, we cannot distinguish whether the population reaches equilibrium because the environment is saturated, or the environment has been saturated *because* the population has reached equilibrium. There are obviously numerous examples against my statement, where it is well determined how the ecological constraints determine a population's carrying capacity (e.g. life history theory approaches the problem in several ways). But the concept, being a cornerstone of population dynamics and genetics, tends to be more phenomenological than mechanistic. I consider that it is possible to study some general properties that determine not only carrying capacity, but also other traits that determine growth.

To start, the Malthusian rate of growth, since the seminal works of Sir Ronald Fisher (1930, Ch. 2), has been conveyed with genetic structure. But other descriptors of growth –like carrying capacity– are not so easily equated into evolutionary and genetic terms. The Malthusian fitness is a very natural

measure of growth, and under fixed population size it links the genetics to a phenotype's rate of growth in simple terms. Other traits involved in growth regulation have to invoke more specific mechanisms in order to be equated to genetic variables in a coherent way. Otherwise, we might work in terms that are very general, but give little evolutionary insight. To follow, since the decomposition of the population growth dynamics into population size and growth rates involves only two parameters in a linear way, it immediately provides a simple set-up to consider evolutionary implications. (Not that a different formulation would not allow it, but the simplicity and generality of the model opens the possibility for simpler evolutionary analyses.)

The question in general terms, is which evolutionary and genetic properties are present in a reproducing system, which determine a growth pattern. The idea can be approached from distinct sides of the evolutionary theory. I will digress in three directions: life-history, invasion analysis (game theoretical approach), and population genetics.

Life history

To begin with, I will give an example on how we can set-up the growth equations in an independent way of the carrying capacity. The first step has already been achieved in the size-rate decomposition, where an explicit dependence on the carrying capacity N_∞ has been achieved[1]. We could say that N_∞, the equilibrium size, is 'hidden' in the initial conditions or the per-capita rate –as I showed in Eq. 1.20, but beware that this equation is not general, since there might not be a carrying ca-

[1]Note that I changed the notation with respect to Ch. 1: I will use now N to denote population size, rather than x, in order to give a more intuitive association to the symbol, and which hopefully avoids miss-understandings in the following equations.

pacity at all. But we can take a different route, and relate the initial growth rate to life-history parameters. For instance, at very low population sizes where the density dependence is the weakest, the rate will be maximal (that is a perturbation of r form the fixed point ρ/θ, Eq. 1.9). Instead of equating the initial rate of growth to a carrying capacity as in Eq. (1.20), we can borrow the interpretation form life history, an equate it to r_{max}, which for textbook examples we can take as a function of other life history parameters (e.g. $r_{\max} \simeq \log(R_0)/T_c$, where R_0 is the lifetime reproductive success, and T_c the generation time; Charnov (1993) Eq. 6.6, pp.118). Hence, equating the initial rate of growth as $r_{(0)} = r_{\max}$, then it follows from Eq. 1.5 that

$$-\alpha = \left(\frac{\theta r_{\max}}{\rho} - 1\right) N_0^{-\theta} = \left(\theta \log(R_0) - 1\right) N_0^{-\theta} \tag{8.1}$$

Now, introducing into 1.5, we get:

$$r_{(t)} = \frac{\rho}{\theta}\left[1 + \left(\theta \log(R_0) - 1\right)\left(\frac{N}{N_0}\right)^{\theta}\right] \tag{8.2}$$

Thus mutants with distinct strategies of ρ, θ and R_0, might exploit the ecological constraints in different ways allowing distinct patterns of growth can be evolutionarily stable, and which may –for example– show different limiting sizes N_∞. Consider the point N^* at which the maximal change of speed of the per-capita rate is attained, that is $d(Nr)/dN = 0 \rightarrow r^* = \rho/(1+\theta)$. Introducing this into Eq. 8.2, we get that

$$1 = (1+\theta)\left(\theta \log(R_0) - 1\right)\left(\frac{N^*}{N_0}\right)^{\theta} \tag{8.3}$$

I assumed that the product ρT_c –the average number that a female would produce in a rarified condition– is of the order of 1. The last expression has some implications. First, if we fix θ,

then we have a relationship between R_0 and N^*. It is convenient to measure N^* relative to the carrying capacity: $\eta = N/N_\infty$, so that the term $N^*/N_0 = \eta^*/\eta_0$. If we introduce and re-arrange Eq. 8.3, we have that

$$\eta^* = \eta_0 \left[(1+\theta)[\theta \log(R_0)]\right]^{-1/\theta} \tag{8.4}$$

If we map to a semi-log scale, we can express this relationship as

$$\eta^* = A - B \log\log(R_0). \tag{8.5}$$

This relationship is known in life-history theory as *Fowler's rule*, explaining that species that share anR_0 will also share η^* (Fowler, 1988). Fowler identified this relationship in an empirical way, and to my knowledge it has not been related to θ-logistic models in the way I have presented. This is relevant, because Sibly et al. (2005) analyzed a substantial data set of population dynamics in order to measure the distribution of the intra-specific competition coefficients, and have shown that the parameter is distributed over a large range (± 100) for distinct taxa: mammals, birds, fish, and insects. The distribution is different across taxa, and remains to be explained. Yet it suggests that Fowler's rule is a coarse-graining of a more detailed version as the one I have presented. Thus, if we plot different curves with different values of θ in the range reported by Sibly et al. (2005), and superimpose Fowler's data (listed in Charnov, 1993, pp.), we can see that considering additional information, that is the value of θ, can give more precise explanations of the data (Fig. 8.1 A). In other words, the competition parameter can explain the deviations from the main trend.

Second, if we fix the point of inflection N^*, then there is an inverse relationship between the degree of competition and the life-time reproductive success. If R_0 is increased, then θ is diminished (intra-specific competition decreases). As a by-product then, the amount of individuals that are maintained

in the population is also increased (the limiting size is bigger). If on the contrary, competition is high, the life time reproductive success decreases, and there is, in equilibrium a smaller population size. But, is there any invariant involved in this relationship? Indeed, some algebra leads to the relationship

$$\log\log(R_0) = A_2 - \log(\theta)\ . \tag{8.6}$$

Formally A is between 0 and $\log(2)$. The choice $A = 0$ gives a perfect fit for big values of θ (as the trend shown in Fig 8.1 B), whilst the choice $A = \log(2)$ fits better the lower values. In any case, the deviations are not big in the log-log scales, and an unbiased survey of data would give probably $A \sim \log(2)/2 = 0.35$. I am enthusiastic and eager to go forward to check my predictions!

Game-theoretical approach

The common approach to density dependence is that it is determined by the ecological conditions. Because the environment would itself be able to support only a maximal amount of individuals, then a carrying capacity would be established. This could well be in many (or even most) situations. But I would like to speculate in a different direction. If a population that does not have a density dependence mechanism that leads to a limit size "consistent" with ecological constraints, then this population would eventually go extinct. This can be most easily pictured with an island model as a proxy (MacArthur and Wilson, 1967), allowing study the evolution of growth rates. For example if mutant individuals that are able to adapt their growth response to the ecology in a coordinated and coherent way invade a resident population, then these new variants would (by definition) be better adapted, and their extinction would be less

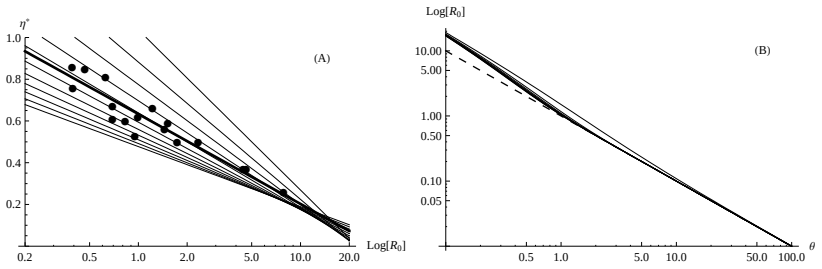


Figure 8.1: Life-history invariants related to growth parameters. (A) Fowler's rule points an inverse relationship (in semi-log scale) between life-time reproductive success, and the point at maximal growth rated. His fit (dark line) to the data (black dots) is given by Eq. 8.5 with $A_1 = 0.633$ and $B = 0.187$. The thin lines show the same relationship but with A_1 computed explicitly with distinct values of θ in the range $(3, 8)$ (spaced every 0.5 units). (B) A predicted invariant between the intra-specific competition parameter θ and life-time reproductive success, in log-log scale. The invariant relationship is surprisingly simple, and independent of any other parameter. The dark dashed line is a simple approximation (Eq. 8.6), whereas the thin lines are the exact values employing a range of inflections N^*/N_0 in the range $(1, 16)$. In this case $A_2 = 0$ (see text for explanations).

likely. Also, they would out-compete any new invader, and dispersing to other island, they would be established.

Thus, to evaluate if the mutant will invade, we can compare their Malthusian rates (Maynard-Smith, 1999). If for example the density dependence affects the mutant equally as any other resident, then the invasion analysis reduces to compare the ratios

$$\frac{\rho_M}{\theta_M} > \frac{\rho_R}{\theta_R} \tag{8.7}$$

where M refers to the mutant strategy, and R to the resident strategy.

An interesting feature, is that if a mutant appears that has θ close to zero (a mutant that is not strongly competitive, or cooperative, with its mates), then it will most likely invade. That

mutant, will of course have a Gompertzian growth rate (because $\theta = 0$ corresponds to this strategy, see Ch. 1). If that is the case, the question is why Gompertzian growth is not conspicuous in populations? Basically, I am just saying that the mutant can invade, which does not mean it can actually replace the resident. That would depend on how the mutants and residents interact. But also, there might be constraints that keeps θ fixed, like strong selection to maintain some degree of competition (e.g. sexual selection, or limited resources).

Gompertzian growth implies that there is no intra - specific competition, yet there is density regulation. At this stage, it matters little what the Malthusian rate of growth is: a Gompertzian growth will outcompete any other determinate population growth strategy. What happens once a Gomperzian strategy has invaded? Naturally then, evolution will favor larger Malthusian parameters. I am of course assuming that there are no trade-offs between ρ and θ. It is not difficult to develop in such direction, but under linear trade-offs it is easy to see that still a Gompertzian strategy will invade, and the Malthusian rate would be maximized.

The analyses of Sibly et al. (2005) have sown that the competition parameter θ has a broad distribution. Although there is a substantial proportion of strategies with $\theta \simeq 0$, other values are frequent. Thus my analysis is clearly missing something. First of all, I have computed only the probability of invasion, and not really analyzed whether the mutant will fix, and reach an evolutionarily stable strategy. Also I have ignored other ecological factors. As simple analysis shows that inclusion of other factors will not lead to distinct results. For example, in an island model, where there is certain probability Px of becoming extinct, then the fitness has to be weighted by these chances.

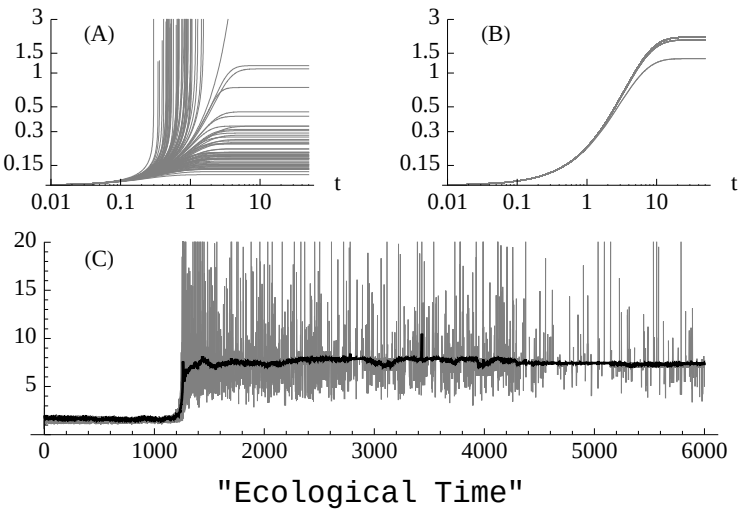


Figure 8.2: Evolution of population patterns. 100 random populations strategies, (ρ, θ) were allowed to evolve for 6000 ecological episodes (events of extinction, re-colonization, or invasion). (A) Dynamics of the initial founder populations. (B) Dynamics of the evolved strategies (after 6000 ecological episodes). In both cases, time is in generations. (C) Sizes of each population once they hay reached equilibrium after an episode –i.e. "carrying capacities" of each population– (gray lines), and the average size of the whole pool of populations (black line), time is measured in ecological episodes.

In short, the invasibility is determined by

$$(1 - Px_M)r_{MR} \geq (1 - Px_R)r_{RR} \tag{8.8}$$

The chances of survival $(1 - Px)$ are related to the size of the population, life-time reproductive success, etc. An thus we can expect that consideration of a more structured model accounting for these factors, can give predictions on how the patterns of growth can be fine tuned. In Fig. 8.2 I performed a simple simulation, where a population has a given chance of becoming

extinct if it is below a certain size, or above a limiting environmental sustainable capacity. The populations are able to evolve their strategies every generation, and if the mutant is able to invade, then it replaces the resident. The value of $r_{\max}$ is fixed, but together with the strategy (ρ, θ) determines the value of N_∞. The survival is decided only after populations have achieved equilibrium. Some combinations of these parameters will lead automatically to extinctions or to explosion, as explained in Ch. 2. In both cases extinction is certain. The results show that (a) the populations do not typically evolve to have a carrying capacity that matches the maximal ecological sustainable density, but rather reach a state of mutation-selection balance, and (b) the populations tend to diminish their intra-specific parameters quite close to zero.

A substantial deal of work is required to gain a clear understanding of the evolution of patterns of growth. Most critically, the way how the mutant and the resident interact will most surely change the Gompertzian outcome.

Genetics

Regarding a growth strategy as a quantitative character can be a complicated matter. Naturally, the distribution of the characters are determined by the frequency of the alleles. But since the evolution of a structured population is driven by its phenotypic distribution, its change feeds back into the distribution of allele frequencies. Ignoring mutation, linkage and drift, Slatkin (1980) studied the effect of population competition on the displacement of a quantitative character. The replication rates (fitness) depended directly on the fitness of the character. In this model the genetic variables were not tracked, but instead it was assumed that both the trait and the genetic variance were changing in time, and thus determining the broadness

and position of the distribution of the trait. Doebeli (1996b) extended the model to include the genetic variables, showing that the distribution of genetic effects will be neither Gaussian, nor in demographic equilibrium. This of course raises the question on how robust is the assumption in population genetics (taken along this thesis) that populations are at a demographic equilibrium. But perhaps most important is the question on how sensitive are the conclusions of population genetics with respect to the demographic equilibrium assumption.

It is possible to extend the Wright-Fisher model studied before to the case where population size is changing. For instance, calling n the number of favorable alleles, then the frequency is given by $p = n/2N$. If we compute the rate of change then we obtain

$$\dot{p} = p\left(\frac{W}{\bar{W}} - 1\right) - pr \tag{8.9}$$

where r is the per-capita growth rate of the population, as above. Thus a quantitative trait that does not affect the growth rate directly will still have some transmission bias due to the effects of population growth. At demographic equilibrium the growth rate vanishes ($r = 0$), and the dynamics would proceed normally through a Wright-Fisher model. If population size changes slowly, then obviously there will be no big deviations from the condition of demographic equilibrium, because the genetic variables mix quickly. However, the situation is typically the contrary: populations tend to change faster than the rate at which mutants appear.

If the dynamics of the population are stochastic, then we can study the joint distribution of allele frequencies and population size. This is not so far from the previous situation, so no particular deviation is expected. In principle, we can easily propose a quasi-equilibrium model, where the local variables $N\mu^*$ and $N\beta^*$ follow a continuous change in N. This should work fine

unless $N\mu^*$ falls below $1/4$. As we saw in Ch. 3, fluctuating selection intensities will still give allow a good approximation from the statistical mechanics methodology. Thus from that side, there is a wide margin for confidence on the demographic equilibrium assumption.

Most interesting is the case where the traits involved in growth (Malthusian and intra-specific competition parameters, as well as "carrying capacity" or any other associated life-history trait) are affected by many genes. As Slatkin (1980); Doebeli (1996b) did, we can define the growth rate of the population as the mean fitness, and then confer it with genetic structure. Equation 8.9 is suggestive of this situation, since from the gene's point of view, selection and growth rates are two kinds of the same. Defining growth rates directly as mean fitness will determine certain patterns of growth, but it will not provide self-regulation. There is still an effect of population size, because it plays a role in the transmission of alleles to the next generation. Nevertheless, the demographic path will depend *only* on the nature of the selective process triggering evolution.

But to conclude the analysis of the evolution of population dynamics, we can consider the genetic effects over the parameters of growth. In that case the rate equation (1.7b) has some extra terms, because its rate of change would be influenced by the rate of change of ρ and θ. In that case, it is not easy even to separate the time scales, because there are two factors affecting r's rate of change. First the density dependent component, which has been studied in the first part of this thesis, and in the above sections of this synthesis. Second, the quantitative rates of change. As indicated by Eq. 8.9 there is a directional effect. This mutual feedback is the source of many complications. So questions about variability maintained by a density dependence term, are not trivial to deal with, but have an interesting richness of properties (see Bürger and Gimelfarb,

2004, for an example of stabilizing selection over a quantitative character mediating competition, and even leading to speciation; Nagylaki, 1979).

For a moment, we can forget about selection, and consider only mutation and drift. This not quite a neutral process, because there is the selective effect of the population dynamics:

$$\frac{d\bar{z}}{dt} = -(2\mu + r)\bar{z} + \text{"drift"} \quad .$$

Where "drift" is Gaussian and has variance proportional to $1/2N$. If the dynamics of the population size is slow (e.g. close to an equilibrium), and the mutation rate high enough ($\mu \gg r/2$), then the rate of change of the trait would be a stochastic process with a stationary Gaussian distribution[2]:

$$\bar{z} \sim \frac{\sqrt{\mu/\pi}}{2N} \exp[-\mu \bar{z}^2/2N].$$

If the trait in question is the parameter θ, then these fluctuations would be of the kind studied in Ch. 2. Then genetic effects over the intra-specific competition parameter are a parsimonious explanation for this kind of stochastic effects leading to density regulation. Of course, if the character is entirely neutral, then there is no mechanism to control population equilibrium size. We would expect, on the other hand, that if selection for this parameter exists, then the mechanism of density dependence would be genetically determined, and given the noise-suppression nature of the dynamics, and would not require a fine tuning. Unfortunately, the regime $\mu \gg r/2$ is not a biological paradigm. Rather the contrary is expected to happen. There is a situation, however, where we expect to find a high mutation rate, which I analyze in the following section, and that is on the evolution of prebiotic replicators.

[2]This would be essentially what is called an *Ornstein-Uhlenbeck process.*

Replicators: density dependence in the origins

In the early stages of life's origin, the proto-organisms would lead the prebiotic world in a democratic way. Virtually any possible mutant that appeared would survive (Szathmary and Gladkih, 1989). These prebiotic entities, essentially replicating RNA molecules, would need of each other to replicate, but in a way that any variant molecule would be able to catalyze the replication of any other molecule. Eigen (1971) has argued that these replicators could not have grown exponentially, because the diversity required for maintaining the evolution of a primitive population of replicators, would simply collapse. This lead to the ideas of hypercycles Eigen (1971) and the stochastic-corrector (Szathmary and Demeter, 1987b), which provide solutions to the information collapse problem, but themselves are liable to other problems of invasion of selfish mutants.

Nevertheless, although these ideas are alive and explain the transition from replicators to protocells (Maynard-Smith and Szathmáry, 1997, Ch. 4), a solution to the replicator information collapse came from a simpler idea. Based on experimental models –or 'artificial replicators'–, (von Kiedrowski, 1986; Scheuring and Szathmary, 2001) found that the rate of growth of the RNA population would be parabolic. Under this assumption of parabolic growth, Szathmary and Gladkih (1989) showed that diversity of RNA molecules would be attained and sustained.

However, although suitable for experimental essays, the system is assumed to be in demographic equilibrium (Eigen, 1971). That is that the total size of the population is constant. It is unlikely, however that there were mechanisms of density dependence in such a primitive world. This assumption of constant population size relies on the fact that the building blocks were a limiting factor. There are reasons to think that this could

have been if the rate of abiotic production of building blocks, i.e. nucleotides, would be slow. But the contrary situation is also plausible, that the building blocks were not a limiting step, specially if the population sizes were rather small and the environment was rich.

Irrespective on the competition details of the replicator system (it might well apply with changing population sizes), indeterminate growth would not be viable. Parabolic growth is indeterminate. Given that reaction catalysis times (on the order from seconds to days) are much faster than evolutionary time, then an explosion of the population size would have occurred practically instantaneous. This sound unlikely, since we would have to explain the transitions to compartmentalization and emergence of mechanisms of density regulation within such a short period. The rate of abiotic production of nucleotides would have been to slow as to allow evolution for very quick reactions if parabolic replicator were to grow to high densities. Here there are two possibilities. Either the dynamic would be driven by the decay and sequestering of the components of unreplicated decaying molecules (Scheuring and Szathmary, 2001), or a mechanism to regulate the growth rate had to be present (or of course both). The former case has been analyzed under the typical population-genetical assumption that the size of the evolving population is constant. Now, I try to give some insights on how this constant size could have been achieved. I argue that the effect studied in Ch. 2, that noise in the parameter θ would stabilize the size of the populations, could have played a role. The question of course is what is the source of this noise. My argument is simple. Although we can assume that in average the order of the replication reaction is 1/2, we can also think that this value is determined by intermediate states of the replication reaction. Thus distinct RNA mutants might have different values of θ, which depends in a

complicated way on the specific sequence. Thus treating this parameter as a quantitative character, then the expected genetic variance would be, following Eq. D.46, $\langle \nu_\theta \rangle = 2\theta\mu/\beta$. As shown in Ch. 5, if fixation of alleles is unlikely, then the character would be normally distributed. Therefore, the strength of the population-genetical fluctuations over the growth rate (the genetic variance of θ) would allow a moderate rate of growth and even an equilibrium if the rate of decay of the molecules would show significant fluctuations.

Now, if two parabolic replicators are competing, the new mutant is able to invade only if $\theta_M < \theta_R$. The later course (whether there is coexistence or displacement of the resident) depends on further details on the density dependence (Mylius and Diekmann, 1995; Metz et al., 2008). Nevertheless, in general terms, selection would favor lower values of θ. A parabolic replicator of exponent $\alpha < 1$ has a per-capita growth rate with exponent $\theta = \alpha - 1$ (Ch. 1), so in general $\theta < 0$.

For parabolic replicators we would need a genetic variance of less than one. In this way, the exponent would be kept between 0 and 1 in average. Otherwise the replicators would cease to be parabolic and we would incur in the domain of the error threshold.

For example, if we were to assume that (1) there are n_e nucleotides which actually affect the value of θ, (2) the mutation rates were relatively high, say $\mu \sim 10^{-3}$, (3) that selection is weak $|\beta| \sim 10^{-2}$, and (4) that $\theta \sim 1/2$ then the genetic variance would be on the order of 10^{-1}. This means that $\epsilon_0 = 50$ (in Eq. 2.2). The rates of decay of the RNA molecules are assumed to be low, but with enough fluctuations (assumed to be environmental), then stabilizing the size of the population would require a variance in the decay rate of the molecules on the order of 10^2, which is too big. On the other hand, if mutation is higher, and or selection lower, the genetic variance can be increased to at about 0.4, which would require a variance on the decay rate of the molecules of about 0.1, even for low decay rates. This is consistent with the co-existence of multiple replicators.

If the genetic variance of θ were too low, for example because selection is strong, then the replicators would not be in a demographic equilibrium. Still, the population as a whole would grow exponentially, rather than exploding in short time.

8.3 RESEARCH PERSPECTIVES

The mechanisms that drive evolution in the wild –natural and sexual selection as well as mutation and drift– are of the same nature at the micro and macro-evolutionary scales. Nevertheless, population and quantitative genetics aim to understand and predict the diversification of populations into two lineages, that is speciation. One of the problems, which I have addressed since the beginning of my dissertation, is that of variability. The state of the art in understanding genetic variability is to some degree, embarrassing. Theoretical models predict much less variability of what is observed in the wild, and the predictions of the evolutionary course are limited to a few generations. Yet we are attached to our current frame of mind, where we continue to use the very same tools that keep failing. If we compare how difficult is to detect selection at a locus with how easy is to measure molecular substitution rates we find that The Neutral Theory of Molecular Evolution (Kimura, 1985) has provided much more pragmatic use than evolutionary and quantitative genetics, even when we are certain that natural selection is the main cause of diversification in the tree of life. This is exciting, because it reveals that even when our current knowledge accounts for micro-evolutionary processes in an accurate way for artificial selection and experimental evolution, the "natural versions" of evolution seem to be hiding something to us. This is not a trivial subject. There are many lenses between the stars that we want to observe, and the light that gets to our eyes.

The amount of assumptions that we make in order to describe an evolutionary response, or to detect natural selection is not negligible.

1. First of all, we typically assume that populations and species are at evolutionary and demographic equilibrium. This translates into a huge bias in what we are able to mea-

sure. We saw in the second part of this thesis that during transient evolution, we might describe the evolving distribution by local parameters. This means that if at a given moment of time we sample a population and intend to characterize evolutionary factors (mutation rates, effective population size, selective gradients, and even breeding values) then we might have a very wrong picture. For example, we might find dominance effects that are not really there, or miss-estimate the size of a population from genetic markers.

2. Density and frequency dependent effects, as discussed above, non only may have selective consequences that have not been considered in typical quantitative genetical empirical studies, but also coupling selection to population dynamics has major impact on the quantitative measures, affecting the rate of change of quantitative characters.

3. The breeding system, which I have not addressed in this thesis, and sexual selection will entirely change the picture of what we are measuring in quantitative genetics.

4. We are assuming that the genetic effects are constant. The theory of reaction norms have provided clear cut evidence that the distribution of phenotypes change with the environment. There is of course possibility to include genotype-environment interactions that could account for the reaction norms. Nevertheless, whether these interactions are the source of the phenotypic changes across environments has not, to my knowledge been shown. There might be an analogous effect as with the genetic co-variances: because there are statistical associations in the evolutionary response to selection it does not imply that there is at all a genetic association. Interestingly, because the

way how we can identify the pleiotropic QTL's is employing correlated responses, if the genetic effects are not constant across environments the genetic associations could be a by-product of the covariant effects.

This essentially translates in that the effects in the wild environments might actually respond in a totally different way than in the lab, or under controlled experimental conditions. Thus there is little that we can say for the long-term evolutionary response out from association analysis derived form captive populations, experimentally derived estimations (populations that are back-crossed, randomly mated, developed under controlled conditions, etc.) QTL's are part of the optimism for modern quantitative and population genetics, since they give a concrete meaning to the reification of the concepts of locus and allele. Not that these are mistaken, but by their phenomenological nature, they are certainly limited, and often need to be narrowed down (in the best of the cases) to single nucleotide substitutions, in order to have a mechanistic understanding of their effects.

Simple mendelian traits behave very well, and there the concepts of loci and allele work very well. Also the nucleotides in molecular genetics fit very well these ideas. Still, *polygenes* is a very abstract entity. Sometimes their segregation will be reducible to several Mendelian loci, but sometimes not. The number of loci might be very variable, and the number of alleles might as well not be a fixed trait. If their effect is infinitesimal, we are in a safe place, since we can model the effects in a simple and convenient way. Yet in that case the predictions don't go beyond micro-evolution. If the effects, on the other hand are not infinitesimal, the population's course will be sensitive

to the number of alleles and loci. This is rarely taken into consideration.

The effects of modifier loci, although they provide a good theoretical aid, are seldom used empirically to predict and forecast consequences on long term evolution.

In general, most components of the genetic architecture tend to be neglected at the time of quantitative estimations. Although these architectonic elements are invoked often in the practice, they are mostly viewed form the molecular perspective and in a neutral context.

5. There is an underestimation on the meaning of selection differentials. A clear understanding of its meaning is not only essential, but also necessary. A selection differential, relies essentially in another statistical association: the correlation between fitness and the trait. For example, so much work is done on the $\mathcal{G}$-matrix, trying to understand its response to selection, mutation, drift, migration, etc. But not only the knowledge of $\mathcal{G}$ is of limited interpretation (Pigliucci, 2006), but also in the absence of a concrete measure of selective gradient, $\mathcal{G}$ would say little. From the empirical side, the gradient of selective values is never computed independently of $\mathcal{G}$ (or H^2, for these matters). In theoretical terms, except in controlled experiments of truncation selection, it is unlikely that we know anything except the local gradient. The result, is that we can only make qualitative predictions. But if we are to be happy with qualitative predictions, then it would make little sense to invest big efforts in accurately determining the genetic structure of a population, if we intend to infer anything about evolution.

6. The last two arguments point to something. If the genetic

effects were constant and the reaction norms would entirely be explained by the genotype-environment interaction, and the micro-evolutionary mechanisms would suffice to explain in the long run macro-evolution, then we would expect to find certain degree of correlation in speciation rates in the bio-geographical gradient. Perhaps this correlation exists, but it is hard at the moments to predict at which (taxonomic) level they should be observed. There are conspicuous cases, like the glaciation periods, where major ecological changes have taken place. But smaller ecological episodes should have not only a notable effect over speciation processes, but also promote it in several lineages.

I would expect this to be so in the case of fixed genetic effects in part because of the universality of the genomic content, in the sense that there are so many genes common (even at distant phylogenetic distances), that at least in traits which are conformed by common genes (e.g. FOX1) there should be correlated responses, at least in particular geographic locations. Honestly I doubt that this would be so. But only because I doubt that these micro-evolutionary mechanisms can be extrapolated to long term under the assumption of fixed genetic values.

Nevertheless, with the local equilibrium approach, we have been able to predict the long term evolutionary response of the genetic co-variances. But these are relying on a constant architecture, a constant population size, and that we know where selection is pointing to. Even with the limitations, this in any case is an improvement, and a constructive methodology, and for which of course I stand. Yet, pointing the failures above is at most useful as a progressive method to see where the limitations are.

These limitations are at best conditions for artificial selection. But if we are to ground the mechanisms of the evolutionary theory only on micro-evolutionary insights, then we might be leaving the back-door open for intelligent absurdities of design, which is a door that we need to close for good.

Bibliography

Bibliography

Abramowitz, M. and Stegun, I. A. (eds): 1972, *Handbook of Mathematical Functions*, Dover Publications Inc., New York.

Ai, B., Wang, X., Liu, G. and Liu, L.: 2003, Correlated noise in a logistic growth model, *Phys. Rev. E* **67**, 022903.

Aita, T. and Husimi, Y.: 1998, Adaptive walks by the fittest among finite random mutants on a Mt. Fuji-type fitness landscape, *J. Theor. Biol.* **193**, 383–405.

Aita, T. and Husimi, Y.: 2003, Thermodynamical interpretation of an adaptive walk on a Mt. Fuji-type fitness landscape: Einstein relation-like formula holds in a stochastic evolution, *J. Theor. Biol.* **225**, 215–288.

Aita, T., Morinaga, S. and Husimi, Y.: 2004, Thermodynamical interpretation of evolutionary dynamics on a fitness landscape in a evolution reactor, I, *Bull. Math. Biol.* **66**, 1371–1403.

Aita, T., Morinaga, S. and Husimi, Y.: 2005, Thermodynamical interpretation of evolutionary dynamics on a fitness landscape in a evolution reactor, II, *Bull. Math. Biol.* **67**, 613–635.

Albert, A. Y. K., Sawaya, S., Vines, T. H., Knecht, A. K., Miller, C. T., Summers, B. R., Balabhadra, S., Kingsley, D. M. and Schluter, D.: 2008, The genetics of adaptive shape shift in stickleback: Pleiotropy and effect size, *Evolution* **62**(1), 76–85.

Allee, W.: 1931, *Animal Aggregations. A study in General Sociology.*, University of Chicago Press, Chicago.

Anteneodo, C. and Tsallis, C.: 2003, Multiplicative noise: A mechanism leading to nonextensive statistical mechanics, *J. Math. Phys.* **44**(11), 5194–5203.

Ao, P.: 2005, Laws in Darwinian evolutionary theory, *Phys. Life Rev.* **2**(2), 117–156.

Ao, P.: 2008, Emerging of stochastic dynamical equalities and steady state thermodynamics from darwinian dynamics, *Commun. Theor. Phys.* **49**(5), 1073–1090.

Arnold, S. J., Bürger, R., Hohenlohe, P. A., Ajie, B. C. and Jones, A. G.: 2008, Understanding the evolution and stability of the G-matrix, *Evolution* **62**(10), 2451–2461.

Arnold, S., Pfrender, M. and Jones, A.: 2001, The adaptive landscape as a conceptual bridge between micro-and macroevolution, *Genetica* **112-113**, 9–32.

Ayala, F., Gilpin, M. and Ehrenfeld, J.: 1973, Competition between species: theoretical models and experimental tests, *Theor. Popul. Biol.* **4**(3), 331–5.

Barton, N. and Coe, J.: 2009, On the application of statistical physics to evolutionary biology, *J. Theor. Biol.* .

Barton, N. H.: 1986, The maintenance of polygenic variation through a balance between mutation and stabilizing selection, *Genet. Res. (Camb.)* **49**, 157–174.

Barton, N. H.: 1989, The divergence of a polygenic system under stabilizing selection,mutation and drift, *Genet. Res. (Camb.)* **54**, 59–77.

Barton, N. H.: 1990, Pleiotropic models of quantitative variation, *Genetics* **124**, 773–782.

Barton, N. H. and de Vladar, H. P.: 2009, Statistical mechanics and the evolution of polygenic quantitative traits, *Genetics* **181**(3), 997–1011.

Barton, N. H. and Keightley, P. D.: 2002, Understanding quantitative genetic variation, *Nat. Rev. Genet.* **3**(1), 11–21.

Barton, N. H. and Rouhani, S.: 1987, The frequency of shifts between alternative equilibria, *J. Theor. Biol.* **125**(4), 397–418.

Barton, N. H. and Shpak, M.: 2000, The stability of symmetric solutions to polygenic models, *Theor. Pop. Biol.* **57**(3), 249–63.

Barton, N. H. and Turelli, M.: 1987, Adaptive landscapes, genetic-distance and the evolution of quantitative characters, *Genetical Research* **49**(2), 157–173.

Barton, N. H. and Turelli, M.: 1989, Evolutionary quantitative genetics: how little do we know?, *Annu. Rev. Genet.* **23**, 337–70.

Barton, N. H. and Turelli, M.: 1991, Natural and sexual selection on many loci, *Genetics* **127**, 229–55.

Barton, N., Turelli, M. and Houle, D.: 2004, Effects of genetic drift on variance components under a general model of epistasis, *Evolution* .

Beavis, W. D.: 1998, QTL analyses: power, precision and accuracy, *in* A. H. Patterson (ed.), *Molecular dissection of complex traits*, CRC Press, New York, USA, pp. 145–162.

Beerenwinkel, N., Pachter, L., Sturmfels, B., Elena, S. and Lenski, R.: 2007, Analysis of epistatic interactions and fitness landscapes using a new geometric approach, *BMC Evolutionary Biology* **7**(1), 60.

Begin, M. and Roff, D.: 2001, An analysis of G matrix variation in two closely related cricket species, Gryllus firmus and G. pennsylvanicus, *J. Evol. Biol* **14**(1), 1–13.

Begin, M. and Roff, D.: 2003, The constancy of the G matrix through species divergence and the effects of quantitative genetic constraints on phenotypic evolution: A case study in crickets, *Evolution* **57**(5), 1107–1120.

Bergland, A. O., Genissel, A., Nuzhdin, S. V. and Tatar, M.: 2008, Quantitative trait loci affecting phenotypic plasticity and the allometric relationship of ovariole number and thorax length in drosophila melanogaster, *Genetics* **180**(1), 567–82.

Biró, T. and Jakovác, A.: 2005, Power-law tails from multiplicative noise, *Phys. Rev. Lett.* **94**(13), 132302.

Björklund, M.: 2004, Constancy of the G matrix in ecological time, *Evolution* **58**(6), 1157–1164.

Blows, M.: 2007, A tale of two matrices: multivariate approaches in evolutionary biology, *J. Evol. Biol* **20**, 1–8.

Boltzmann, L.: 1872, Further studies on the thermal equilibrium of gas molecules, *Sitzungsberichte der Akademie der Wissenschaften, Wien II* **66**, 275–370.

Boyko, A., Williamson, S., Indap, A. and Degenhardt, J.: 2008, Assessing the evolutionary impact of amino acid mutations in the human genome, *PLoS Genet.* **4**(5), e1000083.

Brodie, E.: 1993, Homogeneity of the genetic variance-covariance matrix for antipredator traits in 2 natural-populations of the garter snake thamnophis-ordinoides, *Evolution* **47**(3), 844–854.

Brown, J. H., Gillooly, J., Allen, A. P., Savage, V. M. and West, G. B.: 2004, Toward a metabolic theory of ecology, *Ecology* **85**(7), 1771–1789.

Brown, J. and West, G. (eds): 2000, *Scaling in Biology*, Oxford Univ. Press, Oxford, UK.

Bulmer, M.: 1971, Effect of selection on genetic variability, *Am. Nat.* **105**(943), 201–211.

Bulmer, M.: 1972, The genetic variability of polygenic characters under optimizing selection, mutation and drift., *Genet. Res.* **19**(1), 17–25.

Bulmer, M.: 1980, *Mathematical Theory of Quantitive Genetics*, Oxford Univ. Press, Oxford, UK.

Bürger, R.: 1991, Moments, cumulants, and polygenic dynamics, *J. Math. Biol.* **30**(2), 199–213.

Bürger, R.: 1993, Predictions of the dynamics of a polygenic character under directional selection, *J. Theor. Biol.* **162**, 487–512.

Bürger, R.: 2000, *The Mathematical Theory of Selection, Recombination, and Mutation*, Wiley Series in Mathematical & Computational Biology., Willey, New York, USA.

Bürger, R. and Gimelfarb, A.: 2004, The effects of intraspecific competition and stabilizing selection on a polygenic trait, *Genetics* **167**(3), 1425–1443.

Bürger, R., Wagner, G. P. and Stettinger, F.: 1989, How much heritable variation can be maintained in finite populations by mutation selection balance, *Evolution* **43**(8), 1748–1766.

Calder, W. A.: 1984, *Size, Function, and Life History*, Harvard Univ. Press, Harvard, USA.

Cano, J. M., Laurila, A., Palo, J. and Merilä, J.: 2004, Population differentiation in G matrix structure due to natural selection in Rana temporaria, *Evolution* **58**(9), 2013–2020.

Carter, A., Hermisson, J. and Hansen, T.: 2005, The role of epistatic gene interactions in the response to selection and the evolution of evolvability, *Theor. Pop. Biol.* **68**, 179–196.

Charlesworth, B., Lande, R. and Slatkin, M.: 1982, a neo-Darwinian commentary on macroevolution, *Evolution* **36**(3), 474–498.

Charnov, E. L.: 1993, *Life history invariants*, Oxford Univ. Press, Oxford, UK.

Cheverud, J.: 2000, Detecting epistasis among quantitative trait loci, *in* J. Wolf, E. D. Brodie and M. Wade (eds), *Epistasis and the evolutionary process*, Oxford Univ. Press, New York, USA, pp. 58–81.

Cheverud, J.: 2006, Genetic architecture of quantitative variation, *in* C. Fox and J. Wolf (eds), *Evoltionary genetics: Concepts and case studies*, Oxford Univ. Press, Oxford, UK, pp. 288–309.

Coulson, T., Benton, T., Lundberg, P., Dall, S., Kendall, B. and Gaillard, J.: 2006, Estimating individual contributions to population growth: evolutionary fitness in ecological time, *Proc. R. Soc. Lond. B* **273**(1586), 547–555.

Coyne, J. A., Barton, N. H. and Turelli, M.: 1997, A critique of Wright's shifting balance theory of evolution, *Evolution* **51**, 643–671.

Crow, J. F. and Kimura, M.: 1970, *An Introduction to Population Genetics Theory*, Harper & Row, New York, USA.

Darwin, C.: 1859, *On the origin of species by means of natural selection*, John Murray, London, UK.

Day, T. and Taylor, P. D.: 1997, von Bertalanffy's growth equation should not be use to model age and size at maturity., *Am. Nat.* **149**(2), 381–393.

de Brito, R. A., Pletscher, L. S. and Cheverud, J. M.: 2005, The evolution of genetic architecture. I. diversification of genetic backgrounds by genetic drift, *Evolution* **59**(11), 2333—2342.

De Groot, S. R. and Mazur, P.: 1984, *Non-Equilibrium Thermodynamics*, Dover Publications Inc., New York.

de Vladar, H. P.: 2006, Density dependence as a size-independent mechanism., *J. Theor. Biol.* **238**(2), 245–256.

de Vladar, H. P. and Gonzalez, J. A.: 2004, Dynamic response of cancer under the influence of immunological activity and therapy, *J. Theor. Biol.* **227**(3), 335–348.

de Vladar, H. P. and Pen, I.: 2007, Determinism, noise, and spurious estimations in a generalised model of population growth, *Physica A* **373**, 477–485.

Diserud, P. and Engen, S.: 2000, A general and dynamic species abundance model, embracing the lognormal and the gamma models., *Am. Nat.* **155**(4), 497–511.

Doebeli, M.: 1996a, A quantitative genetic competition model for sympatric speciation, *J. Evol. Biol* **9**(6), 893–909.

Doebeli, M.: 1996b, Quantitative genetics and population dynamics, *Evolution* **50**(2), 532–546.

Doebeli, M. and de Jong, G.: 1999, Genetic variability in sensitivity to population density affects the dynamics of simple ecological models, *Theor. Pop. Biol.* **55**, 37–52.

Doebeli, M. and Koella, J. C.: 1995, Evolution of simple population-dynamics, *Proc. R. Soc. Lond. B* **260**(1358), 119–125.

Doroszuk, A., Wojewodzic, M. W., Gort, G. and Kammenga, J. E.: 2008, Rapid divergence of genetic variance-covariance matrix within a natural population, *Am. Nat.* **171**(3), 291–304.

Ebenman, B., Johansson, A., Jonsson, T. and Wennergren, U.: 1996, Evolution of stable population dynamics through natural selection, *Proc. R. Soc. Lond. B* **263**(1374), 1145–1151.

Edwards, A. W. F.: 1994, The fundamental theorem of natural selection, *Biol. Rev.* **69**(4), 443–474.

Edwards, A. W. F.: 2002, The fundamental theorem of natural selection, *Theor. Pop. Biol.* **61**, 335–337.

Eigen, M.: 1971, Selforganization of matter and the evolution of biological macromolecules, *Naturwissenschaften* **58**(10), 465–523.

Eigen, M. and Schuster, P.: 1979, *The Hypercycle*, Springer - Verlag, Berlin, Germany.

Ellegren, H.: 2000, Microsatellite mutations in the germline: implications for evolutionary inference, *Trends Genet.* **16**(12), 551–558.

Emmeche, C.: 1998, Defining life as a semiotic phenomenon, *Cybern. Hum. Know.* **5**(1), 3–17.

Ewens, W. J.: 1969, A generalized fundamental theorem of natural selection, *Genetics* **63**, 531–537.

Ewens, W. J.: 1976, Remarks on the evolutionary effect of natural selection, *Genetics* **83**, 601–607.

Ewens, W. J.: 1979, *Mathematical Population Genetics*, Springer - Verlag, Berlin, Germany.

Fa, K.: 2003, Linear Langevin equation with time-dependent drift and multiplicative noise term: exact study, *Chem. Phys.* **287**, 1–5.

Falconer, D. S. and Mackay, T. F. C.: 1996, *Introduction to Quantitative Genetics*, 4th edn, Longmans Green, Essex, UK.

Ferriere, R. and Gatto, M.: 1993, Chaotic population-dynamics can result from natural selection, *Proc. R. Soc. Lond. B* **251**(1330), 33–38.

Fisher, R. A.: 1918, The correlation between relatives on the supposition of mendelian inheritance, *Trans. Roy. Soc. Edinb.* **52**, 399–433.

Fisher, R. A.: 1930, *The genetical theory of natural selection*, 1st edn, Clarendon, Oxford, UK.

Fisher, R. A.: 1941, Average excess and average effect of a gene substitution, *Ann. Eugen.* **11**, 53–63.

Fisher, R. A.: 1958, *The genetical theory of natural selection*, 2nd edn, Dover, New York, USA.

Fisher, R. A.: 1999, *The genetical theory of natural selection (Commented by J. H. Bennet)*, 3rd edn, Oxford Univ. Press, Oxford, UK.

Fisher, R. A. and Tippett, L. H. C.: 1928, Limiting forms of the frequency distribution of the largest or smallest member of a sample, *Proc. Camb. Phil. Soc.* **24**, 180–190.

Fleming, G. R. and Hänggi, P.: 1993, *Activated Barrier Crossing: Applications in Physics, Chemistry and Biology*, World Scientific, Singapore.

Fontana, W. and Buss, L.: 1994, What would be conserved if" the tape were played twice"?, *Proc. Natl. Acad. Sci. USA* **91**(2), 757–761.

Fowler, C.: 1988, Population dynamics as related to rate of increase per generation, *Evolutionary Ecology* .

Frank, S.: 1997, The Price equation, Fisher's fundamental theorem, kin selection, and causal analysis, *Evolution* **51**(6), 1712–1729.

Frank, S. and Slatkin, M.: 1990, The distribution of allelic effects under muta on and selection, *Genet. Res.* **55**, 111–117.

Frank, S. and Slatkin, M.: 1992, Fisher's fundamental theorem of natural selection., *Trends Ecol. Evol.* **7**(3), 92–95.

Gardiner, C.: 2004, *Handbook of Stochastic Methods*, Springer - Verlag, Berlin, Germany.

Gardner, A. and West, S. A.: 2004, Spite among siblings, *Science* **305**(413-414).

Gatto, M.: 1993, The evolutionary optimality of oscillatory and chaotic dynamics in simple population models, *Theor. Popul. Biol.* **43**(3), 310–336.

Gavrilets, S. and de Jong, G.: 1993, Pleiotropic models of polygenic variation, stabilizing selection, and epistasis and epistasis, *Genetics* **134**, 609—625.

Gavrilets, S. and Hastings, A.: 1993, Maintenance of genetic variability under strong stabilizing selection: A two-locus model, *Genetics* .

Genovese, W. and noz, M. A. M.: 1999, Recent results on multiplicative noise, *Phys. Rev. E* **60**(1), 69–78.

Georgii, H.: 2003, Probabilistic aspects of entropy, *in* A. Greven (ed.), *Entropy*, Princeton Univ. Press, Princeton, USA.

Gilpin, M. and Ayala, F.: 1973, Global models of growth and competition., *Proc. Natl. Acad. Sci. USA* **70**(12), 3590–3.

Gilpin, M. E., Case, T. J. and Ayala, F. J.: 1976, θ-selection., *Math. Biosci.* **32**, 131–139.

Goldstein, S. and Lebowitz, J.: 2004, On the (Boltzmann) entropy of non-equilibrium systems, *Physica D* **193**, 53–66.

Gompertz, B.: 1825, On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies, *Phil. Trans. R. Soc. London* **123**, 513–585.

Gonik, M., Berezovskaya, F., Malchow, H. and Medvinsky, A.: 2005, Evolution of rotifer population dynamics in a heterogeneous habitat: Mathematical modeling, *Biofizika* **50**(5), 928–933.

González, J. A., Trujillo, L. and Escalante, A. A.: 2003, Intrinsic chaos and external noise in population dynamics., *Physica A* **324**, 723.

Goodnight, C. J.: 1995, Epistasis and the increase in additive genetic variance – implications for phase-1 of Wrigth's shifting balance process, *Evolution* **49**(3), 502–511.

Góra, P.: 2005, Population explosion suppressed by noise: stationary distributions and how to simulate them, *New J. Phys.* **7**, 36.

Grafen, A.: 2003, Fisher the evolutionary biologist, *J. Roy. Stat. Soc. D-Sta.* **52**(3), 319–329.

Guillaume, F. and Whitlock, M. C.: 2007, Effects of migration on the genetic covariance matrix, *Evolution* **61**(10), 2398–2409.

Gzyl, H.: 1995, The maximum entropy method, *Series on Advances in Mathematics for Applied Sciences*, Vol. 29, World Scientific Publising, Singapore.

Haddrill, P. R., Bachtrog, D. and Andolfatto, P.: 2008, Positive and negative selection on noncoding dna in drosophila simulans, *Mol. Biol. Evol.* **25**(9), 1825—1834.

Halley, J. and Kunin, W.: 1999, Extinction risk and the $1/f$ family of noise models, *Theor. Pop. Biol.* **56**, 215–230.

Hamilton, W.: 1963, The evolution of altruistic behavior., *Am. Nat.* **97**, 354–356.

Hamilton, W.: 1964, The genetical evolution of social behavior, I & II., *J. Theor. Biol.* **7**, 1–52.

Hänggi, P.: 1994, Escape over fluctuating barriers driven by colored noise, *Chem. Phys.* **180**, 157–166.

Hansen, T.: 2006, The evolution of genetic architecture, *Annu. Rev. Ecol. Syst.* **37**, 123–157.

Hanski, I., Turchin, P., Korpimaki, E. and Henttonen, H.: 1993, Population oscillations of boreal rodents: regulation by mustelid predators leads to chaos., *Nature* **364**(6434), 232–235.

Hart, D., Shochat, E. and Agur, Z.: 1998, The growth law of primary breast cancer as inferred from mammography screening trials data., *Br. J. Cancer* **78**(3), 382–7.

Hastings, A. and Harrison, S.: 1994, Metapopulation dyamics and genetics, *Annu. Rev. Ecol. Syst.* **25**, 167—188.

Hastings, W.: 1970, Monte carlo sampling methods using markov chains and their applications, *Biometrika* **57**(1), 97.

Hatcher, M. J.: 2000, Persistence of selfish genetic elements: population structure and conflict, *Trends Ecol. Evol.* **15**(7), 271—277.

Haygood, R., Fedrigo, O., Hansonu, B. and Yokoyama, K.-D.: 2007, Promoter regions of many neural- and nutrition related genes have experienced positive selection during human evolution, *Nat. Genet.* **39**(2), 1140—1144.

Heino, M., Metz, J. A. J. and Kaitala, V.: 1998, The enigma of frequency-dependent selection, *Trends Ecol. Evol.* **13**(9), 367–370.

Henle, K., Sarre, S. and Wiegand, K.: 2004a, The role of density regulation in extinction processes and population viability analysis, *Biodiv. Conserv.* **13**, 9–52.

Henle, K., Sarre, S. and Wiegand, K.: 2004b, The role of density regulation in extinction processes and population viability analysis, *Biodiv. Cons.* **13**, 9–52.

Hershey, A. D.: 1939, Factors limiting bacterial growth: IV equations describin the early periods of increase, *J. Gen. Physiol* **23**, 11–19.

Hoede, C., Denamur, E. and Tenaillon, O.: 2006, Selection acts on dna secondary structures to decrease transcriptional mutagenesis, *PLoS Genet.* **2**(11), e176.

Hoekstra, H., Hoekstra, J., Berrigan, D., Vignieri, S., Hoang, A., Hill, C., Beerli, P. and Kingsolver, J.: 2001, Strength and tempo of directional selection in the wild, *Proc. Natl. Acad. Sci. USA* **98**(16), 9157–9160.

Iwasa, Y.: 1988, Free fitness that always increases in evolution, *J. Theor. Biol.* **135**, 265–281.

Jones, A., Arnold, S. and Bürger, R.: 2003, Stability of the G-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift, *Evolution* **57**(8), 1747–1760.

Jones, A., Arnold, S. and Bürger, R.: 2004, Evolution and stability of the G-matrix on a landscape with a moving optimum, *Evolution* **58**(8), 1639–1654.

Jones, A., Arnold, S. and Bürger, R.: 2007, The mutation matrix and the evolution of evolvability, *Evolution* **61**(4), 727–745.

Kadanoff, L.: 2000, *Statistical Physics, Statics, Dynamics, and Renormalization*, World Scientific, London.

Kalisz, S. and Krishnamurthy, M.: 2007, Variation and constraint in plant evolution and development, *Heredity* **100**(2), 1–7.

Kaniadakis, G. and Lapenta, G.: 2000, Microscopic dynamics underlying anomalous diffusion, *Phys. Rev. E* **62**(3), 3246–3249.

Karlin, S. and Feldman, M. W.: 1970, Convergence to equilibrium of the two locus additive viability model, *J. Appl. Probab.* **7**(2), 262–&.

Karlin, S. and Taylor, H. M.: 1975, *A First Course in Stochastic Processes*, 2nd edn, Academic Press, London, UK.

Kawabe, A., Fujimoto, R. and Charlesworth, D.: 2007, High diversity due to balancing selection in the promoter region of the medea gene in arabidopsis lyrata, *Curr. Biol.* **17**, 1885—1889.

Keightley, P. D.: 1991, Genetic variance and fixation probabilities at quantitative trait loci in mutation-selection balance., *Genet. Res.* **58**, 139–144.

Keightley, P. D. and Hill, W.: 1990, Variation maintained in quantitative traits with mutation selection balance - pleiotropic side-effects in fitbess traits, *Proc. R. Soc. London B* **242**(1304), 95–100.

Keightley, P. D. and Hill, W. G.: 1987, Directional selection and variation in finite populations, *Genetics* **117**, 573–582.

Kenney-Hunt, J. P., Wang, B., Norgard, E. A., Fawcett, G., Falk, D., Pletscher, L. S., Jarvis, J. P., Roseman, C., Wolf, J. and Cheverud, J. M.: 2008, Pleiotropic patterns of quantitative trait loci for 70 murine skeletal traits, *Genetics* **178**(4), 2275–2288.

Kim, Y. and Wiehe, T.: 2008, Simulation of dna sequence evolution under models of recent directional selection, *Brief. Bioinform.* **10**(1), 84–96.

Kimura, M.: 1955, Solution of a process of random genetic drift with a continuous model, *Proc. Natl. Acad. Sci. USA* **41**, 144–150.

Kimura, M.: 1965a, A stochastic model concerning the maintenance of genetic variability in quantitative characters, *Proc. Natl. Acad. Sci. USA* **54**, 731–736.

Kimura, M.: 1965b, A stochastic model concerning the maintenance of genetic variability in quantitative characters., *Proc. Nat. Acad. Sci. USA* **54**, 731–736.

Kimura, M.: 1985, *The neutral theory of molecular Evolution*, Cambridge Univ. Press, Cambridge, UK.

Kingman, J. F. C.: 1978, A simple model for the balance between selection and mutation, *J. Appl. Probab.* **15**(1), 1–12.

Kingsolver, J., Hoekstra, H., Hoekstra, J., Berrigan, D., Vignieri, S., Hill, C., Hoang, A., Gibert, P. and Beerli, P.: 2001, The strength of phenotypic selection in natural populations, *Am. Nat.* **157**(3), 245–261.

Kirkpatrick, M., Johnson, T. and Barton, N. H.: 2002, General models of multilocus evolution, *Genetics* **161**, 1727–1750.

Klein, G. and Prigogine, I.: 1953, Sur la mecanique statistique des phenomenes irreversibles I, *Physica* **19**(1), 74–88.

Kondrashov, A. S. and Turelli, M.: 1992, Deleterious mutations, apparent stabilizing selection and the maintenance of quantitative variation, *Genetics* **132**(2), 603–618.

Kopp, M. and Hermisson, J.: 2007, Adaptation of a quantitative trait to a moving optimum, *Genetics* **176**(1), 715.

Kotiaho, J.: 2007, The stability of genetic variance-covariance matrix in the presence of selection., *J. Evol. Biol* **20**, 28–29.

Kozlowski, J. and Konarzewski, M.: 2004, Is West, Brown and Enquist's model for allometric scaling mathematically correct and biologically relevant?, *Funct. Ecol.* **18**, 283–289.

Kozusko, F. and Bajzer, Z.: 2003, Combining Gompertzian growth and cell population dynamics, *Math. Biosci.* **185**(2), 153–167.

Lande, R.: 1975, The maintenance of genetic variability by mutation in a polygenic character with linked loci, *Genet. Res.* **26**, 221–235.

Lande, R.: 1976, Natural selection and random genetic drift in phenotypic evolution, *Evolution* **30**(2), 314–334.

Lande, R.: 1979, Quantitative genetic-analysis of multivariate evolution, applied to brain - body size allometry, *Evolution* **33**(1), 402–416.

Lande, R.: 1980, The genetic covariance between characters maintained by pleiotropic mutations, *Genetics* **94**, 203–215.

Lande, R.: 1981, The minimum number of genes contributing to quantitative variation between and within populations, *Genetics* **99**(3-4), 541–553.

Lande, R., Engen, S. and Saether, B. E.: 2003, *Stochastic Population Dynamics in Ecology and Conservation*, Oxford University Press, New York, USA.

Laurila, A., Karttunen, S. and Merilä, J.: 2002, Adaptive phenotypic plasticity and genetics of larval life histories in two rana temporaria populations, *Evolution* **56**(3), 617–627.

Le Bellac, M., Mortessagne, F. and Batrouni, G. G.: 2004, *Equilibrium and non-equilibrium statistical thermodynamics*, Cambridge Univ. Press, Cambridge, UK.

Leff, H. S. and Rex, A. F.: 2003, *Maxwell's Demon II*, Institute of Physics, Bristol, UK.

Lenormand, T. and Otto, S.: 2000, The evolution of recombination in a heterogeneous environment, *Genetics* **156**(1), 423–438.

León, J. A. and Charlesworth, B.: 1976, Ecological versions of Fisher's fundamental theorem of natural selection, *Adv. Appl. Probab.* **8**(4), 639–641.

León, J. A. and Charlesworth, B.: 1978, Ecological versions of Fisher's fundamental theorem of natural selection, *Ecology* **59**(3), 457–464.

Li, W. H.: 1997, *Molecular Evolution*, Sinauer Associates, Inc, Suntherland, USA.

Lion, S. and van Baalen, M.: 2008, Self-structuring in spatial evolutionary ecology, *Ecol. Lett.* **11**, 277—295.

Loewe, L., Charlesworth, B., Bartolome, C. and Noel, V.: 2006, Estimating selection on nonsynonymous mutations, *Genetics* **172**(2), 1079–1092.

Lwoff, A.: 1965, *Biological Order*, MIT Press, Boston, USA.

Lynch, M. and Walsh, B.: 1998, *Genetics and Analysis of Quantitative Traits*, Sinauer Associates, Sunderland, USA.

Ma, C.-X., Yu, Q., Berg, A., Drost, D., Novaes, E., Fu, G., Yap, J. S., Tan, A., Kirst, M., Cui, Y. and Wu, R.: 2008, A statistical model for testing the pleiotropic control of phenotypic plasticity for a count trait, *Genetics* **179**(1), 627–636.

MacArthur, R. H.: 1962, Some generalized theorems of natural selection., *Proc. Nat. Acad. Sci. USA* **46**, 1893–1897.

MacArthur, R. and Wilson, E. O.: 1967, *The theory of island biogeography*, Princeton Univ. Press, Princeton, USA.

Mackay, T. F. C.: 2001, Quantitative trait loci in drosophila, *Nat. Rev. Genet.* **2**, 11–20.

Mackay, T. F. C.: 2004, Complementing complexity, *Nat. Genet.* **36**(11), 1145–1147.

Malthus, T. R.: 1798, *Principle of Population*, Vol. 1, 6 (1826) edn, John Murray, London, UK.

Mao, X., Marion, G. and Renshaw, E.: 2002, Environmental brownian noise suppresses explosions in population dynamics, *Stoch. Proc. Appl.* **97**, 95–110.

May, R.: 1976, Simple mathematical model with very complicated dynamics., *Nature* **261**, 459–467.

Maynard-Smith, J.: 1983, The genetics of stasis and punctuation, *Annu. Rev. Genet.* **17**, 11–25.

Maynard-Smith, J.: 1999, *Evolutionary genetics*, Oxford University Press, Oxford, UK.

Maynard-Smith, J. and Szathmáry, E.: 1997, *The major transitions in evolution*, Oxford University Press, Oxford, UK.

Maynard-Smith, J. and Szathmáry, E.: 2000, *The Origins of Life: From the Birth of Life to the Origin of Language*, Oxford Univ. Press, Oxford, UK.

Mayr, E.: 1982, *The Growth of Biological Thought. Diversity, Evolution, and Inheritance*, The Belknap Press.

Meszéna, G., Gyllenberg, M., Jacobs, F. and Metz, J. A. J.: 1992, Link between population dynamics and dynamics of darwinian evolution, *Phys. Rev. Lett.* **95**, 078105.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A. and Teller, E.: 1953, Equation of state calculations by fast computing machines, *J. Chem. Phys.* **21**(6), 1087.

Metz, J. A. J., Mylius, S. and Diekmann, O.: 2008, When does evolution optimize?, *Evol. Ecol. Res.* **10**(5), 629–654.

Metz, J. A. J., Nisbet, R. and Geritz, S.: 1992, How should we define'fitness' for general ecological scenarios?, *Trends Ecol. Evol.* **7**(6), 198–202.

Molski, M. and Konarski, J.: 2003, Coherent states of Gompertzian growth, *Phys. Rev. E* **68**(2), 021916.

Morita, A. and Makino, J.: 1986, Simple analytical solution for multiplicative nonlinear stochastic differential equations by a perturbation technique., *Phys. Rev. A* **34**(2), 1595–1598.

Murray, B. G.: 1997, Population dynamics of evolutionary change: Demographic parameters as indicators of fitness, *Theor. Popul. Biol.* **51**(3), 180–184.

Mylius, S. and Diekmann, O.: 1995, On evolutionarily stable life histories, optimization and the need to be specific about density dependence, *Oikos* **74**, 218–224.

Nagylaki, T.: 1979, Dynamics of density-and frequency-dependent selection, *Proc. Natl. Acad. Sci. USA* .

Nagylaki, T.: 1989, The maintenance of genetic variability in two-locus models of stabilizing selection, *Genetics* **122**(1), 235–248.

Nagylaki, T.: 1993, The evolution of multilocus systems under weak selection, *Genetics* **134**(2), 627.

Nicolis, G. and Prigogine, I.: 1977, *Self organization in non-equilibrium systems: From dissipative structures to order through fluctuations*, John Wiley & Sons, New York, USA.

Norton, L. and Simon, R.: 1977, Tumor size, sensitivity to therapy, and design of treatment schedules., *Cancer Treat. Rep.* **61**(7), 1307–1317.

Norton, L., Simon, R., Brereton, H. D. and Bogden, A. E.: 1976, Predicting the course of Gompertzian growth., *Nature* **264**(5586), 542–545.

Nosil, P., Crespi, B., Sandoval, C. and Kirkpatrick, M.: 2006, Migration and the genetic covariance between habitat preference and performance, *Am. Nat.* **167**(3), E66–E78.

Okasha, S.: 2006, *Evolution and the levels of selection*, Oxford Univ. Press, Oxford, UK.

Oksendal, B.: 2002, *Stochastic Differential Equations*, Springer-Verlag, Berlin.

Ollivier, L. and Janss, L. L.: 1993, A note on the estimation of the effective number of additive and dominant loci contributing to quantitative variation, *Genetics* **135**(3), 907–9.

Onsager, L.: 1931, Reciprocal relations in irreversible processes. I., *Phys. Rev.* **37**, 405–426.

Orr, H. A.: 2000, Adaptation and the cost of complexity, *Evolution* **54**(1), 13–20.

Orr, H. A.: 2003, The distribution of fitness effects among beneficial mutations, *Genetics* **163**(4), 1519–1526.

Orr, H. A.: 2005, Theories of adaptation: what they do and don't say, *Genetica* **123**, 3–13.

Orzack, S. H. and Tuljapurkar, S. D.: 1989, Population dynamics in variable environments.7. the demographics and evolution of itroparity, *Am. Nat.* **133**(6), 901–923.

Otto, S. and Jones, C.: 2000, Detecting the undetected: Estimating the total number of loci underlying a quantitative trait, *Genetics* **156**(4), 2093–2107.

Ovaskainen, O., Cano, J. M. and Merilä, J.: 2008, A Bayesian framework for comparative quantitative genetics, *Proc. R. Soc. London B* **275**(1635), 669–678.

Page, K. M. and Nowak, M. A.: 2002, Unifying evolutionary dynamics, *J. Theor. Biol.* **219**, 93—98.

Palo, J., O'Hara, R., Laugen, A., Laurila, A., Primmer, C. and Merilä, J.: 2003, Latitudinal divergence of common frog (Rana temporaria) life history traits by natural selection: evidence from a comparison of molecular and quantitative genetic data, *Mol. Ecol.* **12**(7), 1963–1978.

Paulsen, S. M.: 1996, Quantitative genetics of the wing color pattern in the buckeye butterfly (Precis coenia and Precis evarete): evidence against the constancy of G, *Evolution* **50**, 1585—1597.

Pearl, R.: 1927, The growth of populations., *Q. Rev. Biol.* **2**, 532–548.

Pearson, K.: 1896, Mathematical contributions to the theory of evolution. III.regression, heredity and panmixia., *Philos. Trans. R. Soc. London A* **187**, 253–318.

Pelletier, F., Clutton-Brock, T., Pemberton, J., Tuljapurkar, S. and Coulson, T.: 2007, The evolutionary demography of ecological change: Linking trait variation and population growth, *Science* **315**(5818), 1571–1574.

Peters, R. H. (ed.): 1983, *The Ecological Implications of Body Size*, Cambridge University Press, Cambridge, UK.

Phillips, P., Whitlock, M. C. and Fowler, K.: 2001, Inbreeding changes the shape of the genetic covariance matrix in drosophila melanogaster, *Genetics* **158**(3), 1137–1145.

Pianka, E. R.: 1970, On r and k selection, *Am. Nat.* **104**, 592–597.

Pigliucci, M.: 2006, Genetic variance-covariance matrices: A critique of the evolutionary quantitative genetics research program, *Biol. Philos.* **21**(1), 1–23.

Press, W., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P.: 2007, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, Cambridge Univ. Press, Cambridge, UK.

Price, G.: 1970, Selection and covariance, *Nature* **227**, 520–521.

Prigogine, I.: 1949, Le domaine de validité de la thermodynamique des phénomènes irréversibles, *Physica* **15**(1-2), 272–284.

Provine, W. B.: 1986, *Sewall Wright and Evolutionary Biology*, Univ. Chicago Press, Chicago, USA.

Prugel-Bennett, A.: 1997, Modelling evolving populations, *J. Theor. Biol.* **185**(1), 81–95.

Prugel-Bennett, A. and Shapiro, J.: 1994, Analysis of genetic algorithms using statistical mechanics, *Phys. Rev. Lett.* **72**(9), 1305–1309.

Prugel-Bennett, A. and Shapiro, J.: 1997, The dynamics of a genetic algorithm for simple random ising systems, *Physica D* **104**, 75–114.

Rand, D. M.: 2001, The units of selection on mitochondrial DNA, *Annu. Rev. Ecol. Syst.* **32**, 415—-548.

Rattray, M. and Shapiro, J.: 2001, Cumulant dynamics of a population under multiplicative selection, mutation, and drift, *Theor. Pop. Biol.* **60**, 17–32.

Reeve, J.: 2000, Predicting long-term response to selection, *Genet. Res.* **75**(1), 83–94.

Reif, F.: 1965, *Fundamentals of Statistical and Thermal Physics*, Mcgraw-Hill, New York, USA.

Renaud, S., Auffray, J.-C. and Michaux, J.: 2006, Conserved phenotypic variation patterns, evolution along lines of least resistance, and departure due to selection in fossil rodents, *Evolution* **60**(8), 1701–1717.

Renshaw, E.: 1991, *Modelling biological populations in space and time*, Cambridge Univ. Press, New York, USA.

Renyi, A.: 1961, On measures of entropy and information, *Proc. Fourth. Berkeley Symp. Math. Stat. Prob.*, Vol. 1, University of California Press, Berkeley.

Roff, D.: 1986, Predicting body size with life history models, *BioScience* **36**, 316–23.

Roff, D.: 2000, The evolution of the G matrix: selection or drift?, *Heredity* **84**, 135–142.

Roff, D.: 2007, A centenial celebration for quantitative genetics, *Evolution* **61**(5), 1017–1032.

Roff, D., Emerson, K. and Goodnight, C.: 2006, Epistasis and dominance: evidence for differential effects in life-history versus morphological traits, *Evolution* .

Rogers, A.: 2003, Phase transitions in sexual populations subject to stabilizing selection, *Phys. Rev. Lett.* **90**(15), 158103.

Rogers, A. and Prugel-Bennett, A.: 2000, Evolving populations with overlapping generations, *Theor. Pop. Biol.* **57**(2), 121–129.

Rouhani, S. and Barton, N. H.: 1987, The probability of peak shifts in a founder population., *J. Theor. Biol.* **126**(1), 51–62.

Rouzine, I. M., Brunet, É. and Wilke, C.: 2007, The traveling-wave approach to asexual evolution: Muller's ratchet and speed of adaptation, *Theor. Pop. Biol.* **73**, 24–46.

Rouzine, I. M., Wakeley, J. and Coffin, J. M.: 2003, The solitary wave of asexual evolution, *Proc. Natl. Acad. Sci. USA* **100**(2), 587–92.

Roze, D. and Barton, N. H.: 2006, The Hill-Robertson effect and the evolution of recombination, *Genetics* **173**(3), 1793–1811.

Rueffler, C., Egas, M. and Metz, J. A. J.: 2006, Evolutionary predictions should be based on individual-level traits, *Am. Nat.* **168**(5), 148–162.

Saakian, D. B. and Hu, C.-K.: 2006, Exact solution of the Eigen model with general fitness functions and degradation rates, *Proc. Natl. Acad. Sci. USA* **103**(13), 4935–4939.

Saakian, D. B., Kirakosyan, Z. and Hu, C.-K.: 2008, Diploid biological evolution models with general smooth fitness landscapes and recombination, *Phys. Rev. E* **77**(6), 10.

Saakian, D. B., Muñoz, E., Hu, C.-K. and Deem, M. W.: 2006, Quasispecies theory for multiple-peak fitness landscapes, *Phys. Rev. E* **73**(4), 10.

Saether, B. E., Engen, S., Lande, R., Arcese, P. and Smith, J. N.: 2000, Estimating the time to extinction in an island population of song sparrows., *Proc. R. Soc. Lond. B* **267**(1443), 621–626.

Saether, B. and Engen, S.: 2002, Pattern of variation in avian population growth rates., *Phil. Trans. R. Soc. Lond. B* **357**(1425), 185–195.

Saether, Engen, S., Filli, F., Aanes, R., Schröder, W. and Andersen, R.: 2002, Stochastic population dynamics of an introduced swiss population of the Ibex, *Ecology* **83**(12), 3457–3465.

Sagan, C.: 1973, Life, *The encyclopaedia Britannica*, 15th edn, Vol. 10, William Benton, London, UK.

Scheuring, I. and Szathmary, E.: 2001, Survival of replicators with parabolic growth tendency and exponential decay, *J. Theor. Biol.* **212**, 99–106.

Schliekelman, P. and Ellner, S. P.: 2001, Egg size evolution and energetic constraints on population dynamics, *Theor. Popul. Biol.* **60**(2), 73–92.

Schneider, K.: 2006, A multilocus-multiallele analysis of frequency-dependent selection induced by intraspecific competition, *J. Math. Biol.* **52**(4), 483–523.

Sella, G. and Hirsh, A.: 2005, The application of statistical physics to evolutionary biology, *Proc. Natl. Acad. Sci. USA* **102**(27), 9541–9546.

Shaw, R. G., Platenkamp, G. A., Shaw, F. H. and Podolsky, R. H.: 1995, Quantitative genetics of response to competitors in nemophila menziesii: a field experiment, *Genetics* **139**(1), 397–406.

Shertzer, K. W. and Ellner, S. P.: 2002, Energy storage and the evolution of population dynamics, *J. Theor. Biol.* **215**(2), 183–200.

Shoemaker, J., Painter, I. and Weir, B.: 1999, Bayesian statistics in genetics: a guide for the uninitiated, *Trends Genet.* **15**(9), 354–358.

Shpak, M. and Kondrashov, A. S.: 1999, Applicability of the hypergeometric phenotypic model to haploid and diploid populations, *Evolution* **53**(2), 600–604.

Sibly, R., Baker, D., Denham, M., Hone, J. and Pagel, M.: 2005, On the regulation of populations of mammals, birds, fish, and insects., *Science* **309**, 607–610.

Siefert, M., Kittel, A., Friedrich, R. and Peinke, J.: 2003, On a quantitative method to analyze dynamical and measurement noise, *Europhys. Lett.* **61**(4), 466–472.

Slatkin, M.: 1980, Ecological character displacement, *Ecology* **61**(1), 163–177.

Slatkin, M. and Frank, S.: 1990, The quantitative genetic consequences of pleiotropy under stabilizing and directional selection, *Genetics* **125**, 207–213.

Stearns, S.: 2004, *The evolution of life histories*, Oxford Univ. Press, Oxford, UK.

Steppan, S., Phillips, P. and Houle, D.: 2002, Comparative quantitative genetics: evolution of the G matrix, *Trends Ecol. Evol.* **17**(7), 320–327.

Szathmary, E.: 1991, Simple growth laws and selection consequences., *Trends Ecol. Evol.* **6**(11), 366–370.

Szathmary, E.: 1993, A note on the reduction of the dynamics of multilocus diploid genetic systems with multiplicative fitness., *J. Theor. Biol.* **164**, 351–358.

Szathmary, E.: 1995, A classification of replicators and lambda-calculus models of biological organization, *Proc. Natl. Acad. Sci. USA* **260**(1359), 279–286.

Szathmáry, E. and Demeter, L.: 1987a, Group selection of early replicators and the origin of life., *J. Theor. Biol.* **128**(4), 463–486.

Szathmary, E. and Demeter, L.: 1987b, Group selection of early replicators and the origin of life, *J. Theor. Biol.* **128**(4), 463–486.

Szathmary, E. and Gladkih, I.: 1989, Sub-exponential growth and coexistence of non-enzymatically replicating templates., *J. Theor. Biol.* **138**(1), 55.

Szent-György, A.: 1972, *The Living State*, Academic Press, New York, USA.

Tanaka, Y.: 2005, Constrained evolution of a quantitative character by pleiotropic mutation, *Theor. Pop. Biol.* **68**(4), 243–251.

Travis, J. and Greenwood, J. J. D.: 1990, The interplay of population dynamics and the evolutionary process, *Phil. Trans. R. Soc. Lond. B* **330**(1257), 253–259.

Tsallis, C.: 1988, Possible generalization of Boltzmann-Gibbs statistics, *J. Stat. Phys.* **52**, 479–487.

Tuljapurkar, S.: 1990, *Population dynamics in variable envrironments*, Springer, Berlin, Germany.

Tuljapurkar, S. D.: 1982, Population dynamics in variable environments.3. evolutionary dynamics of r-selection, *Theor. Popul. Biol.* **21**(1), 141–165.

Tuljapurkar, S. D. and Orzack, S. H.: 1980, Population dynamics in variable environments I. long-run growth rates and extinction, *Theor. Pop. Biol.* **18**, 314—342.

Turelli, M.: 1984, Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle, *Theor. Pop. Biol.* **25**(2), 138–193.

Turelli, M.: 1988, Phenotypic evolution, constant covariances, and the maintenance of additive variance, *Evolution* **42**(6), 1342–1347.

Turelli, M. and Barton, N. H.: 1990, Dynamics of polygenic characters under selection, *Theor. Pop. Biol.* **38**, 1–57.

Turelli, M. and Barton, N. H.: 1994, Genetic and statistical analyses of strong selection on polygenic traits: What, me normal?, *Genetics* **138**(3), 913–941.

Turelli, M. and Barton, N. H.: 2004, Polygenic variation maintained by balancing selection pleiotropy, sex-dependent allelic effects and GxE interactions, *Genetics* **166**, 1053–1079.

van Kampen, N. G.: 1957, Derivation of the phenomenological equations from the master equation, *Physica* **101**, 707–719.

Verhulst, P.: 1838, Notice sur la loi due la population suit dans son accrossement., *Corresp. Math. Phys.* **10**, 113–121.

Vlad, M., Szedlacsek, S., Pourmand, N., Cavalli-Sforza, L. L., Oefner, P. and Ross, J.: 2005, Fisher's theorems for multivariable, time- and space-dependent systems, with applications in population genetics and chemical kinetics, *Proc. Natl. Acad. Sci. USA* **102**(28), 948–9853.

von Bertalanffy, L.: 1957, Quantitative laws in metabolism and growth., *Q. Rev. Biol.* **32**, 217–231.

von Bertalanffy, L.: 1966, On the von Bertalanffy growth curve., *Growth* **30**(1), 123–124.

von Kiedrowski, G.: 1986, A self-replicating hexadeoxinucleotide, *Angw. Chem. Intl. Ed.* **25**(10), 932–935.

Wagner, G. P.: 1984, On the eigenvalue distribution of genetic and phenotypic dispersion matrices: evidence for a nonrandom organization of quantitative character variation, *J. Math. Biol.* **21**(1), 77–95.

Wagner, G. P., Kenney-Hunt, J. P., Pavlicev, M., Peck, J. R., Waxman, D. and Cheverud, J. M.: 2008, Pleiotropic scaling of gene effects and the 'cost of complexity', *Nature* **452**, 470—473.

Wang, J., Caballero, A., Keightley, P. and Hill, W.: 1998, Bottleneck effect on genetic variance a theoretical investigation of the role of dominance, *Genetics* .

Wehrl, A.: 1978, General properties of entropy., *Rev. Mod. Phys.* **50**(2), 221–260.

Werren, J. H., Nur, U. and Wu, C.-I.: 1988, Selfish genetic element, *Trends Ecol. Evol.* **3**(11), 297—302.

West, G., Brown, J. and Enquist, B.: 1997, A general model for the origin of allometric scaling laws in biology, *Science* **276**, 122–126.

West, G., Brown, J. and Enquist, B.: 2001, A general model for ontogenetic growth., *Nature* **413**(6856), 628–631.

Wheldon, T.: 1988, *Mathematical models in cancer research*, Adam Hilger, Bristol, UK.

White, A., Greenman, J., Benton, T. and Boots, M.: 2006, Evolutionary behaviour in ecological systems with trade-offs and non-equilibrium population dynamics, *Evol. Ecol. Res.* **8**(3), 387–398.

Wichmann, M., Johstb, K., Schwagerc, M., Blasiusd, B. and Jeltschc, F.: 2005, Extinction risk, coloured noise and the scaling of variance, *Theor. Pop. Biol.* **68**, 29–40.

Widen, B., Andersson, S., Rao, G. and Widen, M.: 2002, Population divergence of genetic (co)variance matrices in a subdivided plant species, Brassica cretica, *J. Evol. Biol* **15**(6), 961–970.

Wilkinson, G. S., Fowler, K. and Partridge, L.: 1990, Resistance of genetic correlation structure to directional selection in drosophila melanogaster, *Evolution* **44**, 1990—2003.

Willensdorfer, M. and Bürger, R.: 2003, The two-locus model of gaussian stabilizing selection, *Theor. Pop. Biol.* **64**, 101–117.

Wright, S.: 1932, The roles of mutation, inbreeding, crossbreeding and selection in evolution, *Proc. Sixth Intl. Cong. Genet.* **1**, 356–366.

Wright, S.: 1935, Evolution in populations in approximate equilibrium, *J. Genet.* **30**, 257–266.

Wright, S.: 1938, The distribution of gene frequencies under irreversible mutation, *Proc. Natl. Acad. Sci. USA* **24**(7), 253–259.

Wright, S.: 1967, 'Surfaces' of selective value, *Proc. Natl. Acad. Sci. USA* **58**(1), 165–172.

Wright, S.: 1968, *Evolution and the Genetics of Populations Volume 1: Genetic and Biometric Foundations.*, University of Chicago Press, Chicago.

Wright, S.: 1988, Surfaces of selective value revisited, *Am. Nat.* **131**(1), 115.

Xu, S.: 2003, Theoretical basis of the Beavis effect, *Genetics* **165**, 2259—-2268.

Yang, Z. H. and Swanson, W. J.: 2002, Codon-substitution models to detect adaptive evolution that account for heterogeneous selective pressures among site classes, *Mol. Biol. Evol.* **19**(1), 49—57.

Yang, Z. H., Wong, W. S. and Nielsen, R.: 2005, Bayes empirical Bayes inference of amino acid sites under positive selection, *Mol. Biol. Evol.* **22**(4), 1107—1118.

Yoshimura, J. and Jansen, V. A. A.: 1996, Evolution and population dynamics in stochastic environments, *Res. Ppopul. Ecol.* **38**(2), 165–182.

Yukilevich, R., Lachance, J., Aoki, F., True, J. and Travisano, M.: 2008, Long-term adaptation of epistatic genetic networks, *Evolution* **62**(9), 2215–2235.

Zhang, X. S. and Hill, W.: 2005, Genetic variability under mutation selection balance, *Trends Ecol. Evol.* **20**(9), 468–470.

Appendix

Appendix A

Maximization of entropy or of mean fitness?

A.1 MAXIMUM ENTROPY

Here, it is shown that the stationary distribution, ψ_0, maximizes the entropy, S_H, subject to constraints on the expected values of a set of observables (see Le Bellac et al., 2004, p. 64), for a treatment in a physical context). We write the potential function in the form $\log\left(\bar{W}\right) + U = \vec{\alpha}.\vec{A}$, where the vector $\vec{A}$ is a function of the allele frequencies $\vec{p}$, and $\vec{\alpha}$ is a vector of coefficients. Crucially, the observables $\vec{A}$ may be a nonlinear function of the microscopic variables, $\vec{p}$. The simplest choice would be to set $\alpha_1 = \mu$, $\alpha_2 = s$, as measures of the rates of mutation and selection. Then, $A_1 = \vec{n}2\sum_{k=1}\left(\theta_k \log\left[p_k\right] + \left(1-\theta_k\right)\log\left[q_k\right]\right)$, where $\theta_k = \mu_{P,k}/\left(\mu_{Q,k}+\mu_{P,k}\right)$, and $A_2 = \log\left(\bar{W}\right)/s$ determines the form of selection. We might further separate $\log\left(\bar{W}\right)$ into separate sources of selection: for example, with stabilizing selection of strength s towards an optimum at z_{opt}, $\log\left(\bar{W}\right) = -s\frac{v}{2} - \frac{s\left(\bar{z}-z_{\text{opt}}\right)^2}{2} = -s\frac{v}{2} - s\frac{\bar{z}^2}{2} + s\bar{z}z_{\text{opt}} -$ constant. Thus, we can set

$$\vec{A} = \left\{\vec{n}2\sum_{k=1}\left(\theta_k \log\left[p_k\right] + \left(1-\theta_k\right)\log\left[q_k\right]\right), -\frac{v}{2}, -\frac{\bar{z}^2}{2}, \bar{z}\right\};$$

the coefficients $\vec{\alpha} = \{\mu, s, s', sz_{\text{opt}}\}$ then represent mutation, selection to reduce variance in the trait, v, stabilizing selection to reduce deviations in the mean, $\bar{z}^2$, and directional selection on the trait mean, $\bar{z}$. We might also add observables that do not affect fitness, but are nevertheless of interest, by setting their α_k to zero. Likewise, setting $s = s' = 0$ but $sz_{\text{opt}} \equiv \beta$ we recover directional selection.

Generalizing the definition of S_H given below Eq. 3.6 we write:

$S_H[\psi] \equiv \int \psi \log\left(\frac{\phi}{\psi}\right) d\vec{p}$

where ϕ is a measure which we take here to be $\phi = \prod_{k=1}^{n}\left(p_k q_k\right)^{-1}$. To find the distribution P_{ME} that maximizes S_H subject to constraints on the expectations, $\left\langle \vec{A} \right\rangle$, we use the method of Lagrange multipliers, setting these multipliers to be proportional to $2N\vec{\alpha}$. We also require the constraint that $\int \psi_{\text{ME}}\, d\vec{p} = 1$, with associated multiplier denoted by

$2N\gamma$:

$$\begin{aligned} 0 =& \delta \mathbf{S}_H + 2N\gamma\delta\left(\int \psi\, d\vec{p}\right) + 2N\vec{\alpha}.\delta\left\langle\right\rangle \\ =& \int (\log(\psi) - 1)\delta\psi d\vec{p} + 2N\gamma \int \delta\psi\, d\vec{p} + 2N \int \vec{\alpha}.\delta\psi d\vec{p} \\ =& \int \left(\log\left(\frac{\phi}{\psi}\right) + (2N\gamma - 1) + 2N\vec{\alpha}.\right)\delta\psi d\vec{p} \end{aligned}$$

Rewriting the normalization as $\mathbb{Z} = \text{Exp}(1 - 2N\gamma)$, we find that the distribution ψ_{ME} that maximizes S_H, for given values of $\left\langle \vec{A} \right\rangle$, is: $\psi = \frac{1}{\mathbb{Z}}\phi e^{2N\vec{\alpha}.}$
where $\mathbb{Z} = \int \phi e^{2N\vec{\alpha}.} d\vec{p}$

The coefficients $\vec{\alpha}$ determine the values of the expectations through the constraint:
$\left\langle \vec{A} \right\rangle = \frac{1}{\mathbb{Z}} \int \phi\ e^{2N\vec{\alpha}.} d\vec{p}$

They can also be found by differentiating the normalization: $\left\langle \vec{A} \right\rangle = \frac{1}{2N} \frac{\partial \log(\mathbb{Z})}{\partial \vec{\alpha}}$, (c.f. Le Bellac et al., 2004, Eq. 2.66).

A.2 MAXIMUM FITNESS

The second part of this thesis seems to be relying on the fact that a mysterious function, the entropy, is maximized at selection-mutation-drift equilibrium. Despite that Iwasa (1988) exemplified that this is so, and as also proven in this thesis, there can be some skepticism in that entropy and its maximization is not a property which is actually realized in nature, but rather a mere mathematical convenience. Perhaps it is, but no less than the mathematical conceptualization, axiomatization, and parametrization of fitness functions, for example. We have seen that each macroscopic (extensive) quantity that we constrain, brings along a corresponding (non-extensive) macroscopic quantity, emerging mathematically form the Lagrange multipliers in the maximization of S. In this sense we have that there is a couple of variables, one intensive and one extensive, that determine the effects

of a given effect which define the evolutionary process, for example under directional selection we have seen that

Process		Extensive	Intensive
Selection over a trait	$\rightsquigarrow$	$\langle \bar{z} \rangle$	β
Dominance	$\rightsquigarrow$	$\langle \nu_z \rangle$	σ
Mutation	$\rightsquigarrow$	$\langle U \rangle$	μ
Genetic Drift	$\rightsquigarrow$	S	$2N$

Between lines, the variables corresponding to genetic drift are actually entropy and population size. But why the effects of drift have to be maximized, why should it have any phenomenological priority over fitness, for example, given that selection is the cornerstone of the evolutionary process?

As a matter of fact, entropy bears no priority, except because it measures the stochasticity allowed by the superposition of evolutionary effects. But this is far from being so fundamental as to ascribe to it all evolutionary causation in quantitative genetics. Nevertheless, knowing that entropy is a necessary variable to include the effects of drift, we still need to take it into consideration. Hence, maximizing the expectancy of log mean fitness over the possible allele frequencies:

$$\langle \log(\bar{W}) \rangle = \int_{(0,1)^n} \log(\bar{W}) \psi d^n p)$$

subject to the constraints

$$\begin{aligned} S &= \int_{(0,1)^n} \log(\psi/\phi) \psi d^n p \\ \langle A_j \rangle &= \int_{(0,1)^n} A_j \psi d^n p \quad j = 1, \dots K \end{aligned}$$

on any necessary set of macroscopics A_j. Variation with respect to ψ leads to the equation

$$0 = \int_{(0,1)^n} \left(\underbrace{\log(\bar{W})}_{\text{Maximization}} - \underbrace{\lambda}_{\text{Normalization}} - \underbrace{\alpha_0 \log(\psi/\phi)}_{\text{Constraint on S}} - \underbrace{\sum_j^K \alpha_j A_j}_{\text{Other constraints}} \right) \delta\psi d^n p \ ,$$

that implies the distribution:

$$\psi = \phi \exp\left[\frac{1}{\alpha_0} \log(\bar{W}) - \frac{\lambda}{\alpha_0} - \frac{\alpha_j}{\alpha_0} A_j \right] \ .$$

Applying the normalization condition and rearranging we get

$$\psi = \frac{\phi}{\mathbb{Z}} \bar{W}^{1/\alpha_0} \exp\left[-\frac{\alpha_j}{\alpha_0} A_j \right] \ .$$

Naturally, we need to solve the Lagrange multipliers. But to make it short, we can identify already that $\alpha_0 = 1/2N$, and the α_j depend on which macroscopics A_j were defined.

To summarize, I briefly showed that the principle of maximal entropy is neither artificial nor arbitrary, and furthermore it is compatible with the idea of fitness maximization, that is at the core of the theory of evolutionary biology.

Nevertheless, the function we have employed as entropy is to some extent arbitrary. We were able to properly define it because we know from the mechanistic theory (Wright-Fisher process, and its diffusion equation) how the distribution looks like. A counter-example, is Prugel-Bennett (1997); Rogers and Prugel-Bennett (2000) approach where the definition of entropy, although functionally similar, differs from the genuine measure by the prior ϕ, hence leading to a wrong microscopic distribution. Similarly, we could choose virtually any Lyapunov function and constrain the macroscopics of interest. As a result we will always have a quantitative description that, by construction, matches our expectations, and couples them with the microscopic variables. This distribution however, as in Rogers and Prugel-Bennett (2000) would lead to an incorrect microscopic distribution, and those variables that were not constrained would necessarily be incorrect estima-

tors, from which it would be possible to falsify the arbitrary Lyapunov measures as descriptors of genetic drift.

Another approach to the maximization idea. The procedure stated above is only one way to show the equivalence between maximizing entropy and log-mean fitness, or as a matter of fact, any other macroscopic constrained to couple the macro and microscopic variables. The entropy is averaging the evolutionary potential (in the sense of chapter 5), which requires k macroscopics. Thus any choice of k variables between the $k+1$ extensive variables (the constrained macroscopics plus the entropy) would lead to the same distribution, because maximization reduces the $k+1$ degrees of freedom to k, and the choice is related by a change of variables.

Assume that entropy S is maximized. In general, calling the intensive variables $\alpha_i,\ i = 1, \ldots, k$ the differential of entropy is

$$dS = \frac{\partial S}{\partial \alpha_1} d\alpha_1 + \ldots \frac{\partial S}{\partial \alpha_k} d\alpha_k \underbrace{= 0}_{\text{at eq.}} .$$

Let ϕ_m be other extensive variable, e.g. $\langle \log(\bar{W}) \rangle$. Its total differential would be

$$d\phi_m = \frac{\partial \phi_m}{\partial \alpha_1} d\alpha_1 + \ldots + \frac{\partial \phi_m}{\partial \alpha_k} d\alpha_k$$

The internal derivative is required to explicitly write the change in variables with respect to our original measure, the entropy: $\frac{\partial \phi_m}{\partial \alpha_j} = \frac{\partial \phi_m}{\partial S} \frac{\partial S}{\partial \alpha_j}$, thus

$$d\phi_m = \frac{\partial \phi_m}{\partial S} \left(\frac{\partial S}{\partial \alpha_1} d\alpha_1 + \ldots + \frac{\partial S}{\partial \alpha_k} d\alpha_k \right)$$

And the quantity in parenthesis is dS. Notice that $\frac{\partial \phi_m}{\partial S} = \alpha_m^{-1}$ which is in general a well defined quantity, thus

$$\begin{aligned} d\phi_m &= \frac{\partial \phi_m}{\partial S}(dS) \\ &= \alpha_m^{-1}(0) \qquad = 0 \ . \end{aligned}$$

Therefore, if entropy is at a maximum, so it is the mean fitness *subject to the constraints of entropy and the other extensive variables.* Furthermore, this also applies to any other extensive variable, like $\langle U \rangle$, for example, or genetic variance, under stabilizing selection.

Appendix B

Notes on the analogy with Statistical Mechanics in Physics

The max-entropic school of statistical mechanics maintains the view that the ensembles in statistical physics show a distribution of energy levels that follows from maximization of entropy. This has been taken as a fundamental principle. However, it is clear that maximization of entropy under constant energy is equivalent as minimization of energy under constant entropy, and they simply conform two different but equivalent and consistent representations of thermodynamics. This is true because a set of known variables of states is uniquely defined by a microscopic distribution. The statistical mechanics formulations had their origins in physics, but the theoretical structure described by these methods (microcanonical, canonical, grand canonical, etc.) have been not only applied to several fields, but proven general for markovian stochastic processes.

Our main concern in this paper, the evolution of quantitative characters under distinct evolutionary forces, has been studied from a max-entropic point of view. This is not the first time that a statistical-mechanics-like approach has been used in the evolutionary context. As mentioned before, we root our ideas in the previous work of PBRS; all of these works, took the statistical mechanics analogy too literal, without taking care of important biological details. Also similar approaches were taken by Iwasa (1988); Kondrashov and Turelli (1992); Barton and Shpak (2000).

Prior distribution, and non equipartition Perhaps the most critical issue, is the choice of the prior distribution. In the hypergeometric model (Kondrashov and Turelli, 1992; Doebeli, 1996a; Shpak and Kondrashov, 1999; Barton and Shpak, 2000), as in PBRS's models, the averages of the ensembles are taken with a uniform weight for any allelic configurations resulting in the same phenotype. Comparing the distribution obtained form the stationary Wright-Fisher model (Eq. 3.1) with the max-entropic distribution (Eq. 3.6) it becomes clear that microscopic states are dependent not only on the value of the trait, but also on the microscopic distribution of the alleles. Therefore, there is no reason to assume that the states are equiprobable. Depending on the selective scheme that is acting over the population, the likelihood of a microstate might also depend on other macroscopic quantities,

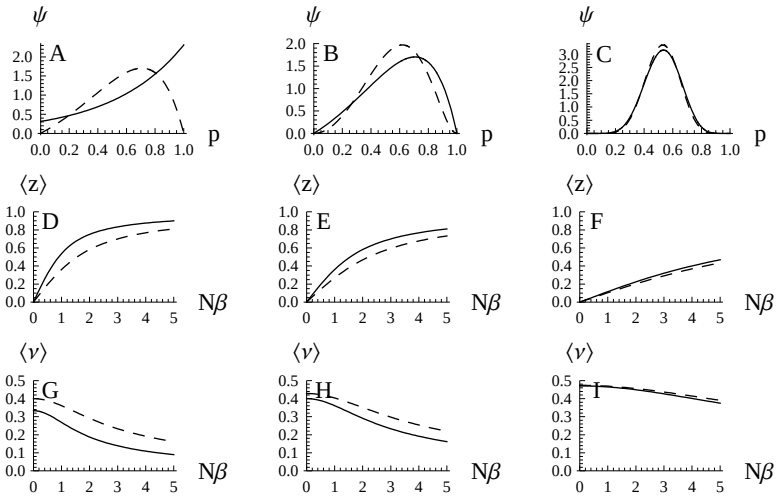


Figure B.1: If we were to assume equipartition (an uniform base distribution $\phi = \text{const.}$ there would be discrepancies in (A-C) the distribution of allele frequencies at each locus ($N\beta = 0.5$), (D-F) the expectations of the trait, and (G-I) the expectation of the genetic variance. In general the discrepancies are less as N/μ is farther from the point $N\mu = 1/4$. (A,D,G) $N\mu = 0.25$, (B,E,H) $N\mu = 0.5$, (C,F,I) $N\mu = 2$.

such as the genetic variance, the variance of the trait, the genetic variability, or other observables. Hence in the entropy measure (Eq. 3.2), the density of states does not reduce to the number of states, as it would be dictated by a uniform distribution, but rather, favour states that in the absence of selection, would result in fixation. In Fig. B.1 we compare the value of the mean trait and genetic variance to the hypothetical case of equipartition of the phase space. Clearly there are deviations from our results, showing the importance of choosing the correct prior distribution.

Intensive and extensive variables Thermodynamics distinguishes between extensive and intensive variables. The formers are those dependent on the amount of matter of the system (total energy, volume, entropy, etc.), and the latters are independent from them (like temperature, pressure, specific heat, etc.). In an extensive system the macroscopic quantities are proportional to the number of particles. The extensiveness in quantitative genetics is given by the number of loci m that are contributing to z. Two physical systems $\mathcal{A}$ and $\mathcal{B}$ with mass M and energy E are equivalent to one system of mass $2M$ and energy $2E$: volume would be doubled, and entropy would be $S_{\mathcal{A}+\mathcal{B}} = S_{\mathcal{A}} + S_{\mathcal{B}}$. This is true also in quantitative genetics, where we have that $\langle \bar{z} \rangle$ and $\langle U \rangle$ are sums over loci, as well as other quantities like genetic variance $\langle \nu_z \rangle$, etc. Doubling the number of loci would double the value of macroscopic variables (if effects are equal; otherwise the extensiveness still applyes to the corresponding effect at each quantity).

Intensive variables, on the other hand, are those that define the constrains of the macrostates. These are selection intensity, mutation rate and population size N. Their value cannot depend on the number of loci. This is a distinction that must be handled with care in our methodology for quanitative genetics; $1/2N$ is analogous to temperature kT (k is Boltzmann' s constant) in physics, in the sense that both reflect the amount of stochastic effects at the micro-level. High temperature indicates strong stochasticity in physical particles. In population genetics, a low number of individual selected by drift show high stochastic fluctuations from generation to generation. Both kT and $1/2N$ result from constraints on the macroscopic systems. But of course, N is itself a quantity that can be confused as extensive. In our problem it has to be regarded as an externally fixed intensive quantity. Actually, the analogy between (the inverse of) temperature and population size reflect the analogous phenomenon: temperature is a measure of the degree of molecular fluctuations, just as population size is a measure of the intensity of drift.

Another aspect that must be clarified in the analogy is about the entities that conform the system. The simplest microscopic system in physics is the ideal monoatomic gas. It consists of a closed reservoir of a big number if independent (not interacting) particles. These particles

are physical entities. The analogy to QG is more abstract. The obvious or natural guess would be that the biologically independent entities are individuals, each having a genetic system of m loci contributing to its phenotype. Our calculations employ a more extravagant definition of what our biological 'particles' are. We consider the average value of each of the m loci across the N individuals as the atomic units. Therefore non-interacting physical particles are analogous to the situation of having infinite recombination and no epistasis, and therefore linkage equilibrium.

Still, the theory of statistical mechanics is robust to the assumption of independency of the particles in the ideal gas, for example. The structure of the theory holds even when interactions are strong (maybe an exception is low temperature physics). Similarly, our calculations are valid with arbitrary degrees of recombination and epistasis. A formal treatment for the former was introduced by Prugel - Bennet (2001). Of course the mathematical simplicity goes hand in hand with the independency assumptions, and therefore they are the cornerstone to the understanding of how a static macroscopic world is built from a continously changing microscopic universe.

Appendix C

Estimations of the genetic architecture of *Rana temporaria*

C.1 GENETIC EFFECTS AND EFFECTIVE NUMBER OF LOCI

C.1.1 Methods

Notice on the one hand from Eq. 4.1 that at $p_\ell = 1/2$, the mean traits will vanish and genetic variances will be maximal, $\tilde{\nu} = \frac{1}{2}\sum_{\ell=1}^{n}\gamma_\ell^2$. This situation is ideally realized at neutrality ($\beta_k = 0$)), but it can be also realized in any distribution of allele frequencies since there is always a non-zero probability that a population has these frequencies $p_\ell = 1/2$. On the other hand, the extreme values that an additive mean trait can have would be achieved at fixation of every locus $p_\ell, q_\ell = 1$, e.g. under extreme selective pressure, and would have values of $\tilde{z} = \pm\sum_{\ell=1}^{n}\gamma_\ell$. It follows that $\frac{2\tilde{\nu}}{n} = \frac{1}{n}\sum_{\ell=1}^{n}\gamma_\ell^2$ and that ,$\frac{\tilde{z}_j}{n} = \frac{1}{n}\sum_{\ell=1}^{n}\gamma_\ell = \bar{\gamma}$. We can apply Jensen's inequality to these quantities: $\frac{1}{n}\sum_{\ell=1}^{n}\gamma_\ell^2 \geq \left(\frac{1}{n}\sum_{\ell=1}^{n}\gamma_\ell\right)^2 \Rightarrow \frac{2\tilde{\nu}}{n} \geq \left(\frac{\tilde{z}}{n}\right)^2$. Thus the estimators for $\underset{\sim}{n}$ and $\bar{\gamma}$ follow from the equalities.

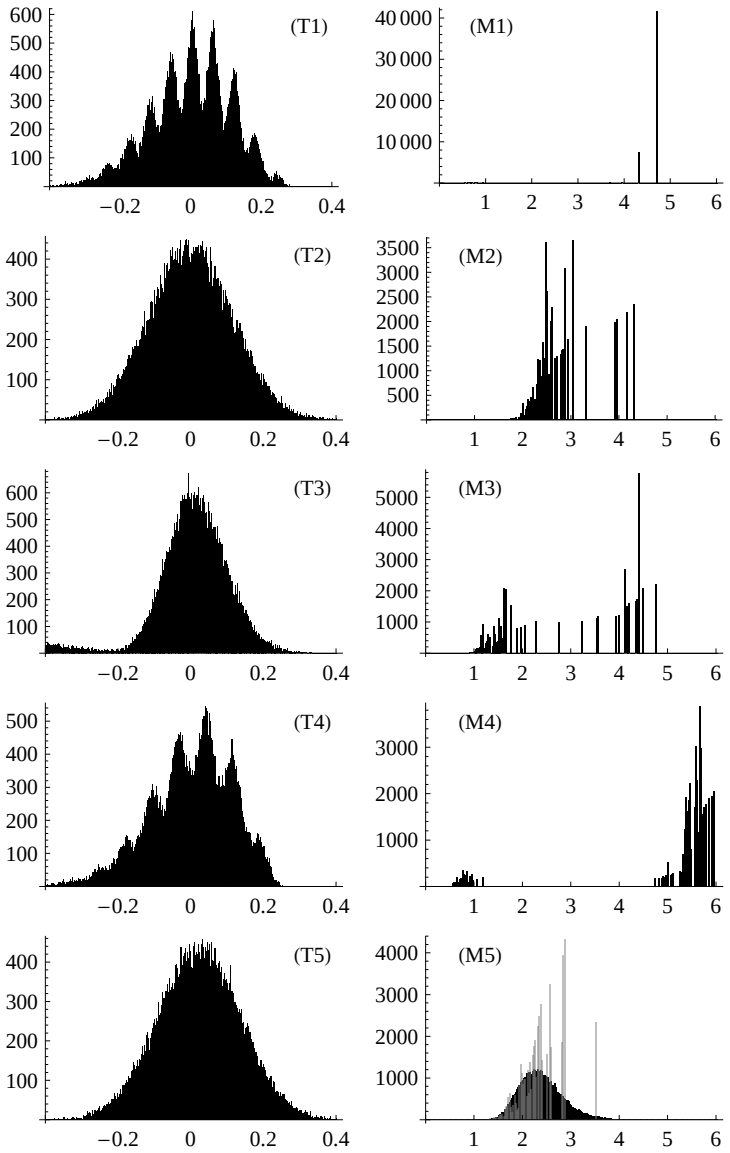
Thus following this results, we proceeded to identify from the data the corresponding quantities. First we pooled that data (1550 F_1 population of 18 full-sib individuals) in order to perform a bootstrap analysis over the maximal trait in a subsample of fixed size $m = 72$ (i.e. the number of experimental replicates). Independent bootstrappings were performed with 5.10^6 iterations for every trait (to avoid inherent constraints or trade-offs), to calculate the mean and variance of the maximal mean trait values. We also computed the maximum value of a sib population (ordering the data, and taking the largest 72 values); and the mean maximum from the experiments. As a conservative estimator we took the mean value from the bootstrap plus half of the standard deviation. This was also compared with the biggest individual sample from the data, keeping the greatest among them.

The bootstraping was compared with an equivalent methodology where the pool of data is substituted by a sample of 1550 random numbers from a normal distribution. Since the pool distribution is finite, the distribution of the maximum of a finite sub-sample deviates substantially from the theoretical expectations.

However, the patterns of this deviation are similar from the pooled data, consistent with the approximate normality of the trait distribution, as dictated by theoretical grounds (Fig C.1).

The deviations of the bootstrapped distribution from the theoretical one, which follows a Gompertzian distribution (Fisher and Tippett, 1928), is evident (Figs. C.1 and C.2), not only because of (i) finite sample effects, but also because (ii) the sub-sample size is small for the requirements of the limit theorem (Fisher and Tippett, 1928), and probably also because (iii) the real distribution of the trait is truncated, property whose consequences over the estimations herein presented were not investigated.

The issue is that in theory the trait has a maximal value that we want to know, but we don't want to overestimate. If we would know the theoretical distribution density of the maximum for a truncated Gaussian (that approximates the mean trait distribution if the number of loci n is large, (Turelli and Barton, 1994), we could compute the deviation of the expectance of the maxima to the real maximum, for replicas of a given size. In this scenario we could state the null hypothesis that the bootstraped mean maximum is significantly equal to the expectance, and from it we could calculate the real maximum. But this distribution is not known, at least to us. Even then, if the mean trait distribution can be approximated by a Gaussian, then the cumulative distribution of the maximal $\tilde{z}$ would be Gompertzian, that is of the form Exp $\left[ae^{-b\tilde{z}}\right]$, where the parameters a and b depend on the number of samples m (Fisher and Tippett, 1928). Therefore the expected value of the maximum in a Gaussian distribution would be bigger than the expected maximum in a truncated Gaussian. Knowing the latter would give at least an upper bound for overestimations. Unfortunately this is a limit theorem with very slow convergence with m and for the data employed here, the Gompertz distribution highly underestimates the actual one (Fig C.2).


(T1)
(M1)
(T2)
(M2)
(T3)
(M3)
(T4)
(M4)
(T5)
(M5)

Estimating the biggest genetic variance, suffered from more serious problems, since the values that we have were inferred from an animal model (Lynch and Walsh, 1998, pp 755-758) that eliminated maternal and dominance effects. Thus they do not follow directly as the variance of the raw data. As a first subterfuge, we calculated the total variance, assuming that the data are subsamples of the pool, thus $\mathrm{Var}_{\mathrm{tot}} = \mathrm{Mean}(\nu) + \mathrm{Var}\,(\bar{z})$. Second, we employed the data of both populations, their controls, and both treatments (fast and slow dessication), for a total of 6 points for each trait, to regress the mean trait value and genetic variance calculations. From it, we took the intercept of the regression and added half of the standard error (square root of the sum of squares of the regression). The maximum between these two approaches, together with the maximal observed genetic variance, was taken as the estimator of $\tilde{\nu}$.

C.1.2 Results

The maximal trait values estimated from the resampling proved higher than the experimental ones for traits 1 and 4, while for traits 2 and 3 the empirical maxima were used (Table 1). The total genetic variance for an idealized pool resulted smaller than the maximal empirical genetic variances, except for trait 3, with which the difference is very small. The method with least squares always gave a smaller maximal genetic variance Fig. C.2.

Figure C.1: (Opposite page) Bootstrapping results. The left column (T) shows the distributions of the four mean traits, and the right column (M) the distribution of their maximum values. The distributions were calculated bootstrapping the pooled data, with a sub-sampling size of m=72. The mean traits were centered in the average, and scaled by the variance of the pool (to approximate normality). From top to bottom: (1) develpment time, (2) mass, (3) body length, (4) tail length, and (5) the null hypothesis (samples from a normal distribution). Notice that the resulting distribution of the maximum in M5 differ when the distribution can be resampled infinitely, leading to a Gompertzian distribution (gray bulk density), to when they are finite and resampled (gray overimposed bars).

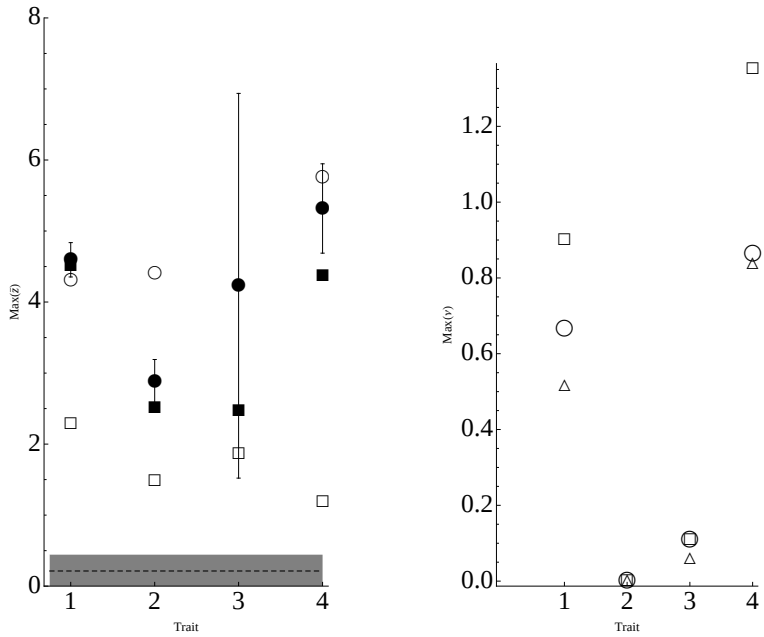


Figure C.2: Average maxima of the mean traits computed from subsampling the bootstrap (black dots, bars indicate standard error), maximum subsample from the pool (black squares), maximum mean trait from the population (open squares), and maximum sampled mean (open circle). The dotted line represents the expected value of the maximum following the theoretical predictions, and the gray shade extends up to half the standard deviation. (Right) Maximum genetic variance. Pooled variance (open circles), maximal variance in the populations (open squares) and maximal variance by linear regression (open triangles). The traits are represented as (1) development time, (2) mass, (3) body length, (3) tail length.

C.2 GENETIC ARCHITECTURE AND M.S.D. EQUILIBRIUM

C.2.1 Methods

Assuming that the effects of each loci over a given trait is equal to the average effect, the possible combinations of pleiotropic structures are drastically reduced. Still, if the number of loci is big, which in practice would be on the order of 10 or more, finding all combinations is a tedious enterprise. Thus we designed an algorithm to search a representative subset of these pleiotropic structures. The contributing loci are represented in an array whose positions can either be empty, or occupied. The array has two dimensions. The first dimension, say the rows, represents the different traits, while the second dimension, the columns are slots where the contributing loci can be. The second dimension can be as big as the sum of the total number of loci contributing to all traits. That is no overlap among loci, and therefore the traits are independent. The other extreme, when there is maximum degree of overlap (pleiotropy is strong) is when the second dimension is as small as the bigest number of loci that contribute to any of the traits. Any case in between represents partial overlap between some of the loci. Thus distinct structures with varying degrees of pleiotropy are generated by randomly moving and shifting these loci to different slots in each row. The movements in the algorithm are: a row of loci (that is all the effects for one trait) can be shifted left or right one step; the other alternative is to shift only one of the loci to an empty slot on its right or left. To optimize the algorithm, we assigned probabilities for the shifting of a locus, proportional to how fragmented a row of loci is (that is how many empty slots among contributing loci). If the row is too fragmented, it tends to condense, and if it is a single block, it tends to fragment. In addition, empty columns are deleted, since they do not represent any biological effect. Since all the effects over a given trait are the same at all loci, the algotithmic exploration of the pleiotropic configurations discards those which have a degree of overlap that was already sampled (they are redundant in macroscopic terms). Perform-

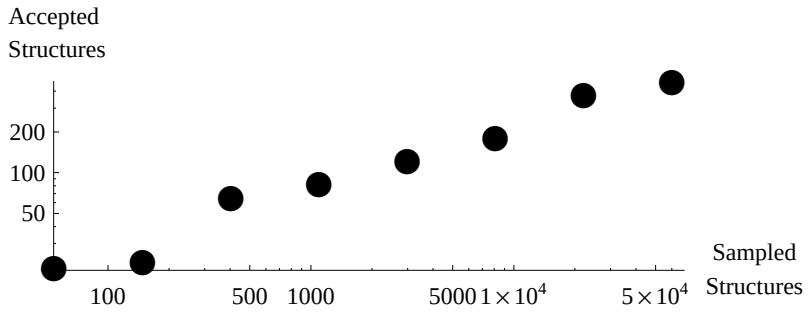


Figure C.3: Performance of the Monte Carlo method for generating new possible pleiotropic structures.

ing a random search with these rules allows to find a big number of possible combinations.

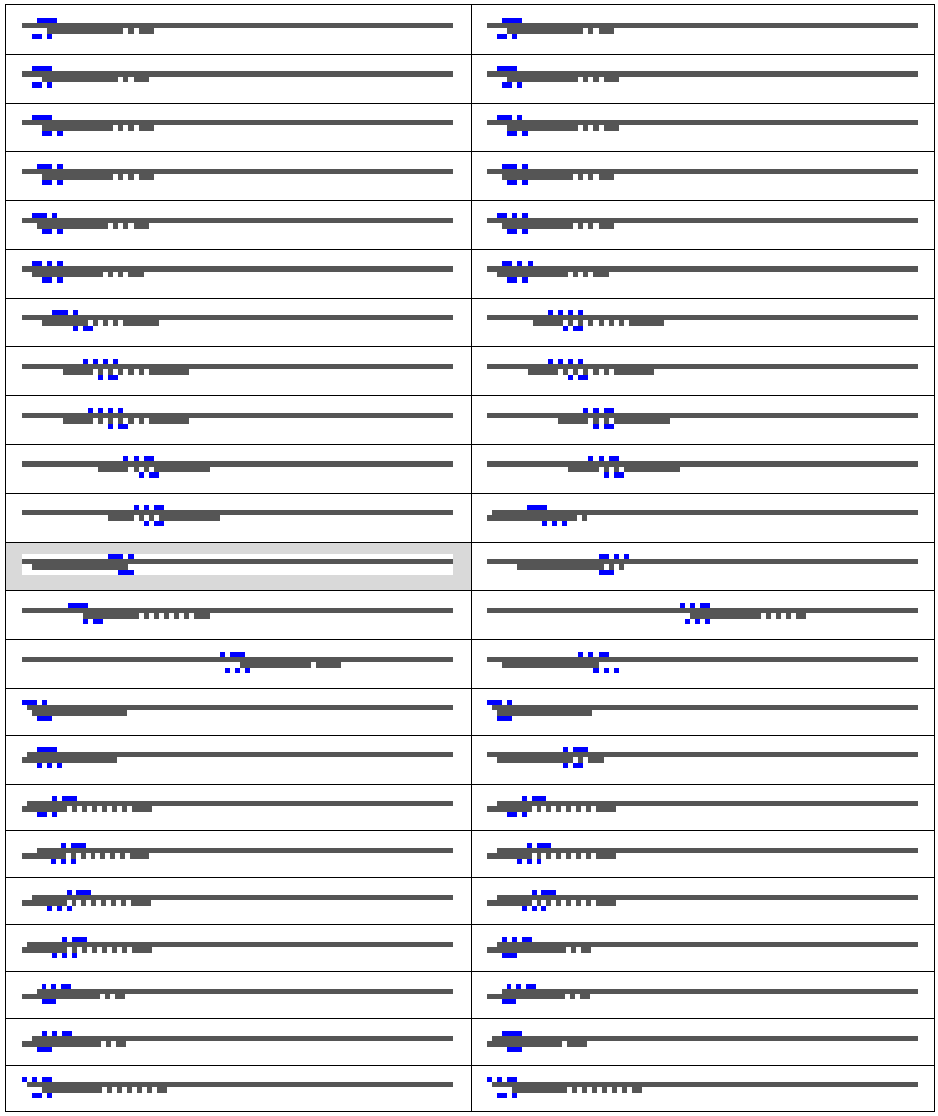
C.2.2 Results

Figure C.3 shows the results of the Monte Carlo procedure to identify the pleiotropic structures. Although it is clear that more structures can be found, the idea is to have a representative sample from the space. Thus in our analyses we employed only 60000 trial structures. Supplementary Material B illustrates this search an animated representation. As a rule of thumb, about 5% of the tested structures will be considered.

C.3 PLEIOTROPIC STRUCTURES AND FITNESS DIFFERENTIALS (β)

Of the 450 structures that were produced with the previous algorithm, only 47 allowed convergence of the numerical fit of the empirical traits with the theoretical formulas. When numerical solutions could not be found with precision of at lest 3 significant digits, the structures were discarded as impossible to allow a fit between microscopic and macroscopic states. However, those structures for which there was convergence, had a precision of at least 10^{-8}. Furthermore, although the estimations were performed independently for the data of each locations, the consistent pleiotropic structures were in both cases the same (Fig C.4).

Nevertheless, the estimations gave distinct distribution of β at each location, as seen in Fig C.5. Notice that the distribution changes for both locations, indicating the action of selection.



C.4 EIGENSTRUCTURES OF THE $\mathcal{G}$ MATRICES

A consistent way to compare matrices, is to compare their eigenstructure. (There are many other ways, by the way, as well as statistical tests. But it looks the choice on the method remains arbitrary.) We computed the eigenstructures for the empirical, expectances, and drift samples matrices. In Fig C.6. it is revealed that the eigenvalues of the expectancies $\langle G \rangle$, in most cases are close to those of the empirical $\mathcal{G}$ s, except only for the third eigenvalue, and the fourth for the Southern population.

Figure C.4: (Opposite page) Pleiotropic structures consistent with the data. Each of these pleiotropic structures correspond to an estimated value of $\vec{\beta}$. Notice that most of them represent a high pleiotropic coupling of the traits. The rows in each plot represent the effects of the loci over the first to the fourth traits (development time, (2) mass, (3) body length, and (4) tail length. The structure shaded in gray was the one employed to forecast the evolutionary dynamics in Fig. 4.1.

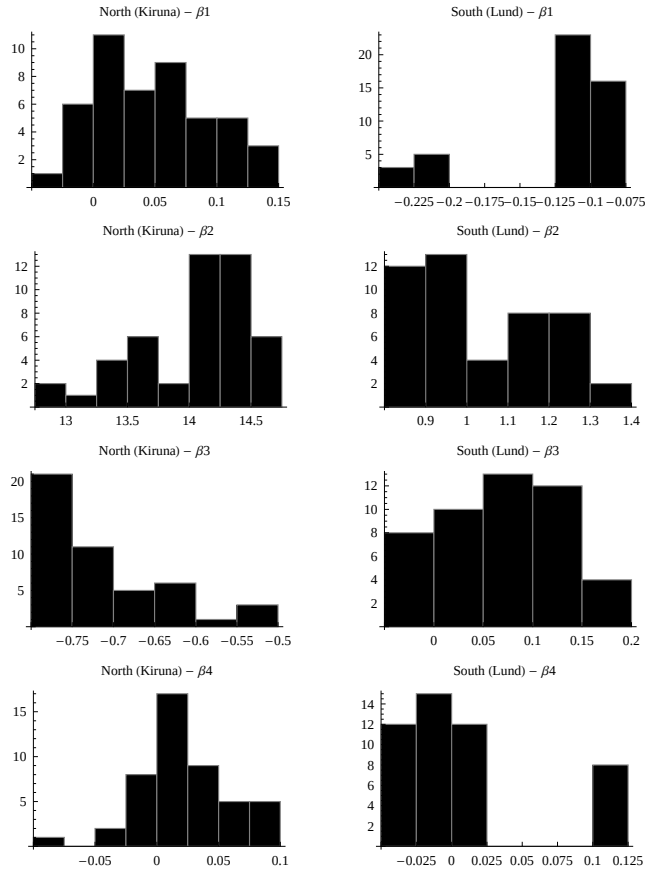


Figure C.5: (Opposite page) Distribution of selection coefficients for each location: Kiruna (Northen population), first column, and Lund (Southern population) second column. Each histogram βk, represents the distribution of estimated selection coefficients for each trait k: (1) development time, (2) mass, (3) body length, and (4) tail length.

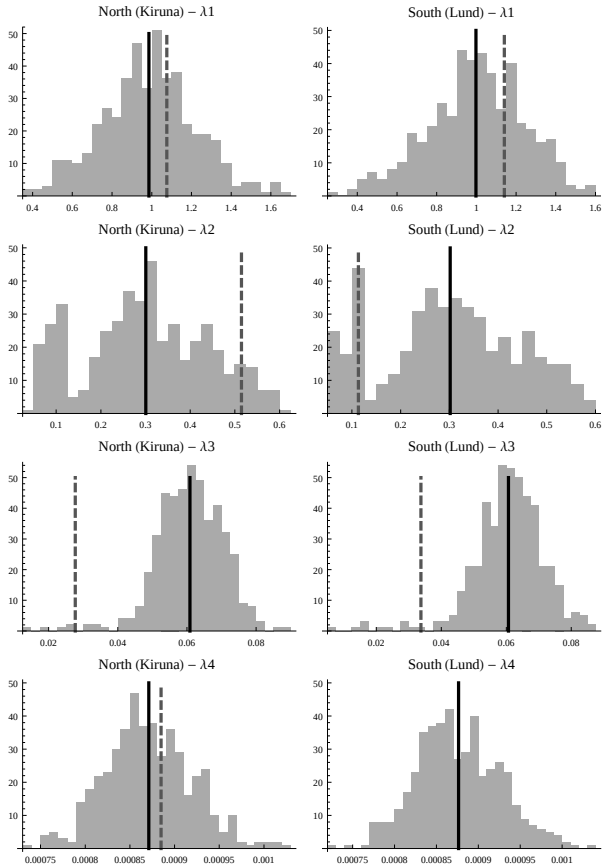


Figure C.6: Distribution of the eigenvalues of $\mathcal{G}$ for each location: Kiruna (Northern population), first column, and Lund (Southern population) second column. Each histogram λk, represents the distribution of k'th eigenvalues of the $\mathcal{G}$ matrices resampled from the distribution of allele frequencies, where the components k are: (1) development time, (2) mass, (3) body length, and (4) tail length. The gray dashed lines are the eigenvalues of the empirical $\mathcal{G}$. The black solid lines are the eigenvalues of the expectancies.

Appendix D

Formulas for the quantitative variables

D.1 DIRECTIONAL SELECTION

The main goal, from the technical side, is to have explicit formulas for the expectations of the macroscopic variables, in order to be able to make predictions of the course of evolution. That is, to put the statistical mechanical method 'at work'. The formulas for polygenic traits are closely related to those of a single locus, which I will give in the following pages. The polygenic versions will be given in the following sub-section.

Remarks about the notation. The 'per locus' formulas will be always denoted with the subscript ℓ. The formulas can be represented by either regularized confluent hypergeometrics ${}_0F_1\left(.;.\right)$, or modified Bessel functions of fractional order $\mathcal{I}_\nu\left(.\right)$. Although entirely equivalent, for some of the formulas I will provide both expressions. It is just a matter of taste which one to use. In general I am using variables scaled by population size, so mutation rate and selective values will be expressed as $N\mu$ and $N\beta$. Although simplifications with respect to N is often possible, I will leave all expressions in terms of the scaled variables. The reason is that the scaling *always* appears, and population size N rarely appears on its own.

The formulas here defined, as well as the evolutionary dynamics are implemented in Mathematica 6.0 packages, that are available upon request.

D.1.1 Statistics for traits with a single locus

Distribution of allele frequencies and the partition function

- The maxentropic distribution is

$$\psi = \frac{\phi}{\mathbb{Z}} \exp\left[2N\beta\bar{z} + 2N\mu U\right] \tag{D.1}$$

where the base density is

$$\phi = 1/p(1-p) \tag{D.2}$$

- The partition function is

$$\begin{aligned} \mathbb{Z}_{\ell(N\mu,N\beta)} &= \mathcal{I}_{4N\mu+\frac{1}{2}}\left(2N\beta\right) \\ &= 2^{1-8N\mu}\Gamma(4N\mu)\ {}_0F_1\left(4N\mu+\frac{1}{2};(N\beta)^2\right) \end{aligned} \tag{D.3}$$

  which is also the generating function for the macroscopics defined below.

— *Low mutation rates limit ($4N\mu < 1$).* The distribution has only two peaks at $\psi(0)$ and $\psi(1)$, and the partition function is

$$\mathbb{Z}_{\ell(N\beta)} = \cosh(2N\beta)\ . \tag{D.4}$$

Genetic mutation variability

- Definition

$$U = 2\log[p(1-p)] \tag{D.5}$$

- Expectancy

$$\begin{aligned} \langle U\rangle_{\ell(N\mu,N\beta)} &= \frac{\partial\log(\mathbb{Z}_{\ell(N\mu,N\beta)})}{2N\partial\mu} \\ = 2\Psi(4N\mu)-\log(16) &+ 2\frac{{}_0F_1^{(1,0)}\left(4N\mu+\frac{1}{2};(N\beta)^2\right)}{{}_0F_1\left(4N\mu+\frac{1}{2};(N\beta)^2\right)} \end{aligned} \tag{D.6}$$

  where Ψ is the digamma function. The equation is to be computed numerically, or approximated at small β, for which it simplifies as

$$\langle U\rangle \simeq \Psi(8N\mu)-\log(16)+\left(\frac{\beta}{4N\mu+1/2}\right)^2\ .$$

— *Low mutation rates limit ($4N\mu < 1$).* This macroscopic does not need to be defined for low mutation rates to couple the distribution of the micro and macroscopic variables.

- Variance

$$\mathrm{var}(U)_{\ell(N\mu,N\beta)} = \frac{\partial^2 \log(\mathbb{Z}_{\ell(N\mu,N\beta)})}{(2N)^2\partial\mu^2} = \frac{\partial\langle U\rangle}{2N\partial\mu} \tag{D.7}$$

$$= 4\Psi'(4N\mu) - 4\frac{{}_0F_1^{(2,0)}\left(4N\mu+1/2;(N\beta)^2\right)}{{}_0F_1\left(4N\mu+1/2;(N\beta)^2\right)} \tag{D.8}$$

$$+ 4\left[\frac{{}_0F_1^{(1,0)}\left(N\mu+1/2;(N\beta)^2\right)}{{}_0F_1\left(4N\mu+1/2;(N\beta)^2\right)}\right]^2$$

'Genetic variance' of genetic variability This quantity is defined by the rate of change of U (Eq. 3.22), in analogy to the rate of change of the mean trait, that is proportional to genetic variance:

- Definition

$$H = 2(p(1-p))^{-1} - 4 \tag{D.9}$$

- Expectancy

$$\langle H\rangle_{\ell(N\mu,N\beta)} = -8 + 2\frac{\mathbb{Z}_{\ell\ (N\mu-1/4,N\beta)}}{\mathbb{Z}_{\ell(N\mu,N\beta))}} \tag{D.10}$$

$$= -4 + \frac{1}{4N\mu-1/4}\frac{{}_0F_1\left(4N\mu-1/2;(N\beta)^2\right)}{{}_0F_1\left(4N\mu+1/2;(N\beta)^2\right)}$$

$$= 4\frac{N\mu+1/4}{N\mu-1/4} + \frac{N\beta\langle\bar{z}\rangle_{\ell(N\mu,N\beta)}}{(N\mu-1/2)(N\mu+1/2)} \tag{D.11}$$

— *Low mutation rates limit ($4N\mu < 1$).* As $4N\mu \downarrow 1$, $\langle U\rangle \uparrow \infty$. This macroscopic however does not need to be defined for low mutation rates to couple the dynamics of the micro and macroscopic variables.

Mean Trait

- Definition

$$\bar{z} = 2(p-1) \tag{D.12}$$

- Expectancy

$$\begin{aligned}\langle \bar{z} \rangle_{\ell(N\mu,N\beta)} &= \frac{\partial \log(\mathbb{Z}_{\ell(N\mu,N\beta)})}{2N\partial\beta} \\ &= \frac{\mathcal{I}_{4N\mu+\frac{1}{2}}(2N\beta)}{\mathcal{I}_{4N\mu-\frac{1}{2}}(2N\beta)} \\ &= N\beta \frac{{}_0F_1\left(4N\mu+\frac{3}{2};(N\beta)^2\right)}{{}_0F_1\left(4N\mu+\frac{1}{2};(N\beta)^2\right)}\end{aligned} \tag{D.13}$$

— *Low mutation rates limit* ($4N\mu < 1$).

$$\langle \bar{z} \rangle_{\ell(N\beta)} = \tanh(2N\beta) \tag{D.14}$$

- Variance

$$\mathrm{var}(\bar{z})_{\ell(N\mu,N\beta)} \frac{\partial^2 \log(\mathbb{Z}_{\ell(N\mu,N\beta)})}{(2N)^2\partial\beta^2} = \frac{\partial\langle \bar{z} \rangle}{2N\partial\beta} \tag{D.15}$$

$$\begin{aligned}&= 1 - \frac{\mathcal{I}_{4N\mu+1/2}(2N\beta)}{\mathcal{I}_{4N\mu-1/2}(2N\beta)} \left(\frac{4N\mu}{2N\beta} + \frac{\mathcal{I}_{4N\mu+1/2}(2N\beta)}{\mathcal{I}_{4N\mu-1/2}(2N\beta)} \right) \\ &= 1 - \langle \bar{z} \rangle \left(\frac{4N\mu}{2N\beta} + \langle \bar{z} \rangle \right)\end{aligned} \tag{D.16}$$

— *Low mutation rates limit* ($4N\mu < 1$).

$$\mathrm{var}(\bar{z})_{\ell(N\beta)} = \mathrm{sech}^2(2N\beta) \tag{D.17}$$

Genetic variance

- Definition

$$\nu = 2p(1-p) \tag{D.18}$$

- Expectancy

$$\langle \nu \rangle_{\ell(N\mu,N\beta)} = 2\frac{\mathbb{Z}_{\ell\ (N\mu+1/4,N\beta)}}{\mathbb{Z}_{\ell(N\mu,N\beta))}} \tag{D.19}$$

$$\begin{aligned}&= 2N\mu \frac{{}_0F_1\left(4N\mu+\frac{3}{2};(N\beta)^2\right)}{{}_0F_1\left(4N\mu+\frac{1}{2};(N\beta)^2\right)} \\ &= \frac{2N\mu}{N\beta}\langle \bar{z} \rangle_{\ell(N\mu,N\beta)}\end{aligned} \tag{D.20}$$

— *Low mutation rates limit ($4N\mu < 1$).*

$$\langle \nu \rangle_{\ell(N\mu,N\beta)} = \frac{2N\mu}{N\beta} \tanh(2N\beta) \tag{D.21}$$

Covariance between meant trait and genetic variability

$$\begin{aligned} \mathrm{cov}(\bar{z}, U)_{\ell(N\mu,N\beta)} &= \frac{\partial^2 \log(\mathbb{Z}_{\ell(N\mu,N\beta)})}{(2N)^2 \partial\beta\partial\mu} \\ &= \frac{\partial\langle\bar{z}\rangle}{2N\partial\mu} = \frac{\partial\langle U\rangle}{2N\partial\beta} = \mathrm{cov}(U,\bar{z}) \\ &= 2\left[\frac{\mathcal{I}^{(1,0)}_{4N\mu+1/2}(2N\beta)}{\mathcal{I}_{4N\mu-1/2}(2N\beta)} - \frac{\mathcal{I}_{4N\mu+1/2}(2N\beta)\,\mathcal{I}^{(1,0)}_{4N\mu-1/2}(2N\beta)}{\left(\mathcal{I}_{4N\mu-1/2}(2N\beta)\right)^2}\right] \end{aligned} \tag{D.22}$$

D.1.2 Statistics for multivariate traits with unequal (and pleiotropic) effects

The statistics for the moments of a vector of mean traits $\langle\vec{z}\rangle$ results as an extension of the per-locus functions. I will give only the multivariate expressions, since the uni-variate traits are a special case. For the latter, some properties (similar as in the one-locus case) hold, which I will highlight.

Remarks about the notation. The 'per locus' sub-index will be used here also as an iterator (index for the summations and products). Hence, whenever a quantity has a subindex, it refers to the 'per locus' quantity evaluated at the effects of a given locus. These effects of each locus ℓ over every trait k are summarized in a matrix Υ; each column of this matrix, $\vec{\gamma}_\ell$ contains the effects of one locus ℓ over each of the m traits; the effect of a locus ℓ over a trait k is then $\gamma_{k\ell}$, which for the univariate case will simply be written as γ_ℓ. :

$$\Upsilon = \begin{array}{c} \vdash \text{loci} \longrightarrow \\ \left(\begin{array}{cccccc} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1\ell} & \cdots & \gamma_{1n} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2\ell} & \cdots & \gamma_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \gamma_{m1} & \gamma_{m2} & \cdots & \gamma_{m\ell} & \cdots & \gamma_{mn} \end{array}\right) \end{array} \begin{array}{c} \vec{\gamma}_\ell \\ \\ \top \\ \text{traits} \\ \downarrow \end{array}$$

Similarly, each row of Υ represent the vector of effects of all loci over each of the traits.

The symbol $\otimes$ represents the Kronecker (tensor, dyadic, or external) product, and will appear when using matrix notation.

Partition function

- Partition function

$$\mathbb{Z}_{(N\mu,N\vec{\beta}|\Upsilon)} = \prod_{\ell=1}^{n} \mathbb{Z}_{\ell(N\mu,N\vec{\gamma}_\ell\cdot\vec{\beta})} \tag{D.23}$$

— *Low mutation rates limit ($4N\mu < 1$).*

$$\mathbb{Z}_{(N\vec{\beta}|\Upsilon)} = \prod_{\ell=1}^{n} \cosh[2N\vec{\gamma}_\ell \cdot \vec{\beta}] \tag{D.24}$$

Genetic mutation variability

- Definition

$$U = 2\sum_{\ell=1}^{n} \log[p_\ell(1-p_\ell)] \tag{D.25}$$

- Expectancy

$$\begin{aligned} \langle U\rangle_{(N\mu,N\vec{\beta}|\Upsilon)} &= \frac{\partial \log(\mathbb{Z}_{(N\mu,N\vec{\beta}|\Upsilon)})}{2N\partial\mu} \qquad \text{(D.26)} \\ &= 2\sum_{\ell=1}^{n} \langle U\rangle_{\ell(N\mu,N\vec{\gamma}_\ell\cdot\vec{\beta})} \end{aligned}$$

- Variance

$$\begin{aligned} \mathrm{var}(U)_{(N\mu,N\vec{\beta}|\Upsilon)} &= \frac{\partial^2 \log(\mathbb{Z}_{(N\mu,N\vec{\beta}|\Upsilon)})}{(2N)^2\partial\mu^2} \qquad \text{(D.27)} \\ &= \frac{\partial\langle U\rangle}{2N\partial\mu} \\ &= \sum_{\ell=1}^{n} \mathrm{var}(U)_{\ell(N\mu,N\vec{\gamma}_\ell\cdot\vec{\beta})} \qquad \text{(D.28)} \end{aligned}$$

'Genetic variance' of genetic variability

- Definition

$$H = \sum_{\ell=1}^{n} 2[p_\ell(1-p_\ell)]^{-1} - 4 \tag{D.29}$$

- Expectancy

$$\langle H \rangle_{(N\mu, N\vec{\beta}|\Upsilon)} = 2\frac{\mathbb{Z}_{(N\mu-1/4, N\vec{\beta}|\Upsilon)}}{\mathbb{Z}_{(N\mu, N\vec{\beta}|\Upsilon)}} - 4n \tag{D.30}$$

$$= -4n + 2(N\mu - 1/4)^{-1} \sum_{\ell=1}^{n} \frac{{}_0F_1\left(4N\mu - 1/2; (N\vec{\gamma}_\ell \cdot \vec{\beta})^2\right)}{{}_0F_1\left(4N\mu + 1/2; (N\vec{\gamma}_\ell \cdot \vec{\beta})^2\right)}$$

$$= 4n\frac{N\mu + 1/4}{N\mu - 1/4} + \frac{N\vec{\beta} \cdot \langle \vec{z} \rangle_{(N\mu, N\vec{\beta}|\Upsilon)}}{(N\mu - 1/2)(N\mu + 1/2)} \tag{D.31}$$

Mean traits

- Defintion

$$\vec{z} = \sum_{\ell=1}^{n} \vec{\gamma}_\ell (2p_\ell - 1) \tag{D.32}$$

- Expectancies

$$\langle \vec{z} \rangle_{(N\mu, N\vec{\beta}|\Upsilon)} = \frac{\partial \log(\mathbb{Z}_{(N\mu, N\vec{\beta}|\Upsilon)})}{2N\partial\vec{\beta}} \tag{D.33}$$

$$= \sum_{\ell=1}^{n} \vec{\gamma}_\ell \frac{\mathcal{I}_{4N\mu+\frac{1}{2}}\left(2N\vec{\gamma}_\ell \cdot \vec{\beta}\right)}{\mathcal{I}_{4N\mu-\frac{1}{2}}\left(2N\vec{\gamma}_\ell \cdot \vec{\beta}\right)}$$

$$= \sum_{\ell=1}^{n} \vec{\gamma}_\ell \langle \bar{z} \rangle_{\ell(N\mu, N\vec{\gamma}_\ell \cdot \vec{\beta})} \tag{D.34}$$

— *Low mutation rates limit* ($4N\mu < 1$).

$$\langle \vec{z} \rangle_{(N\vec{\beta}|\Upsilon)} = \sum_{\ell=1}^{n} \vec{\gamma}_\ell \tanh[2N\vec{\gamma}_\ell \cdot \vec{\beta}] \tag{D.35}$$

- Covariances of the mean traits.

$$\begin{aligned}\mathrm{covar}(\bar{z}_j, \bar{z_k})_{(N\mu, N\vec{\beta}|\Upsilon)} =& \frac{\partial \log(\mathbb{Z}_{(N\mu, N\vec{\beta}|\Upsilon)})}{(2N)^2 \partial\beta_j \partial\beta_k} \qquad \text{(D.36)}\\ =& \frac{\partial\langle\bar{z}_j\rangle}{2N\partial\beta_k} = \frac{\partial\langle\bar{z}_k\rangle}{2N\partial\beta_j}\\ =\ 2(\nu_{jk}^{\max} - \langle\nu_{jk}\rangle) &- \sum_{\ell=1}^{n} \gamma_{j\ell}\gamma_{k\ell}\langle\bar{z}\rangle^2_{\ell(N\mu, N\vec{\gamma}_\ell\cdot\vec{\beta})} \qquad \text{(D.37)}\end{aligned}$$

  where

$$\nu_{jk}^{\max} = \frac{1}{2}\sum_{\ell=1}^{n}\gamma_{j\ell}\gamma_{k\ell} \qquad \text{(D.38)}$$

  are the maximal genetic variances.

- Variances of the mean traits. The variance of a trait follows directly for $j = k$ in Eq. D.36. For univariate traits of equal effects

$$\mathrm{var}(\bar{z}) = \nu^{\max} - \frac{2N\mu}{N\beta}\langle\bar{z}\rangle - \langle\bar{z}\rangle^2 \ . \qquad \text{(D.39)}$$

  We can express the covariances in matrix form, as it is needed for the dynamics; $\mathcal{C} = \{\mathrm{covar}(\bar{z}_j, \bar{z}_k)\}_{j,k=1}^{m}$ which results in:

$$\mathcal{C} = 2(\mathcal{G}^{\max} - \mathcal{G}) - \sum_{\ell=1}^{n}\vec{\gamma}_\ell \otimes \vec{\gamma}_\ell \langle\bar{z}\rangle^2_{\ell(N\mu, N\vec{\gamma}_\ell\cdot\vec{\beta})} \ , \qquad \text{(D.40)}$$

  where the relation between the $\mathcal{G}$-matrix and the variance across phenotypes is revealed.

— *Low mutation rates limit ($4N\mu < 1$).*

$$\mathcal{C}_{(N\vec{\beta}|\Upsilon)} = \sum_{\ell=1}^{n}\vec{\gamma}_\ell \otimes \vec{\gamma}_\ell \mathrm{sech}^2[2N\vec{\gamma}_\ell\cdot\vec{\beta}] \qquad \text{(D.41)}$$

Genetic variances, covariances, and the $\mathcal{G}$-matrix

- Definition

$$\nu_{jk} = 2\sum_{\ell=1}^{n}\gamma_{j\ell}\gamma_{k\ell}p_\ell(1-p_\ell) \qquad \text{(D.42)}$$

- Expectancies of the covariances

$$\langle \nu_{jk} \rangle_{(N\mu, N\vec{\beta}|\Upsilon)} = 2\frac{\mathbb{Z}_{(N\mu+1/4, N\vec{\beta}|\Upsilon)}}{\mathbb{Z}_{(N\mu, N\vec{\beta}|\Upsilon)}} \tag{D.43}$$

$$= 2N\mu \sum_{\ell=1}^{n} \frac{\gamma_{j\ell}\gamma_{k\ell}}{2N\vec{\gamma}_\ell \cdot \vec{\beta}} \langle \bar{z} \rangle_{\ell(N\mu, N\vec{\gamma}_\ell \cdot \vec{\beta})} \tag{D.44}$$

- Expectancies of the genetic variances. These are defined by

$$\langle \nu_k \rangle \equiv \langle \nu_{kk} \rangle \ . \tag{D.45}$$

  In the case of univariate traits, setting $(m = 1)$, comparing Eqns. D.32 and D.42 it holds true that

$$\langle \nu \rangle = \frac{2N\mu}{N\beta} \langle \bar{z} \rangle \ , \tag{D.46}$$

  for arbitrary number of loci and effects.

- Definition of the $\mathcal{G}$-matrix. Expressed in matrix form, the covariances are

$$\mathcal{G} = \{\nu_{jk}\}_{j,k=1}^{m} = \begin{pmatrix} \nu_1 & \nu_{12} & \cdots & \nu_{1m} \\ \nu_{21} & \nu_2 & \cdots & \nu_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \nu_{m1} & \nu_{m2} & \cdots & \nu_m \end{pmatrix} \tag{D.47}$$

- Expectancies

$$\langle \mathcal{G} \rangle_{(N\mu, N\vec{\beta}|\Upsilon)} = \left\langle \begin{matrix} \nu_1 & \nu_{12} & \cdots & \nu_{1m} \\ \nu_{21} & \nu_2 & \cdots & \nu_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \nu_{m1} & \nu_{m2} & \cdots & \nu_m \end{matrix} \right\rangle \tag{D.48}$$

$$= 2N\mu \sum_{\ell=1}^{n} \vec{\gamma}_\ell \otimes \vec{\gamma}_\ell \frac{\langle \bar{z} \rangle_{\ell(N\mu, N\vec{\gamma}_\ell \cdot \vec{\beta})}}{2N\vec{\gamma}_\ell \cdot \vec{\beta}} \tag{D.49}$$

— *Low mutation rates limit* $(4N\mu < 1)$.

$$\langle \mathcal{G} \rangle_{(N\mu, N\vec{\beta}|\Upsilon)} = 2N\mu \sum_{\ell=1}^{n} \frac{\vec{\gamma}_\ell \otimes \vec{\gamma}_\ell}{2N\vec{\gamma}_\ell \cdot \vec{\beta}} \tanh[2N\vec{\gamma}_\ell \cdot \vec{\beta}] \tag{D.50}$$

- Variance of the genetic co-variances

$$\mathrm{var}(\nu_{jk}) = 4\sum_{\ell=1}^{n} \vec{\gamma}_\ell^2 \otimes \vec{\gamma}_\ell^2 \frac{\mathbb{Z}_{(N\mu+1/2,2N\vec{\gamma}_\ell\cdot\vec{\beta})}}{\mathbb{Z}_{(N\mu,N\vec{\beta}|\Upsilon)}} \tag{D.51}$$

Covariance between meant trait and genetic variability

$$\begin{aligned}
\mathrm{covar}(\vec{z},U)_{(N\mu,N\vec{\beta}|\Upsilon)} =& \frac{\partial^2 \log(\mathbb{Z}_{(N\mu,N\vec{\beta}|\Upsilon)})}{(2N)^2\partial\vec{\beta}\partial\mu} && \text{(D.52)}\\
=& \frac{\partial\langle U\rangle}{2N\partial\vec{\beta}} = \frac{\partial\langle\vec{z}\rangle}{2N\partial\mu} && \text{(D.53)}\\
=& \sum_{\ell=1}^{n} \vec{\gamma}_\ell \mathrm{covar}(\bar{z},U)_{\ell(N\mu,N\vec{\gamma}_\ell\cdot\vec{\beta})} \ . && \text{(D.54)}
\end{aligned}$$

Notice that this is a vector of the covariances between each mean trait and U.

D.2 STABILIZING SELECTION

Similar to the case of directional selection, in linkage equilibrium the statistics for polygenic characters depend on those of single loci, so I will first review these, and in the following subsection, the polygenic formulas.

Remarks about the notation. For clarity in reading the formulas I will use the following convention. The single locus formula will always be denoted by the symbol $\langle \ldots \rangle_\ell$; the subscript ℓ is used to indicate single locus statistic. In the polygenic formulas this subscript will also be used as an iterator across the loci. Furthermore, the statistics for single locus are function of three parameters, that is the vector $(N\mu, N\beta, N\sigma)$. Also for notational simplicity this vector will be omitted, bearing in mind that single locus statistics require these three variables. The statistics for polygenic systems will not have any subscript. These statistics depend on the vector $(N\mu, N\beta, N\sigma, N\alpha)$ which will also be left implicit. However, some formulas (one and multiple loci) sometimes require evaluation of the parameters at different values. In those cases I will express this dependence explicitly. For example if I were to evaluate a given statistic X at a mutation rate $N\mu + 1$ (as it will be necessary to point out some properties and simplifications on the implementations) and at a selective gradient of $\beta\gamma$ I'd write $\langle X \rangle_{(N\mu = N\mu + 1, N\beta = N\beta\gamma)}$, where the rest of the parameters are left untouched (in this case, $N\sigma$ and $N\alpha$).

D.2.1 Statistics for traits with a single locus

The distribution of allele frequencies and the partition function

- Distribution of allele frequencies

$$\begin{aligned} \psi_{\ell(p;N\mu,N\beta,N\sigma)} &:= \psi_\ell \\ &= \frac{\phi}{\mathbb{Z}_\ell} \exp\left[2N\beta(2p-1) + 4N\sigma p(1-p) + 4N\mu \log(p(1-p))\right] . \end{aligned} \tag{D.55}$$

- Partition function. The partition function for the statistics of single locus characters under SSMD, unfortunately cannot be analitically integrated to give a closed solution. Yet properties and

relations exist that simplify calculations and give understanding of the dynamics.

$$\begin{aligned}\mathbb{Z}_{\ell(N\mu,N\beta,N\sigma)} &:= \mathbb{Z}_\ell \\ &= \int_0^1 dp[p(1-p)]^{-1}\times \\ &\times\ \exp\left[2N\beta(2p-1)+4N\sigma p(1-p)+4N\mu\log[p(1-p)]\right].\end{aligned} \tag{D.56}$$

— *Approximation in series.* For practical purposes, it is sometimes useful to represent the partition function as a series, expanding over σ, which yields the following expression:

$$\begin{aligned}\mathbb{Z}_\ell &\simeq \sqrt{\pi}2^{1-8N\mu}\times \\ &\times\sum_{i=0}^{\infty}\Gamma(4N\mu+i)\ {}_0F_1\left(4N\mu+\frac{1}{2}+i;(N\beta)^2\right)\frac{(16N\sigma)^i}{i!}\end{aligned} \tag{D.57}$$

The series is convergent, and numerically can be computed using Aitken's method (Abramowitz and Stegun, 1972, Sect. 3.9.5, p. 18), since the Hypergeometric function decreases exponentially fast with the expansion index i. However, for weak $N\beta$ between 8 and 10 terms would be enough for a precision of $\sim 10^{-5}$. For strong $N\beta$ between 15 and 20 terms would give that precision. The calculations using the series are usually faster than the numerical or Monte Carlo integration.

Genetic mutation variability

- Definition. As in Eq. D.5.
- Expectancy

$$\begin{aligned}\langle U\rangle_\ell &= \frac{\partial\log(\mathbb{Z}_\ell)}{2N\partial\mu} \\ &= \int_0^1 U\psi_\ell dp\end{aligned} \tag{D.58}$$

- Variance.
- Variance. Calculated from the definition (see Eq. D.60 below)

$$\mathrm{var}(U) = \langle U^2\rangle - \langle U\rangle^2\ . \tag{D.59}$$

Square genetic mutation variability

- Expectancy

$$\langle U^2 \rangle_\ell = \int_0^1 U^2 \psi_\ell dp \tag{D.60}$$

'Genetic variance' of mutational variability

- Definition. As in Eq. D.9.
- Expectancy

$$\langle H \rangle_\ell = -4 + 2\frac{\mathbb{Z}_{\ell(N\mu=N\mu-1/4)}}{\mathbb{Z}_\ell} \tag{D.61}$$

Mean trait

- Definition. As in Eq. D.12
- Expectancy

$$\begin{aligned} \langle \bar{z} \rangle_\ell =& \frac{\partial \log(\mathbb{Z}_\ell)}{2N\partial\beta} \\ =& \int_0^1 \bar{z}\psi_\ell dp \end{aligned} \tag{D.62}$$

- Variance. Calculated from the definition (see Eq. D.65 below)

$$\text{var}(\bar{z}) = \langle \bar{z}^2 \rangle - \langle \bar{z} \rangle^2 \ . \tag{D.63}$$

Squared mean trait

- Definition

$$\begin{aligned} \bar{z}^2 =& (2p-1)^2 \\ =& 1 - 2\nu \end{aligned} \tag{D.64}$$

- Expectancy

$$\langle \bar{z}^2 \rangle_\ell = 1 - 2\langle \nu \rangle_\ell \tag{D.65}$$

Genetic variance

- Definition. As in Eq. D.18.
- Expectancy

$$\langle \nu \rangle_\ell = 2\frac{\mathbb{Z}_{\ell(N\mu=N\mu+1/4)}}{\mathbb{Z}_\ell} \tag{D.66}$$

Third moment of allele frequencies

- Definition

$$m_{3z} \equiv \bar{z}\nu 2(2p-1)(1-p)p \tag{D.67}$$

- Expectancy

$$\begin{aligned} \langle m_{3z} \rangle \equiv & \langle \bar{z}\nu \rangle_\ell = \langle \bar{z} \rangle_{\ell(N\mu=N\mu+1/4)} \langle \nu \rangle_\ell \\ = & \frac{N\beta}{N\sigma}\langle \nu \rangle_\ell - \frac{2N\mu}{N\sigma}\langle \bar{z} \rangle_\ell \end{aligned} \tag{D.68}$$

Fourth moment of allele frequencies

- Definition

$$m_{4z} \equiv \bar{z}^2\nu = 2(2p-1)^2(1-p)p \tag{D.69}$$

- Expectancy

$$\langle m_{4z} \rangle \equiv \langle \bar{z}^2\nu \rangle_\ell = \langle \bar{z}^2 \rangle_{\ell(N\mu=N\mu+1/4)} \langle \nu \rangle_\ell \tag{D.70}$$

$$= \left(\frac{(N\beta)^2}{N\sigma} + 4N\mu \right) \langle \nu \rangle_\ell - \frac{2N\mu N\beta}{N\sigma}\langle \bar{z} \rangle_\ell - 2N\mu \tag{D.71}$$

D.2.2 Statistics for polygenic traits with unequal effects

The computation of multi-locus statistics requires some tricks. There are several options. The first, is to compute the multidimensional integrals using Monte Carlo sampling. This is the least desirable, specially when the distribution of allele frequencies has most of the density near the borders (e.g. under disruptive selection, when we select for genetic variance, and/or at low mutation rates).

However, a neat approach is to transform the integrals from the variables $p \to r = p(1-p)$. Then they are much better behaved to deal numerically (the peaks near the borders are not so steep, and the density is spread across the interval $r \in (0, 1/4)$. Then this integral can be partitioned to sum around the borders $p = \{0, 1\}$, and adding the results of the 2^N corners.

Another option, the one taken for now to achieve some level of analytic results, is to integrate the space of n variables ($\vec{p} = \{p_\ell\}_{\ell=1}^n$) constraining that they conform to a given value of the mean trait, and then integrate in the space of traits. In Section 7.2 I already showed how this method goes. I will focus now in the ways to compute the expectances out of it.

Remarks about the notation. In this section I will use the following notation. Quantities $\langle A \rangle_{\gamma_\ell}$ are per-locus statistics, as above, but evaluated with scaled the scaled arguments $N\beta \to \gamma_\ell N\beta$ and $N\sigma \to \gamma_\ell^2 N\sigma$. That is

$$\langle A \rangle_{\gamma_\ell} = \langle A \rangle_{\ell(N\mu, \gamma_\ell N\beta, \gamma_\ell{}^2 N\sigma)} \ .$$

Partition function

$$\mathbb{Z} = \sqrt{\frac{\pi}{|2N\alpha|}} \int_{-\infty}^{\infty} \exp\left[-\frac{(2N\beta - i\omega)^2}{8N\alpha}\right] \prod_{\ell=1}^{n} \mathbb{Z}_{\ell\ (N\beta = -\frac{i\omega}{2N}\gamma_\ell)} d\omega \qquad \text{(D.72)}$$

(This is the same as Eq. 7.10). If we call

$$\mathbb{Z}_{\text{Ep}} = \sqrt{\frac{\pi}{|2N\alpha|}} \exp\left[-\frac{(2N\beta)^2}{8N\alpha}\right] \qquad \text{(D.73)}$$

$$\mathbb{Z}_{\text{Ad}} = \prod_{\ell=1}^{n} \mathbb{Z}_\ell \qquad \text{(D.74)}$$

we notice that:

1. $\mathbb{Z}_{\text{Ep}}$ is a Gaussian function only of α, β and $i\omega$,
2. $\mathbb{Z}_{\text{Ad}}$ is a function only of σ, μ and $i\omega$,
3. $\mathbb{Z}$ is the convolution (in the variable $2N\beta$) of the 'extensive' single loci partition functions (without epistasis) ($\mathbb{Z}_{\text{Ad}}$), and the Gaussian distribution which induces the epistatic coupling ($\mathbb{Z}_{\text{Ep}}$).

Partitioning into the two terms, and expressing $\mathbb{Z}$ as a convolution simplifies the calculations and the notation. The convolution is defined as:

$$[\![F_{1(x)} * F_{2(x)}]\!]_{(x)} := \int_{-\infty}^{\infty} F_{1(\omega)} F_{2(x-\omega)} d\omega \ .$$

Thus we can write the partition function as

$$\mathbb{Z} = [\![\mathbb{Z}_{\text{Ad}} * \mathbb{Z}_{\text{Ep}}]\!]_{(2N\beta)} \ , \tag{D.75}$$

form equivalent to Eq. D.72, but in convolution notation. The convolution brackets are 'permeable' to the derivatives withe respect to any of the parameters. Thus the statistical mechanical methods of calculating the observables from the partition function apply.

Genetic mutational variability

- Definition

$$U = 2 \sum_{\ell=1}^{n} \log[p_\ell (1 - p_\ell)] \tag{D.76}$$

- Expectancy

$$\langle U \rangle_\ell = \frac{\partial \log(\mathbb{Z})}{2N \partial \mu} \tag{D.77}$$

$$= \mathbb{Z}^{-1} \left[\!\!\left[\left(\mathbb{Z}_{\text{Ad}} \sum_{\ell}^{n} \langle U \rangle_{\gamma_\ell} \right) * \mathbb{Z}_{\text{Ep}} \right]\!\!\right]_{(2N\beta)} \tag{D.78}$$

$$\tag{D.79}$$

Squared mutational genetic variability

$$\langle U^2 \rangle = \mathbb{Z}^{-1} \left[\!\!\left[\left(\mathbb{Z}_{\text{Ad}} \sum_{\ell}^{n} \langle U^2 \rangle_{\gamma_\ell} \right) * \mathbb{Z}_{\text{Ep}} \right]\!\!\right]_{(2N\beta)} \tag{D.80}$$

'Genetic variance' of mutational variability

- Definition

$$H = 2 \sum_{\ell=1}^{n} \left([p_\ell (1 - p_\ell)]^{-1} - 2 \right) \tag{D.81}$$

- Expectancy

$$\langle H \rangle = -4n + 2\frac{\mathbb{Z}_{(N\mu=N\mu-1/4)}}{\mathbb{Z}} \tag{D.82}$$

$$= \mathbb{Z}^{-1}\left[\!\!\left[\left(\mathbb{Z}_{\text{Ad}}\sum_{\ell}^{n}\langle H\rangle_{\gamma_\ell}\right) * \mathbb{Z}_{\text{Ep}}\right]\!\!\right]_{(2N\beta)} \tag{D.83}$$

Mean trait

- Definition

$$\bar{z} = \sum_{\ell=1}^{n}\gamma_\ell(2p_\ell - 1) \tag{D.84}$$

- Expectancy

$$\langle z \rangle = \frac{\partial \log(\mathbb{Z})}{2N\partial\beta} \tag{D.85}$$

$$= \mathbb{Z}^{-1}\left[\!\!\left[\mathbb{Z}_{\text{Ad}} * \left(\frac{-2N\beta}{4N\alpha}\mathbb{Z}_{\text{Ep}}\right)\right]\!\!\right]_{(2N\beta)}$$

Squared mean trait

- Expectancy

$$\langle \bar{z}^2 \rangle = \frac{\partial \log(\mathbb{Z})}{2N\partial\alpha} \tag{D.86}$$

$$= \mathbb{Z}^{-1}\left[\!\!\left[\mathbb{Z}_{\text{Ad}} * \left(\mathbb{Z}_{\text{Ep}}\frac{2N\beta - 16N\alpha}{(8N\alpha)^2}\right)\right]\!\!\right]_{(2N\beta)} \tag{D.87}$$

Cubic mean trait

- Expectancy

$$\langle \bar{z}^3 \rangle = \mathbb{Z}^{-1}\left[\!\!\left[\mathbb{Z}_{\text{Ad}} * \left(\frac{8N\alpha + 2N\beta(16N\alpha - 8N\beta)}{(8N\alpha)^3}\mathbb{Z}_{\text{Ep}}\right)\right]\!\!\right]_{(2N\beta)} \tag{D.88}$$

Quartic mean trait

- Expectancy

$$\langle \bar{z}^4 \rangle = \tag{D.89}$$
$$\mathbb{Z}^{-1}\left[\!\!\left[\mathbb{Z}_{\text{Ad}} * \left(\frac{(2N\beta - 16N\alpha)^2 + 2(16N\alpha)^2 - 128N\alpha N\beta}{(8N\alpha)^4}\mathbb{Z}_{\text{Ep}}\right)\right]\!\!\right]_{(2N\beta)}$$

Genetic variance

- Definition

$$\nu_z = 2\sum_{\ell=1}^{n}\gamma_\ell p_\ell(1-p_\ell) \tag{D.90}$$

- Expectancy

$$\langle \nu_z \rangle = 2\frac{\mathbb{Z}_{(N\mu = N\mu + 1/4)}}{\mathbb{Z}} \tag{D.91}$$
$$= \mathbb{Z}^{-1}\left[\!\!\left[\left(\mathbb{Z}_{\text{Ad}}\sum_{\ell}^{n}\gamma_\ell^2\langle\nu_z\rangle_{\gamma_\ell}\right) * \mathbb{Z}_{\text{Ep}}\right]\!\!\right]_{(2N\beta)} \tag{D.92}$$

Squared genetic variance

- Expectancy

$$\langle \nu_z^2 \rangle = \mathbb{Z}^{-1}\left[\!\!\left[\left(\mathbb{Z}_{\text{Ad}}\sum_{\ell}^{n}\gamma_\ell^4\langle\nu_z^2\rangle_{\gamma_\ell}\right) * \mathbb{Z}_{\text{Ep}}\right]\!\!\right]_{(2N\beta)} \tag{D.93}$$

Third moment of allele frequencies

- Definition

$$m_{3z} = \sum_{\ell=1}^{n}\gamma_\ell^3 2(1-2p_\ell)(1-p_\ell)p_\ell \tag{D.94}$$

- Expectancy

$$m_{3z} = -\mathbb{Z}^{-1}\left[\!\!\left[\left(\mathbb{Z}_{\text{Ad}}\sum_{\ell}^{n}\gamma_\ell^3\langle\bar{z}\nu\rangle_{\gamma_\ell}\right) * \mathbb{Z}_{\text{Ep}}\right]\!\!\right]_{(2N\beta)} \tag{D.95}$$

Fourth moment of allele frequencies

- Definition

$$m_{4_z} = 2\sum_{\ell=1}^{n} \gamma_\ell^4 (2p_\ell - 1)^2 (1 - p_\ell) p_\ell \tag{D.96}$$

- Expectancy

$$\langle m_{4_z} \rangle = \mathbb{Z}^{-1} \left[\!\!\left[\left(\mathbb{Z}_{\text{Ad}} \sum_{\ell}^{n} \gamma_\ell^4 \langle \bar{z}^2 \nu \rangle_{\gamma_\ell} \right) * \mathbb{Z}_{\text{Ep}} \right]\!\!\right]_{(2N\beta)} \tag{D.97}$$

Crossed moment: mean trait and mutational genetic variability

$$\langle \bar{z} U \rangle = \mathbb{Z}^{-1} \left[\!\!\left[\left(\mathbb{Z}_{\text{Ad}} \sum_{\ell}^{n} U_{\gamma_\ell} \right) * \left(\frac{-2N\beta}{4N\alpha} \mathbb{Z}_{\text{Ep}} \right) \right]\!\!\right]_{(2N\beta)} \tag{D.98}$$

Crossed moment: squared mean trait and mutational genetic variability

$$\langle \bar{z}^2 U \rangle = \mathbb{Z}^{-1} \left[\!\!\left[\left(\mathbb{Z}_{\text{Ad}} \sum_{\ell}^{n} \langle U \rangle_{\gamma_\ell} \right) * \left(\frac{2N\beta - 16N\alpha}{(8N\alpha)^2} \mathbb{Z}_{\text{Ep}} \right) \right]\!\!\right]_{(2N\beta)} \tag{D.99}$$

Crossed moment: mean trait and genetic variance

$$\langle \bar{z} \nu_z \rangle = \mathbb{Z}^{-1} \left[\!\!\left[\left(\mathbb{Z}_{\text{Ad}} \sum_{\ell}^{n} \gamma_\ell^2 \langle \nu_z \rangle_{\gamma_\ell} \right) * \left(\frac{-2N\beta}{4N\alpha} \mathbb{Z}_{\text{Ep}} \right) \right]\!\!\right]_{(2N\beta)} \tag{D.100}$$

Crossed moment: squared mean trait and genetic variance

$$\langle \bar{z}^2 \nu_z \rangle = \mathbb{Z}^{-1} \left[\!\!\left[\left(\mathbb{Z}_{\text{Ad}} \sum_{\ell}^{n} \gamma_\ell^2 \langle \nu_z \rangle_{\gamma_\ell} \right) * \left(\frac{2N\beta - 16N\alpha}{(8N\alpha)^2} \mathbb{Z}_{\text{Ep}} \right) \right]\!\!\right]_{(2N\beta)} \tag{D.101}$$

Crossed moment: mean trait and third moment of allele frequencies

$$\langle \bar{z} m_{3_z} \rangle = -\mathbb{Z}^{-1} \left[\!\!\left[\left(\mathbb{Z}_{\text{Ad}} \sum_{\ell}^{n} \gamma_\ell^3 \langle \bar{z} \nu \rangle_{\gamma_\ell} \right) * \left(\frac{-2N\beta}{4N\alpha} \mathbb{Z}_{\text{Ep}} \right) \right]\!\!\right]_{(2N\beta)} \tag{D.102}$$

Covariances of A_j A_k (not including the traits)

$$\mathrm{cov}(A_j, A_k) = \mathbb{Z}^{-1}\left[\!\!\left[\left(\mathbb{Z}_{\mathrm{Ad}}\sum_{\ell}^{n}\gamma_\ell^2\langle A_j\rangle_{\gamma\ell}\langle A_k\rangle_{\gamma\ell} + \mathrm{cov}(A_j, A_k)_{\gamma\ell}\right) * \mathbb{Z}_{\mathrm{Ep}}\right]\!\!\right]_{(2N\beta)} - \langle A_j\rangle\langle A_k\rangle \tag{D.103}$$

Variances of A_j (not including powers of the traits)

From the previous identity, it follows that

$$\mathrm{var}(A_j) = \mathbb{Z}^{-1}\left[\!\!\left[\left(\mathbb{Z}_{\mathrm{Ad}}\sum_{\ell}^{n}\gamma_\ell^2\langle A_j\rangle_{\gamma\ell}^2 + \mathrm{var}(A_j)_{\gamma\ell}\right) * \mathbb{Z}_{\mathrm{Ep}}\right]\!\!\right]_{(2N\beta)} - \langle A_j\rangle^2 \,. \tag{D.104}$$