

Springer Theses

Recognizing Outstanding Ph.D. Research

Aims and Scope

The series “Springer Theses” brings together a selection of the very best Ph.D. theses from around the world and across the physical sciences. Nominated and endorsed by two recognized specialists, each published volume has been selected for its scientific excellence and the high impact of its contents for the pertinent field of research. For greater accessibility to non-specialists, the published versions include an extended introduction, as well as a foreword by the student’s supervisor explaining the special relevance of the work for the field. As a whole, the series will provide a valuable resource both for newcomers to the research fields described, and for other scientists seeking detailed background information on special questions. Finally, it provides an accredited documentation of the valuable contributions made by today’s younger generation of scientists.

Theses are accepted into the series by invited nomination only and must fulfill all of the following criteria

- They must be written in good English.
- The topic should fall within the confines of Chemistry, Physics, Earth Sciences, Engineering and related interdisciplinary fields such as Materials, Nanoscience, Chemical Engineering, Complex Systems and Biophysics.
- The work reported in the thesis must represent a significant scientific advance.
- If the thesis includes previously published material, permission to reproduce this must be gained from the respective copyright holder.
- They must have been examined and passed during the 12 months prior to nomination.
- Each thesis should include a foreword by the supervisor outlining the significance of its content.
- The theses should have a clearly defined structure including an introduction accessible to scientists not expert in that particular field.

More information about this series at <http://www.springer.com/series/8790>

George William Albert Constable

Fast Variables in Stochastic Population Dynamics

Doctoral Thesis accepted by
the University of Manchester, UK

Author

Dr. George William Albert Constable
Department of Ecology
and Evolutionary Biology
Princeton University
Princeton, NJ
USA

Supervisor

Prof. Alan McKane
School of Physics and Astronomy
The University of Manchester
Manchester
UK

ISSN 2190-5053

Springer Theses

ISBN 978-3-319-21217-3

DOI 10.1007/978-3-319-21218-0

ISSN 2190-5061 (electronic)

ISBN 978-3-319-21218-0 (eBook)

Library of Congress Control Number: 2015943440

Springer Cham Heidelberg New York Dordrecht London

© Springer International Publishing Switzerland 2015

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media
(www.springer.com)

Parts of this thesis have been published in the following journal articles:

G. W. A. Constable, A. J. McKane and T. Rogers

Stochastic dynamics on slow manifolds

J. Phys. A: Math. Theor. Vol 46, 295002 (2013)

G. W. A. Constable and A. J. McKane

Fast-mode elimination in stochastic metapopulation models

Phys. Rev. E Vol 89, 032141 (2014)

G. W. A. Constable and A. J. McKane

Population genetics on islands connected by an arbitrary network: An analytic approach

J. Theor. Biol. Vol 358, 149–165 (2014)

G. W. A. Constable and A. J. McKane

Models of genetic drift as limiting forms of the Lotka-Volterra competition model

Phys. Rev. Lett. Vol 114, 038101 (2015)

G. W. A. Constable and A. J. McKane

Stationary solutions for metapopulation Moran models with mutation and selection

Phys. Rev. E Vol 91, 032711 (2015)

Supervisor's Foreword

It is with the greatest of pleasure that I write the foreword to George Constable's Ph.D. thesis. At one level the thesis is a very clear description of a method of analysing models in population genetics, which I will outline in more detail below. But on a more general level it should also be very useful as an introduction to those wishing to understand the formalism of continuous-time stochastic processes. There are not so many textbooks on this subject, and those which do exist can be technically quite forbidding. This need not be so—the essential ideas and techniques are relatively straightforward to understand—and the opening chapters of this thesis provide an accessible introduction to them.

The main subject of the thesis is the stochastic time-evolution of populations when the individuals making up the population have a very simple genetic make-up: they are haploid and the focus is on a single gene which has only two alleles. The ideas can certainly be extended to more complex systems, but the aim was to develop techniques to allow models to be analysed, and testing them on the simplest situations to start with. Attention was also directed to aspects of the models which have perhaps received less attention than they deserve.

The first of these aspects is the careful specification of the model and its subsequent simplification through the use of a diffusion approximation. Theoretical physicists tend to take great care in distinguishing between microscopic, mesoscopic and macroscopic descriptions of the same system, and the specification of the approximations that are made to go from one level of description to another. There is less of a tradition of doing this in the context of biological systems. In this thesis there is a careful separation of the modelling and approximation processes, so that both the starting point and the nature of the approximations subsequently made are absolutely clear.

The second, and more substantial, contribution described in the thesis is in the development of an additional approximation which makes an intractable equation amenable to analysis. The standard method of specifying a model after the diffusion approximation has been made goes back to the work of Fisher, but was popularised by Kimura in the 1950s. It takes the form of a partial differential equation known as

a Fokker-Planck equation to physicists and as a Kolmogorov forward equation to those in many other disciplines. In simple situations it can be analysed but if, for instance, the population is subdivided into islands (or demes, as they are sometimes called), then this is a partial differential equation in many variables, and as such any in-depth analysis appears to be almost impossible.

As a consequence in some areas of population genetics, the general equations governing the dynamics have not been addressed directly, because of their intractability, and the focus has moved to simpler systems or to those where progress could be made. This is understandable, but what the work in this thesis shows is that this is not necessary; a simple approximation is available which reduces a many-island description to one which is effectively a one-island model, but with effective parameters. The resulting Fokker-Planck equation can now be used to calculate the probability that a particular allele becomes fixed, how long this will take on average or what the nature of the stationary probability distribution of alleles is. There had been a few previous attempts to do this, but the procedure outlined here is both more general and easier to understand than previous studies.

The procedure is based on a fast-mode elimination technique. The idea is very simple and long established in the theory of dynamical systems. Essentially, the variables in the model are decomposed into a set which decay at different rates, $|\lambda^{(i)}|^{-1}$, $i = 1, 2, \dots$. It turns out that in many cases of interest there is a 'gap' between the smallest $|\lambda^{(i)}|$ (taken to be $|\lambda^{(1)}|$) and all the others. This means that after quite a short time, compared to the timescales of interest to us, only the mode characterised by $|\lambda^{(1)}|$ is left in the model describing the system. This is the 'slow' mode—all the other 'fast' modes have decayed away and dropped out of the theoretical description. What is left is a model with just one degree of freedom, which can then be analysed systematically.

Of course, although the idea just described seems simple enough, finding a concrete procedure which works, and which can be turned into a calculational tool, is not. However, remarkably, a method was found which is both rather straightforward to apply and which also gives results in excellent agreement with computer simulations of the original individual-based (that is, microscopic) model. It results in a 'reduced' model in which the parameters are given in terms of the island sizes, the scale of migration between the islands, or whatever parameters were present in the original (full) model. Although I have used the example of subdivided populations to illustrate the method, it should be more generally applicable to the reduction of complex population genetics models down to much simpler ones with just a few effective parameters which are explicitly given in terms of those of the full model.

I have already mentioned that those looking for an easy-to-understand introduction to the formalism of continuous-time stochastic processes would benefit from reading this thesis. But I would also hope that it would appeal to theoretical biologists seeking to extend the scope of problems it can be applied to, to mathematicians wanting to make the approach more rigorous, and to theoretical physicists looking for an application of the ideas and techniques of non-equilibrium

statistical mechanics. So my hope is that the publication of this thesis will allow a much wider range of people to appreciate the power of the methodology presented here, and also enable them to contribute to extending its range of applicability.

Manchester
April 2015

Alan McKane

Abstract

In this thesis, I present two methods of fast variable elimination in stochastic systems. Their application to models of population dynamics from ecology, epidemiology and population genetics, is explored. In each application, care is taken to develop the models at the microscale, in terms of interactions between individuals. Such an approach leads to well-defined stochastic systems for finite population sizes. These systems are then approximated at the mesoscale, and expressed as stochastic differential equations. It is in this setting that elimination techniques are developed.

In each model a deterministically stable state is assumed to exist, about which the system is linearised. The eigenvalues of the system's Jacobian are used to identify the existence of a separation of timescales. The fast and slow directions are then given locally by the associated eigenvectors. These are used as approximations for the fast and slow directions in the full nonlinear system. The general aim is then to remove these fast degrees of freedom and thus arrive at an approximate, reduced-variable description of the dynamics on a slow subspace of the full system.

In the first of the methods introduced, the conditioning method, the noise of the system is constrained so that it cannot leave the slow subspace. The technique is applied to an ecological model and a susceptible-exposed-infectious-recovered epidemiological model, in both instances providing a reduced system which preserves the behaviour of the full model to high precision.

The second method is referred to as the projection matrix method. It isolates the components of the noise on the slow subspace to provide its reduced description. The method is applied to a generalised Moran model of population genetics on islands, between which there is migration. The model is successfully reduced from a system in as many variables as there are islands, to an effective description in a single variable. The same methodology is later applied to the Lotka-Volterra competition model, which is found under certain conditions to behave as a Moran model. In both cases the agreement between the reduced system and stochastic simulations of the full model is excellent.

It is emphasised that the ideas behind both the conditioning and projection matrix methods are simple, their application systematic, and the results in very good agreement with simulations for a range of parameter values. When the methods are compared, however, the projection matrix method is found in general to provide better results.

Acknowledgments

Writing this thesis has been a surprisingly enjoyable experience. It would not have been so without the help and support of those around me.

I would like to offer my most sincere gratitude to my supervisor, Alan McKane, whose advice on all matters has been invaluable, and whose company I have thoroughly enjoyed.

I thank Tim Rogers, who by virtue of infectious enthusiasm, tempered by bouts of cynicism, is a great guy to work with.

Over the past three years I have been lucky enough to be surrounded by people who are happy to take time out of their day to talk about ideas and clarify thoughts. In this regard I am grateful to Tobias, Diana, Louise and Luis, as well as the many who have shared Office 7.26—Alex, Joe, Tommaso, James, Toby, Peter and Cesar.

The funding for this Ph.D. has been provided by a University of Manchester Engineering and Physical Sciences scholarship, and so I thank the EPS for their generosity.

Finally, I am indebted to my family, which have been unfailingly supportive—David and Julie, Deborah and June.

Contents

1	Introduction	1
	References.	10
2	Technical Background	11
2.1	Stochastic Processes and the Markov Assumption	12
2.2	The Master Equation and the Limit of Continuous Time	13
2.3	The Gillespie Algorithm.	16
2.4	The Expansion of the Master Equation	18
2.5	The Fokker-Planck Equation and Some Useful Manipulations.	22
	2.5.1 The Stationary Probability Distribution	23
	2.5.2 First Passage Problems	25
2.6	Stochastic Differential Equations.	28
2.7	Dynamical Systems Theory	30
2.8	Fast-Variable Elimination: The Origins of a Complicated Problem	33
2.9	The Moran Model.	37
	2.9.1 The Moran Model with Selection	38
2.10	Technical Background Overview.	41
	References.	42
3	The Conditioning Method	45
3.1	Identifying the Deterministic Slow Subspace	46
3.2	Illustrative Example.	48
3.3	General Formulation	53
3.4	Application: Seasonally Forced Epidemics	55
	3.4.1 Model Definition and Deterministic Treatment	55
	3.4.2 Stochastic Treatment Exploiting the Slow Manifold	59
3.5	Discussion	62
	References.	64

4	The Projection Matrix Method	65
4.1	Introduction to the Projection Matrix Method	65
4.2	The Moran Model with Migration	67
4.3	The Metapopulation Moran Model	69
4.3.1	The Metapopulation Moran Model with Selection	72
4.4	The Effective Metapopulation Moran Model	73
4.4.1	The Case with Selection	78
	References.	81
5	Analysis of the Reduced Metapopulation Moran Model	83
5.1	Analysis—Neutral Case	83
5.2	Analysis—Case with Selection	87
5.2.1	First Order in s	88
5.2.2	Second Order in s	88
5.3	Estimating the Range of Validity of the Method	91
5.4	Migration-Selection Balance	93
5.5	Hub.	96
5.5.1	Hub: $s = 0$	99
5.5.2	Hub with Selection	100
	References.	101
6	Further Developments.	103
6.1	Mutation and the Metapopulation Moran Model	103
6.2	Comparing the Conditioning and the Projection Matrix Method	109
6.2.1	Applying the Methods to a Linear System with Additive Noise	111
6.2.2	Application to Neutral Two-Deme Metapopulation Moran Model	114
6.3	Reducing the Lotka-Volterra Model.	116
6.3.1	The Neutral Case	118
6.3.2	The Non-neutral Case	121
	References.	125
7	Conclusion	127
	References.	131
	Appendix A: The Conditional and Marginal of a Multivariate Gaussian Distribution	133
	Appendix B: Floquet Theory	139
	Appendix C: Derivation of the Fokker-Planck Equation for the Metapopulation Moran Model	143

Contents	xvii
Appendix D: Specification of Parameters Used in Figures	147
Appendix E: Moran Model with Selection: Fixation Time	151
Appendix F: Calculation of the Metapopulation Moran Model Dynamics on the Slow Subspace	155
Appendix G: The Probability of Fixation in the Metapopulation Moran Model with Selection	159
About the Author	163