

Xiaoying Han
Peter E. Kloeden

Random Ordinary Differential Equations and Their Numerical Solution

Probability Theory and Stochastic Modelling

Volume 85

Editors-in-chief

Søren Asmussen, Aarhus, Denmark
Peter W. Glynn, Stanford, CA, USA
Yves Le Jan, Orsay, France

Advisory Board

Martin Hairer, Coventry, UK
Peter Jagers, Gothenburg, Sweden
Ioannis Karatzas, New York, NY, USA
Frank P. Kelly, Cambridge, UK
Andreas Kyprianou, Bath, UK
Bernt Øksendal, Oslo, Norway
George Papanicolaou, Stanford, CA, USA
Etienne Pardoux, Marseille, France
Edwin Perkins, Vancouver, BC, Canada
Halil Mete Soner, Zürich, Switzerland

The **Probability Theory and Stochastic Modelling** series is a merger and continuation of Springer's two well established series Stochastic Modelling and Applied Probability and Probability and Its Applications series. It publishes research monographs that make a significant contribution to probability theory or an applications domain in which advanced probability methods are fundamental. Books in this series are expected to follow rigorous mathematical standards, while also displaying the expository quality necessary to make them useful and accessible to advanced students as well as researchers. The series covers all aspects of modern probability theory including

- Gaussian processes
- Markov processes
- Random fields, point processes and random sets
- Random matrices
- Statistical mechanics and random media
- Stochastic analysis

as well as applications that include (but are not restricted to):

- Branching processes and other models of population growth
- Communications and processing networks
- Computational methods in probability and stochastic processes, including simulation
- Genetics and other stochastic models in biology and the life sciences
- Information theory, signal processing, and image synthesis
- Mathematical economics and finance
- Statistical methods (e.g. empirical processes, MCMC)
- Statistics for stochastic processes
- Stochastic control
- Stochastic models in operations research and stochastic optimization
- Stochastic models in the physical sciences

More information about this series at <http://www.springer.com/series/13205>

Xiaoying Han · Peter E. Kloeden

Random Ordinary Differential Equations and Their Numerical Solution



Springer

Xiaoying Han
Auburn, AL
USA

Peter E. Kloeden
School of Mathematics and Statistics
Huazhong University of Science
and Technology
Wuhan, Hubei
China

ISSN 2199-3130

ISSN 2199-3149 (electronic)

Probability Theory and Stochastic Modelling

ISBN 978-981-10-6264-3

ISBN 978-981-10-6265-0 (eBook)

DOI 10.1007/978-981-10-6265-0

Library of Congress Control Number: 2017951158

Mathematics Subject Classification (2010): 37H10, 60H10, 60H35, 34F05, 37H70, 60H30, 65LC30, 65L05, 65L06, 65L20, 92-08

© Springer Nature Singapore Pte Ltd. 2017

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. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature

The registered company is Springer Nature Singapore Pte Ltd.

The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

For my parents

—Xiaoying Han

Für Karin

—Peter E. Kloeden

Preface

Random ordinary differential equations (RODEs) are ordinary differential equations (ODEs) that include a stochastic process in their vector field. They seem to have had a shadow existence to Itô stochastic differential equations (SODEs), but have been around for as long as if not longer and have many important applications. In particular, RODEs play a fundamental role in the theory of random dynamical systems.

In the older engineering and physics literature, a simpler kind of RODE is investigated with the vector field being chosen randomly rather than depending on a stochastic process. Such RODEs are still of interest in the uncertainty quantification community. They are a special case of those considered in this book, but will not be treated separately.

RODEs, unlike SODEs, can be analysed pathwise with deterministic calculus, but require further treatment beyond that of classical ODE theory. Specifically, since the driving process in a RODE has at most Hölder continuous sample paths, the solution sample paths are continuously differentiable, but the sample paths of the derivative are no more than Hölder continuous in time. The resulting vector field after insertion of the driving stochastic process is at most Hölder continuous in time, no matter how smooth the vector field is in its original variables. Thus, the solutions of RODEs do not have sufficient smoothness to have Taylor expansions in the usual sense.

Taylor expansions are a very basic tool in numerical analysis. They allow one to derive one-step numerical schemes for ODEs of arbitrary high order. In practice, such Taylor schemes are rarely implemented, but are used instead as a theoretical comparison for determining the convergence orders of other schemes that have been derived by more heuristic methods. This theory does not apply directly to RODEs or SODEs. Although classical numerical schemes for ODEs can be used pathwise for RODEs, they rarely attain their traditional order.

The situation with RODEs is similar to that for Itô SODEs, the sample paths of which are just Hölder continuous and not even differentiable. For Itô SODEs, stochastic Taylor expansions can be derived using an iterated application of the Itô formula, the stochastic chain rule, in integral form. These stochastic Taylor

expansions were the starting point for the derivation of consistent higher order numerical schemes for SODEs. A similar approach will be used in this book for RODEs.

A major motivation for writing this book is to make more widely known recent results on the derivation of higher order numerical schemes for RODEs. Another is to make RODEs themselves and the closely associated theory of random dynamical systems better known too. A particular motivation of personal interest to us is to show how RODEs are being used in the biological sciences, where non-Gaussian and bounded noise are often more realistic than the conventionally used Itô calculus.

Auburn, USA
Wuhan, China
March 2017

Xiaoying Han
Peter E. Kloeden

Acknowledgements

Both authors were substantially supported by the Thousand Talent Program for Foreign Experts of the Chinese Central Government and the Huazhong University of Science and Technology.

We thank Pamela Burrage, Denise Chang, Jiayin Jin, Thomas Lorenz, Andreas Neuenkirch, Yuanling Niu, Andreas Rößler, Florian Rupp, Peng Wang, Xiaojie Wang and Fuke Wu for carefully reading parts of the manuscript and for their useful feedback. In addition, we thank Fritz Colonius, Lars Grüne, Thomas Lorenz and Thomas Wanner for their assistance with the literature. The valuable insights and suggestions of the anonymous reviewers of an early manuscript have been very helpful to us and are deeply appreciated.

We are particularly grateful to Yusuke Asai, who was working towards his doctorate on the numerical analysis of RODEs at the Goethe University in Frankfurt am Main, as we were writing this book. He read various earlier versions of the manuscript with great care and provided us with many of the figures and simulation results as well as very useful comments.

Guide to Reading This Book

We suggest that all readers first read Chap. 1 where we introduce random ordinary differential equations (RODEs) and briefly indicate various numerical methods for them that have appeared in the literature in the past.

Readers interested in their applications, in particular in a biological context, can obtain an overview by reading the first pages of each chapter in Part IV at any time. The later parts of these chapters apply the numerical methods and various results from the theory of random dynamical systems that are developed in the book.

Those readers who are mainly interested in numerical methods could skip Chap. 4 on random dynamical systems and Chap. 5 on numerical dynamics or return to them later when the material in them is needed, e.g. as in some of the applications in Part IV.

The numerical schemes for RODEs in this book are derived systematically using various kinds of generalised Taylor expansions in integral form, depending on the structure of the RODEs or the nature of the noise in them. These are introduced and investigated in Part II. This is essential reading for those who plan to contribute to further mathematical developments of the subject. Numerical schemes based on them are presented in Part III, with the notation and Taylor approximations from Part II briefly restated for the readers' convenience.

It is not essential to read the chapters in these parts in a linear fashion. A reader who is interested only in a particular type of RODE or driving noise could just read the corresponding chapters in the Parts II and III. Nevertheless, we recommend that readers skim first over all chapters, here and elsewhere, without focusing on technical details, to obtain an overview of the contents and to see how the different aspects are interrelated.

Contents

Part I Random and Stochastic Ordinary Differential Equations

1	Introduction	3
1.1	Simple Numerical Schemes for RODEs	4
1.2	Taylor Expansions for RODEs	6
1.2.1	ODE Case	6
1.2.2	SODE Case	7
1.2.3	RODE Case	8
1.3	RODEs with Bounded Noise	11
1.4	RODEs and Carathéodory ODEs	12
1.5	Endnotes	13
2	Random Ordinary Differential Equations	15
2.1	Existence and Uniqueness Theorems	16
2.1.1	Classical Assumptions	16
2.1.2	Measurability of Solutions	20
2.1.3	Carathéodory Assumptions	21
2.1.4	Positivity of Solutions	24
2.2	RODEs with Canonical Noise	25
2.3	Endnotes	26
3	Stochastic Differential Equations	29
3.1	Wiener Processes and Itô Integrals	30
3.2	Itô Stochastic Differential Equations	31
3.3	The Itô Formula: The Stochastic Chain Rule	33
3.4	Stratonovich SODEs	34
3.5	Relationship Between RODEs and SODEs	35
3.5.1	Doss–Sussmann Transformation	35
3.6	Endnotes	36

4 Random Dynamical Systems	37
4.1 Nontrivial Equilibrium Solutions	37
4.2 Random Dynamical Systems	42
4.2.1 Random Attractors.	43
4.2.2 Contractive Cocycles.	46
4.3 Endnotes	47
5 Numerical Dynamics	49
5.1 Discretisation of Random Attractors	49
5.2 Discretisation of a Random Hyperbolic Point	53
5.3 Endnotes	58
 Part II Taylor Expansions	
6 Taylor Expansions for Ordinary and Stochastic Differential Equations	61
6.1 Taylor Approximations for ODEs.	61
6.2 Taylor Approximations of Itô SODEs.	62
6.2.1 Multi-indices	64
6.2.2 Multiple Integrals of Stochastic Processes	64
6.2.3 Coefficient Functions.	65
6.2.4 Hierarchical and Remainder Sets.	66
6.3 General Itô–Taylor Expansions.	66
6.4 Strong Itô–Taylor Approximations	69
6.4.1 Examples.	70
6.4.2 Pathwise Convergence.	70
6.5 Endnotes.	72
7 Taylor Expansions for RODEs with Affine Noise	73
7.1 An Illustrative Example	74
7.2 Affine-RODE Taylor Expansions	76
7.3 General Affine-RODE-Taylor Approximations	77
7.4 Endnotes.	79
8 Taylor-Like Expansions for General Random Ordinary Differential Equations	81
8.1 Preliminaries and Notation	81
8.1.1 Regularity of the Driving Stochastic Process.	82
8.1.2 Multi-index Notation.	83
8.1.3 Iterated Integrals	84
8.1.4 Function Spaces.	84
8.1.5 Iterated Differential Operators	85
8.2 Integral Equation Expansions	86
8.3 RODE-Taylor Approximations	88
8.4 Essential RODE-Taylor Approximations.	89

8.5	Examples	91
8.6	Proof of Theorem 8.1	94
8.7	Endnotes	98
Part III Numerical Schemes for Random Ordinary Differential Equations		
9	Numerical Methods for Ordinary and Stochastic Differential Equations	101
9.1	One-Step Numerical Schemes for ODEs	101
9.2	One-Step Numerical Schemes for Itô SODEs	104
9.3	Strong Taylor Schemes for Itô SODEs	105
9.4	Endnotes	108
10	Itô–Taylor Schemes for RODEs with Itô Noise	109
10.1	One-Step Schemes	110
10.1.1	Scalar Case	110
10.1.2	Vector Case	111
10.1.3	Examples	113
10.1.4	Derivative-Free Explicit Strong Schemes	114
10.2	Implicit Strong Schemes	116
10.3	Multi-step Schemes	117
10.4	RODEs with Affine Noise	124
10.5	Proof of Theorem 10.1	125
10.6	Endnotes	127
11	Numerical Schemes for RODEs with Affine Noise	129
11.1	Affine-RODE Taylor Schemes for Bounded Noise	130
11.2	Affine-RODEs with Special Structure	133
11.2.1	Additive Noise	133
11.2.2	Commutative Noise	133
11.3	Affine-RODE Derivative-Free Schemes	134
11.3.1	Finite Difference Approximation of Derivatives	134
11.3.2	Runge–Kutta Schemes for Affine-RODE	136
11.4	Linear Multi-step Methods for Affine RODEs	139
11.5	Endnotes	142
12	RODE-Taylor Schemes: General Case	143
12.1	RODE-Taylor Schemes	143
12.1.1	The Essential RODE-Taylor Schemes	146
12.2	Examples of the RODE-Taylor Schemes	147
12.3	RODEs with Affine Noise	149
12.4	Other Numerical Schemes for RODEs	150
12.4.1	The Local Linearisation Scheme for RODEs	150
12.4.2	The Averaged Euler Scheme	152

12.4.3	Heuristic RODE-Taylor Schemes	152
12.5	Endnotes	153
13	Numerical Stability	155
13.1	B-Stability of the Implicit Averaged Schemes	156
13.2	B-Stability of the Implicit Multi-step Schemes	158
13.3	Endnotes	161
14	Stochastic Integrals: Simulation and Approximation	163
14.1	Calculating a Finer Approximation of the Same Sample Path	164
14.2	Integral of a Wiener Process	165
14.3	Integral of an Ornstein–Uhlenbeck Process	165
14.4	Fractional Brownian Motion	167
14.4.1	Riemann Integral of an fBm	172
14.4.2	Riemann Sums Approximation	173
14.4.3	Comparison of Computational Costs	174
14.5	Integrals of Compound Poisson Processes	177
14.6	Endnotes	179

Part IV Random Ordinary Differential Equations in the Life Sciences

15	Comparative Simulations of Biological Systems	183
15.1	Tumor Inhibition Model	183
15.2	Population Dynamics	185
15.3	Toggle Switch Model	187
15.4	Sea Shell Pattern Model	189
15.5	Endnotes	192
16	Chemostat	193
16.1	Random Chemostat Models	194
16.2	RDS Generated by Random Chemostat	196
16.3	Existence of a Random Attractor	198
16.4	Endnotes	203
17	Immune System Virus Model	205
17.1	Properties of Solutions	206
17.2	Existence of Global Random Attractors	208
17.3	Numerical Simulations	212
17.4	Endnotes	214
18	Random Markov Chains	215
18.1	Random Environment	216
18.2	Positivity of Solutions of Linear RODEs	217
18.3	Linear Random Dynamical Systems	220

Contents	xvii
18.4 Random Attractor Under Discretisation	221
18.5 Endnotes	222
Appendix A: Probability Spaces	223
Appendix B: Chain Rule for Affine RODEs	227
Appendix C: Covariance Matrix of a Fractional Brownian Motion and Its Integral	231
References	241
Index	247

Part I

**Random and Stochastic Ordinary
Differential Equations**

Chapter 1

Introduction

The inclusion of noise terms in differential equations dates back to the early 1900s when Einstein provided an explanation for the physical phenomenon of Brownian motion and modeled it as a diffusion [45]. Langevin was interested in the actual motion of the Brownian particles and formulated noisy differential equations of the form [97]

$$\frac{dx}{dt} = f(x) + \sigma(x)\eta_t, \quad (1.1)$$

where η_t is a random force due to random density fluctuations in the fluid. More precisely η_t is a stochastic process giving the effect of background noise due to the fluid on the Brownian particle, with its first and second moments satisfying

$$\mathbb{E}[\eta_t] = 0, \quad \mathbb{E}[\eta_{t_1}\eta_{t_2}] = \rho\delta(t_1 - t_2), \quad \forall t, t_1, t_2 \in [0, \infty),$$

where ρ is the measure of the strength of the fluctuation force and δ is the Dirac delta function.

Many technical difficulties arose, for mathematicians at least, since the noise process η_t was meant to be Gaussian white noise and it is not at all straightforward to obtain the existence of a unique solution to the differential Eq. (1.1), nor even the existence of dx/dt . If η_t is continuous, standard existence theorems for ordinary differential equations guarantee the existence of a local solution to (1.1) that exists in some neighborhood of the point at which the initial value is given. Even if a solution exists, it may be only local, or it may not be unique, unless some stronger conditions are imposed on the coefficient functions.

It took a good half century before K. Itô [73] could develop a stochastic calculus which allowed a rigorous formulation and mathematical development of a theory of

stochastic differential equations (SODEs¹). Itô stochastic calculus is a mean-square or L_2 calculus with different transformation rules to deterministic calculus. This has some critical implications for the development of effective numerical schemes for SODEs.

In many physical applications the noise often has a wide band rather than white spectrum, i.e., it is an h -correlated stationary Gaussian process $\eta_t^{(h)}$ with a Gaussian white noise limit as $h \rightarrow 0$. In this case the noisy differential Eq. (1.1) is in fact an ordinary differential equation (ODE)

$$\frac{dx}{dt} = f(x) + \sigma(x) \eta_t^{(h)},$$

which can be handled pathwise by the methods of deterministic calculus. The kind of SODE that might arise in the limit as $h \rightarrow 0$ has been investigated intensively. For example, Wong and Zakai [140] obtained the Stratonovich SODE

$$\frac{dx}{dt} = f(x) + \sigma(x) \circ dW_t$$

under certain conditions.

More generally, a *random ordinary differential equation* (RODE) is formulated pathwise as an ODE

$$\frac{dx}{dt} = g(x, \eta_t), \quad (1.2)$$

where η_t is a stochastic process with Hölder continuous sample paths or even a Poisson process with piecewise continuous sample paths. In some of the older engineering literature, e.g., Soong [124], Srinivasan and Vasudevan [126], a simpler kind of RODE was investigated with the vector field being chosen randomly, i.e., of the form $g(x, \omega)$, rather than depending on a stochastic process. Such RODEs are special case of those considered in this book.

RODEs seem to have had a shadow existence to SODEs, but have been around for as long as if not longer than SODEs themselves and have many important applications, see e.g., Bunke [20], Neckel and Rupp [108], Sobczyk [123]. They also play a fundamental role in the theory of random dynamical systems, see Arnold [4].

1.1 Simple Numerical Schemes for RODEs

The rules of deterministic calculus apply pathwise to RODEs, but the vector field function in $G_\omega(t, x) := g(x, \eta_t(\omega))$ of the RODE (1.2) for a fixed sample path (labeled by ω) is not smooth in t . It is at most Hölder continuous in time when the

¹The “O” here stands for “ordinary” and is often included to distinguish them from stochastic “partial” differential equations (SPDEs). It has nothing to do with deterministic ordinary differential equations.

driving stochastic process η_t is Hölder continuous and thus lacks sufficient smoothness needed to justify the Taylor expansions and the error analysis of traditional numerical methods for ODEs. Such methods can be used but will attain at best a low convergence order, so new higher order numerical schemes must be derived for RODEs.

For example, let η_t be pathwise Hölder continuous of order $\frac{1}{2}$. Then the Euler scheme with step size h_n

$$x_{n+1} = (1 - h_n) x_n + \eta_{t_n} h_n$$

for the RODE

$$\frac{dx}{dt} = -x + \eta_t,$$

attains the pathwise order $\frac{1}{2}$.

An Averaged Euler Scheme

One can do better, however, by using the pathwise *averaged Euler scheme*

$$x_{n+1} = (1 - h_n) x_n + \int_{t_n}^{t_{n+1}} \eta_s ds,$$

which was proposed by Grüne and Kloeden [55]. It attains the pathwise order 1 provided the integral is approximated with Riemann sums

$$\int_{t_n}^{t_{n+1}} \eta_s ds \approx \sum_{j=1}^{N_{h_n}} \eta_{t_n + j \Delta t_n} \Delta t_n$$

with the step size Δt_n satisfying $\sqrt{\Delta t_n} \approx h_n$ and $\Delta t_n \cdot N_{h_n} = h_n = t_{n+1} - t_n$. In fact, this was done more generally in [55] for RODEs with an affine structure, i.e., of the form

$$\frac{dx}{dt} = f(x) + \sigma(x)\eta_t,$$

where $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^d \times \mathbb{R}^m$ and η_t is m -dimensional.

The explicit averaged Euler scheme then reads

$$x_{n+1} = x_n + [f(x_n) + \sigma(x_n) I_n] h_n,$$

where

$$I_n := \frac{1}{h_n} \int_{t_n}^{t_{n+1}} \eta_s ds.$$

For the general RODE (1.2) this suggests that one should pathwise average the vector field, i.e.,

$$\frac{1}{h_n} \int_{t_n}^{t_{n+1}} g(x_n, \eta_s(\omega)) \, ds,$$

which is computationally expensive even for low dimensional systems. An alternative is to use the averaged noise within the vector field, which leads to the explicit *averaged Euler scheme*

$$x_{n+1} = x_n + g(x_n, I_n) h_n.$$

A Local Linearisation Method

Carbonell, Jimenez, Biscay and de la Cruz [30] introduced a local linearisation method for RODEs of the form

$$\frac{dx}{dt} = g(x, \eta_t),$$

in \mathbb{R}^d , where $\eta_t = (\eta_t^1, \dots, \eta_t^m)$ is an m -dimensional noise. This has the form

$$x_{n+1} = x_n + \int_0^h e^{g_x(h-s)} g \, ds + \sum_{j=1}^m \int_0^h e^{g_x(h-s)} g_{\eta_t^j} \Delta \eta_{t_n}^j \, ds,$$

where $\Delta \eta_{t_n}^j = \eta_{t_n+s}^j - \eta_{t_n}^j$, and the $d \times d$ matrix-valued function $g_x = \partial_x g$ and the d dimensional vector-valued functions g and $g_{\eta_t^j} = \partial_{\eta_t^j} g$ for $j = 1, \dots, m$ are evaluated at (x_n, η_n) . It was shown in [30] that the scheme converges pathwise with order 2β , where β is the common Hölder exponent of the noise and the component, under the assumption that the vector field g is globally Lipschitz continuous in the state variable and uniformly continuous in the noise variables.

1.2 Taylor Expansions for RODEs

Taylor expansions are used in numerical analysis to derive systematically higher order numerical schemes for ODEs or to check the order of convergence of schemes that have been found by heuristic methods.

1.2.1 ODE Case

The situation is quite straightforward for an ODE

$$\frac{dx}{dt} = f(t, x), \quad x(t_0) = x_0. \tag{1.3}$$

By the chain rule, the total derivative of a function $u(t, x)$ along the solution $x(t) = x(t, t_0, x_0)$ of the ODE (1.3) is given by

$$Du(t, x) := \frac{\partial u}{\partial t}(t, x) + f(t, x) \frac{\partial u}{\partial x}(t, x), \quad (1.4)$$

so the derivatives of the solution can be written compactly as

$$x^{(j)}(t) = D^{j-1}f(t, x(t)), \quad j = 1, 2, \dots, p,$$

where $D^{j+1} = DD^j$ (with $D^0 \equiv Id$ for notational convenience), provided the vector field in (1.3) is p times continuously differentiable in its variables. Substituting into the Taylor expansion

$$\begin{aligned} x(t_0 + h) &= x(t_0) + x'(t_0) h + \dots + \frac{1}{p!} x^{(p)}(t_0) h^p \\ &\quad + \frac{1}{(p+1)!} x^{(p+1)}(\tau_h) h^{p+1}, \quad \tau_h \in [t_0, t_0 + h], \end{aligned}$$

and truncating the remainder gives the p -Taylor approximation

$$\Phi_p(t_0, x_0, h) = x_0 + \sum_{j=1}^p \frac{h^j}{j!} D^{j-1}f(t_0, x_0),$$

of the solution value $x(t_0 + h) = x(t_0 + h, t_0, x_0)$, which is the basis of the p th order Taylor scheme. For more details see Chap. 6.

1.2.2 SODE Case

This above does not work for SODEs because the sample paths are only Hölder continuous, but an integral version of the stochastic chain rule, the Itô formula can be used instead to derive stochastic Taylor expansions in integral form.

This will be developed in Chap. 6. To avoid introducing the required notation now, the idea will be illustrated here in terms of integral versions the ODE (1.3) and the deterministic chain rule (1.4). Essentially, the chain rule is first applied to the function $u(t, x) = f(t, x)$ and then to $u(t, x) = Df(t, x)$, and so on, to obtain

$$\begin{aligned} x(t) &= x_0 + \int_{t_0}^t f(s, x(s)) \, ds \\ &= x_0 + \int_{t_0}^t \left[f(t_0, x_0) + \int_{t_0}^s Df(\tau, x(\tau)) \, d\tau \right] \, ds \end{aligned}$$

$$= x_0 + f(t_0, x_0) \int_{t_0}^t ds + \int_{t_0}^t \int_{t_0}^s Df(\tau, x(\tau)) d\tau ds,$$

which is the first order Taylor expansion. Then, applying (1.4) over the interval $[t_0, \tau]$ to the integrand Df in the double integral remainder term leads to

$$\begin{aligned} x(t) &= x_0 + f(t_0, x_0) \int_{t_0}^t ds + Df(t_0, x_0) \int_{t_0}^t \int_{t_0}^s d\tau ds \\ &\quad + \int_{t_0}^t \int_{t_0}^s \int_{t_0}^\tau D^2 f(r, x(r)) dr d\tau ds. \end{aligned}$$

In this way one obtains the Taylor expansion in integral form

$$\begin{aligned} x(t) &= x(t_0) + \underbrace{\sum_{j=1}^p D^{j-1} f(t_0, x_0) \int_{t_0}^t \int_{t_0}^{s_1} \cdots \int_{t_0}^{s_{j-1}} ds_j \cdots ds_1}_{\text{Taylor approximation}} \\ &\quad + \underbrace{\int_{t_0}^t \int_{t_0}^{s_1} \cdots \int_{t_0}^{s_j} D^p f(s_{j+1}, x(s_{j+1})) ds_{j+1} \cdots ds_1}_{\text{remainder}}. \end{aligned}$$

The SODE case is more complicated by the presence of an additional stochastic integral, but the idea is essentially the same. Importantly, with the corresponding stochastic chain rule, the sample paths of the solutions do not need to be differentiable.

1.2.3 RODE Case

For RODEs with Hölder continuous noise the solutions have Hölder continuous derivatives, but are not further differentiable, while for more irregular noise the RODE solutions are at most continuous. The integral version of the Taylor expansion for ODEs does not work here because it would have to be applied to the vector field $G_\omega(t, x) = g(x, \eta_t(\omega(t)))$, which is not differentiable in the t variable. Instead, three different approaches to derive higher order numerical schemes for RODEs will be presented in this book.

Instead of looking for a Taylor expansion of the solution itself, Jentzen and Kloeden [75, 76, 78, 87] used a Taylor expansion of the vector field of the RODE (1.2) in both of its variables. Replace the process η_t by its Hölder continuous sample paths $t \mapsto \omega(t)$, then the vector field $(x, \omega) \mapsto g(x, \omega)$ has the Taylor expansion²

²Note that the vector field $(x, \omega) \mapsto g(x, \omega)$ can be smooth in its variables, in particular in the ω variable. Later the continuous, but not differentiable sample path $t \mapsto \omega(t)$ of the noise is inserted into this position in the function.

$$g(x(s), \omega(s)) = \sum_{|\alpha| \leq p} \frac{1}{\alpha!} \partial^\alpha g(x_0, \omega_0) (\Delta x_s)^{\alpha_1} (\Delta \omega_s)^{\alpha_2} + R_{p+1}(s),$$

where $\omega_0 := \omega(t_0)$, $x_0 := x(t_0)$ and

$$\Delta \omega_s := \omega(s) - \omega_0, \quad \Delta x_s := x(s) - x_0,$$

with remainder term $R_{p+1}(s)$ and the multi-indice $\alpha = (\alpha_1, \alpha_2) \in \mathbb{N}_0^2$ satisfying

$$|\alpha| := \alpha_1 + \alpha_2, \quad \alpha! := \alpha_1! \alpha_2!.$$

Substituting this into the integral equation version of the RODE gives

$$\Delta x_t = \underbrace{\sum_{|\alpha| \leq p} \frac{1}{\alpha!} \partial^\alpha g(x_0, \omega_0) \int_{t_0}^t (\Delta x_s)^{\alpha_1} (\Delta \omega_s)^{\alpha_2} ds}_{\text{Taylor-like approximation}} + \underbrace{\int_{t_0}^t R_{p+1}(s) ds}_{\text{remainder}}.$$

Then apply the above expression on sub-intervals $[t_n, t_{n+1}]$ with step size $h_n = t_{n+1} - t_n$ and discard the remainder.

The simplest case for $p = 0$ and $\alpha = (0, 0)$ gives the Euler scheme

$$x_{n+1} = x_n + h_n g(x_n, \omega(t_n)).$$

The higher order Taylor-like approximations are implicit in Δx_t , but this difficulty can be resolved by replacing Δx_s inside the integral by a lower order approximation, e.g., for $p = 1$ and $\alpha = (1, 0)$ one has

$$\begin{aligned} x(t_{n+1}) &\approx x(t_n) + h_n g(x(t_n), \omega(t_n)) + \partial_x g(x(t_n), \omega(t_n)) \int_{t_n}^{t_{n+1}} \Delta x_s ds \\ &\approx x(t_n) + h_n g(x(t_n), \omega(t_n)) \\ &\quad + \partial_x g(x(t_n), \omega(t_n)) \int_{t_n}^{t_{n+1}} \underbrace{[(s - t_n) g(x(t_n), \omega(t_n))] ds}_{\text{Euler approximation}}. \end{aligned}$$

For example, the Taylor numerical scheme with multi-indices $\{(0, 0), (1, 0)\}$ is given by

$$x_{n+1} = x_n + h_n g(x_n, \omega(t_n)) + \frac{1}{2} h_n^2 g(x_n, \omega(t_n)) \partial_x g(x_n, \omega(t_n)),$$

and the Taylor numerical scheme with multi-indices $\{(0, 0), (0, 1), (2, 0), (1, 0)\}$ is given by

$$x_{n+1} = x_n + h_n g + \partial_x g \int_{t_n}^{t_{n+1}} \Delta \omega_s \, ds + \frac{1}{2} \partial_x^2 g \int_{t_n}^{t_{n+1}} (\Delta \omega_s)^2 \, ds + (\partial_x g) g \frac{h_n^2}{2}$$

where the functions g , $\partial_x g$ and $\partial_x^2 g$ are evaluated at $(x(t_n), \omega(t_n))$.

The next schemes soon become very cumbersome, but they can be represented compactly with a notation introduced in Jentzen and Kloeden [76, 87], see also [75]. A different, but related approach taken from Jentzen and Kloeden [78] will be used in Chap. 8, since it provides optimal schemes without superfluous terms.

RODEs with Special Structure: Itô Noise and Affine Noise

When the RODEs have a special structure this can be exploited to obtain other kinds of Taylor approximations and numerical schemes, which are often simpler.

In Chap. 10 the noise in the RODE is an Itô diffusion Y_t . The coupled RODE and Itô SODE, in the scalar case,

$$\frac{dx}{dt} = g(x, Y_t), \quad dY_t = f(Y_t)dt + \sigma(Y_t) dW_t,$$

is a system of Itô SODEs

$$dX_t = g(X_t, Y_t)dt + 0 dW_t, \quad dY_t = f(Y_t)dt + \sigma(Y_t) dW_t.$$

In vector notation this is a 2-dimensional Itô SODE

$$d \begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} g(X_t, Y_t) \\ f(Y_t) \end{pmatrix} dt + \begin{pmatrix} 0 \\ \sigma(Y_t) \end{pmatrix} dW_t.$$

Since the order γ Itô-Taylor expansions and schemes (developed in Kloeden and Platen [91] and to be recalled in Chap. 6) can be shown to converge pathwise with order $\gamma - \varepsilon$ for arbitrarily small $\varepsilon > 0$, the first component of these schemes is a pathwise convergent scheme of that order for the RODE.

For example, with $\gamma = 1.5$, the order 1.5 RODE-Taylor scheme for X_t is given by

$$x_{n+1} = x_n + g(x_n, y_n)h_n + \sigma(y_n)\partial_y g(x_n, y_n) I_{(1,0),n} + \frac{1}{2} \left(g(x_n, y_n)\partial_x g(x_n, y_n) + f(y_n)\partial_y g(x_n, y_n) + \frac{1}{2}\sigma^2(y_n)\partial_y^2 g(x_n, y_n) \right) \Delta W_n^2,$$

which includes the correlated stochastic integrals

$$\Delta W_n = I_{(1),n} = \int_{t_n}^{t_{n+1}} dW_s, \quad I_{(1,0),n} = \int_{t_n}^{t_{n+1}} \int_{t_n}^t dW_s dt$$

and

$$h_n = I_{(0),n} = \int_{t_n}^{t_{n+1}} ds = t_{n+1} - t_n.$$

Another special structure is affine noise, which means that the RODE has the form

$$\frac{dx}{dt} = f(t, x) + \sigma(t, x) \eta_t, \quad (1.5)$$

where the noise may now have bounded but only measurable sample paths. These are very similar to affine control systems with bounded measurable controls, which are interpreted as ODEs in the sense of Carathéodory. In a way they also resemble Stratonovich SODEs. The fact that the Stratonovich chain rule is analogous to the deterministic one was exploited by Grüne and Kloeden [55] to adapt Stratonovich stochastic Taylor expansions to affine control systems to derive systematically higher order numerical schemes for deterministic affine control systems. It was later used by Asai and Kloeden [11] to do the same for RODEs with affine structure. These will be discussed in Chaps. 7 and 11. A significant contrast with the other approaches is that the driving noise process does not need to have continuous sample paths.

1.3 RODEs with Bounded Noise

The complexity of biological systems is often a consequence of uncertainty and noise, and thus consideration of noise terms is necessary in mathematical models. There are various ways in which this can be done, see e.g., [1, 114], which usually lead to an Itô SODE.

In biological models bounded noise is often more realistic, see e.g., d’Onofrio [41]. For example, a mass-action interaction is often assumed under idealistic perfect mixing conditions. If the multiplicative mass-action parameter is made noisy with values in a bounded interval, then the resulting RODE has an affine noise structure, i.e., it is of the form (1.5).

Bounded noise can be introduced by allowing the mass-action parameter to vary randomly within a bounded interval about its idealised value. Two possibilities from Asai and Kloeden [9] are given here. Both modify an input noise such as an Ornstein–Uhlenbeck process, denoted by O_t , although in principle any other input process could be used.

In one of the examples, a positive parameter ζ may be replaced by the bounded stochastic process

$$\zeta(O_t) := \zeta_0 \left(1 - 2\nu \frac{O_t}{1 + O_t^2} \right), \quad (1.6)$$

where ζ_0 and ν are positive constants with $\nu \in (0, 1]$. The noise here tends to peak around $\zeta_0(1 \pm \nu)$, and is thus suitable for a noisy switching scenario (see Fig. 1.1).

In the other example, a positive parameter μ may be replaced by the stochastic process

$$\mu(O_t) := \mu_0 \left(1 - \frac{2\nu}{\pi} \arctan O_t \right), \quad (1.7)$$

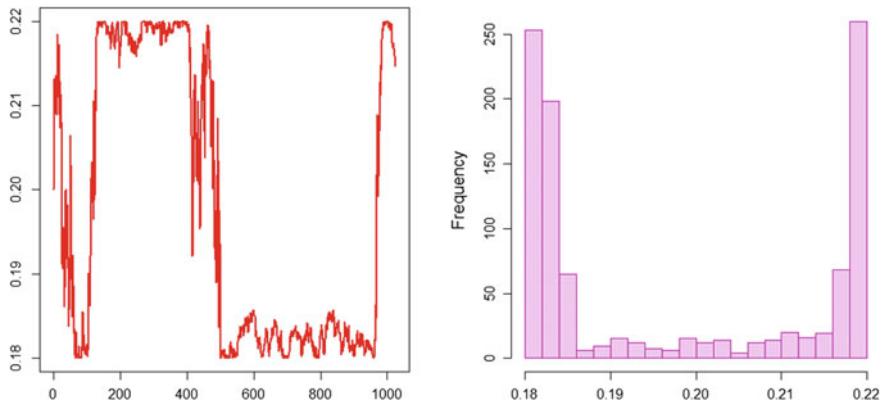


Fig. 1.1 Switching noise (1.6) driven by an Ornstein–Uhlenbeck process

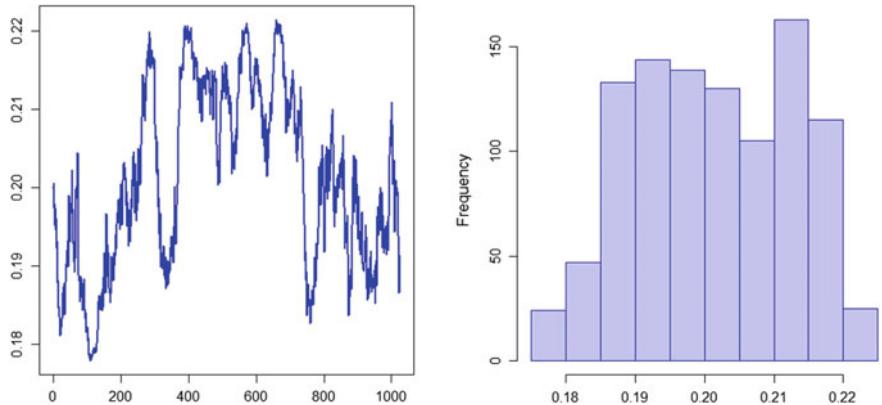


Fig. 1.2 Centered noise (1.7) driven by an Ornstein–Uhlenbeck process

where μ_0 and ν are positive constants with $\nu \in (0, 1]$. The noise then takes values in the interval $(\mu_0(1 - \nu), \mu_0(1 + \nu))$ with the probability density taking its maximum at μ_0 (see Fig. 1.2).

1.4 RODEs and Carathéodory ODEs

If the noise process in a RODE is only measurable rather than continuous in time, the RODE must be interpreted as a Carathéodory ODE, i.e., with absolutely continuous solutions which satisfy the ODE only almost surely. Affine control systems [55], i.e., with an affine structure like (1.5) but with a measurable control function instead of noise are Carathéodory ODEs. Switching ODEs [56] can be formulated in a similar

way with the controls taking only unit coordinate vector values. Similarly, RODEs with an affine structure in the noise must also be considered as Carathéodory ODEs when the noise process is only measurable in time.

A Random Euler Scheme for Carathéodory ODEs

In [79] Jentzen and Neuenkirch analysed a numerical scheme for a general deterministic Carathéodory ODE

$$\frac{dx}{dt} = f(x, t),$$

which converges with order $\frac{1}{2}$. This scheme, called the *random Euler scheme*, is given by

$$x_{n+1} = x_n + h_n \cdot f(x_n, t_n + R_n \cdot h_n),$$

where R_n are independent uniformly distributed random variables on $[0, 1]$.

1.5 Endnotes

A mathematical theory of Brownian motion was, in fact, formulated by Bachelier five years before Einstein's classical paper. In [13] Bachelier assumed that the stock price is a Brownian motion without drift, which implies that the stock prices have a normal distribution.

I. Gikhman [50] and W. Döblin [40] also, independently, developed a similar version of stochastic calculus to Itô's. See also Doob [42].

Numerical schemes of higher order for RODEs will be discussed extensively in this book. RODEs are used by Swishchuk and Wu [130] to model the evolution of biological systems in random media. Bounded noise in biological systems as well as from engineering and physics are discussed in the collection of articles edited by d'Onofrio [41], where other kinds of bounded noise to those above can be found.

Soong [124], Srinivasan and Vasudevan [126] are two older books, which discuss RODEs with random vector fields rather than those driven by a stochastic process. Older references include Wonham [141] who used such RODEs in control theory, while Tiwari and Hobbie [133] and vom Scheidt, Starkloff and Wunderlich [136] considered other applications. Bharucha-Reid [16] and Tsokos and Padgett [134] are two other older books on related random integral equations.

Khasminskii [83] discussed the stability of RODEs and provided references on them to the Russian language up to 1967, when the first edition of his book appeared. Kac and Krasovski [143] investigated the stability of RODEs driven by a finite Markov chain.

Itô SODEs will be discussed in Chap. 3. Stratonovich [127] introduced an alternative form of SODEs which, unlike Itô SODEs, have a similar chain rule to deterministic calculus. For SODEs with Markovian switching see Mao and Yuan [100].

Chapter 2

Random Ordinary Differential Equations

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, where \mathcal{F} is a σ -algebra on Ω and \mathbb{P} is a probability measure, and let $\eta: [0, T] \times \Omega \rightarrow \mathbb{R}^m$ be an \mathbb{R}^m -valued stochastic process with continuous sample paths. In addition, let $g: \mathbb{R}^d \times \mathbb{R}^m \rightarrow \mathbb{R}^d$ be a continuous function.

A random ordinary differential equation (RODE) in \mathbb{R}^d ,

$$\frac{dx}{dt} = g(x, \eta_t(\omega)), \quad x \in \mathbb{R}^d, \tag{2.1}$$

is a nonautonomous ordinary differential equation (ODE)

$$\frac{dx}{dt} = G_\omega(t, x) := g(x, \eta_t(\omega)) \tag{2.2}$$

for almost every realization $\omega \in \Omega$.

A simple example of a scalar RODE is

$$\frac{dx}{dt} = -x + \sin W_t(\omega),$$

where W_t is a scalar Wiener process. Here $g(x, z) = -x + \sin z$ and $d = m = 1$. RODEs with other kinds of noise such as fractional Brownian motion have also been used.

For convenience, it will be assumed that the RODE (2.2) holds for all $\omega \in \Omega$, by restricting Ω to a subset of full probability if necessary, and that g is infinitely often continuously differentiable in its variables, although k -times continuously differentiable with k sufficiently large would suffice. In particular, g is then locally Lipschitz in x , so the initial value problem

$$\frac{dx}{dt} = g(x(t, \omega), \eta_t(\omega)), \quad x(0, \omega) = x_0(\omega), \quad (2.3)$$

where the initial value x_0 is an \mathbb{R}^d -valued random variable, has a unique pathwise solution $x(t, \omega)$ for every $\omega \in \Omega$, which will be assumed to exist on the finite time interval $[0, T]$ under consideration. Sufficient conditions that guarantee the existence and uniqueness of such solutions are similar to those for ODEs and will be considered in Sect. 2.1 of this chapter. The situation is more complicated when the sample paths of the driving noise process η_t are only measurable in t , because then function $G_\omega(t, x)$ is only measurable in t and existence and uniqueness of solutions must now be understood in the sense of Carathéodory.

The solution of the RODE (2.3) is a stochastic process X_t on the interval $[0, T]$. Its sample paths $t \mapsto X_t(\omega)$ are continuously differentiable, but need not be further differentiable, since the vector field $G_\omega(t, x)$ of the nonautonomous ODE (2.2) is usually only at most continuous, but not differentiable in t , no matter how smooth the function g is in its variables.

2.1 Existence and Uniqueness Theorems

Once a sample path of the noise has been fixed a RODE (2.1) is an ODE, in fact a nonautonomous ODE (2.2) since the noise changes the vector field with time.

If the vector field function g in the RODE (2.1) is continuous in both of its variables and the sample paths of the noise process η_t are continuous too, then the vector field function $G_\omega(t, x) := g(x, \eta_t(\omega))$ of the corresponding nonautonomous ODE (2.2) is continuous in both of its variables for each fixed ω . Classical existence and uniqueness theorems for ODEs apply to RODEs in this case. On the other hand, if the sample paths of the noise process η_t are only measurable in t , then function $G_\omega(t, x)$ is only measurable in t and the existence and uniqueness of solutions are to be understood in the sense of Carathéodory.

2.1.1 Classical Assumptions

Suppose that the vector field g in the RODE (2.1) is at least continuous in both of its variables and the sample paths of the noise η_t are continuous. Fix a sample path, i.e., ω , write $G(t, x) := g(x, \eta_t(\omega))$, and consider the initial value problem (IVP)

$$\frac{dx}{dt} = G(t, x), \quad x(t_0) = x_0, \quad x \in \mathbb{R}^d. \quad (2.4)$$

where G is at least continuous.

A solution of the IVP (2.4) is a continuously differentiable function $x : [t_0, T] \rightarrow \mathbb{R}^d$ with $x(t_0) = x_0$ such that

$$\frac{d}{dt}x(t) = G(t, x(t)) \quad \text{for all } t \in (t_0, T).$$

Integrating the RODE (2.4) gives the integral equation

$$x(t) = x_0 + \int_{t_0}^t G(s, x(s))ds, \quad t \in [t_0, T]. \quad (2.5)$$

A solution of the IVP (2.4) is thus a solution of the integral equation (2.5). The converse also holds.

Lemma 2.1 *A continuous function $x : [t_0, T] \rightarrow \mathbb{R}^d$ satisfying the integral equation (2.5) is a solution of the IVP (2.4). In particular, it is continuously differentiable in (t_0, T) .*

Proof The mapping $t \mapsto G(t, x(t))$ is continuous, because the mappings $t \mapsto x(t)$ and $(t, x) \mapsto G(t, x)$ are continuous. Hence, the fundamental theorem of integral and differential calculus applies here and gives

$$\frac{d}{dt} \int_{t_0}^t G(s, x(s))ds = G(t, x(t)) \quad \text{for each } t \in (t_0, T).$$

This means that the right-hand side of the integral equation (2.5) is continuously differentiable. Hence, when $x(t)$ satisfies the integral equation (2.5), it too is continuously differentiable and satisfies

$$\frac{d}{dt}x(t) = G(t, x(t)) \quad \text{for each } t \in (t_0, T).$$

Finally, for $t = t_0$, the integral equation reduces to $x(t_0) = x_0$, so x is a solution of the Eq. (2.4). \square

The IVP (2.4) and the integral equation (2.5) are equivalent, but the integral equation is theoretically more convenient because it requires only continuity and not continuous differentiability. Another advantage is that the solution is a fixed point of an integral operator in the integral equation representation of the IVP.

The classical existence and uniqueness theorem due to Picard and Lindelöf is proved using a convergent sequence of successive approximations. The result holds under a local Lipschitz assumption, which is satisfied if G is continuously differentiable in x , but existence may only hold on a smaller time interval, the length of which may depend on the initial value.

Theorem 2.1 (The Picard–Lindelöf Theorem) *Let $G : [t_0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ be continuous on a parallelepiped $R := \{(t, x) : t_0 \leq t \leq t_0 + a, |x - x_0| \leq b\}$ and*

uniformly Lipschitz continuous in x and continuous in t . In addition, let M be a bound for $|G(t, x)|$ on R and denote by $\mu := \min\{a, b/M\}$. Then the initial value problem (2.4) has a unique solution $x^* = x^*(t)$ on $[t_0, t_0 + \mu]$.

Proof Let $x_0(t) = x_0$. Suppose that $x_k(t)$ has been defined on $[t_0, t_0 + \mu]$, is continuous, and satisfies $|x_k(t) - x_0| \leq b$ for $k = 0, \dots, n$. Put

$$x_{n+1}(t) = x_0 + \int_{t_0}^t G(s, x_n(s))ds. \quad (2.6)$$

Then $x_{n+1}(t)$ is defined and continuous on $[t_0, t_0 + \mu]$, since $G(t, x_n(t))$ is so. Also it is clear that

$$|x_{n+1}(t) - x_0| \leq \int_{t_0}^t |G(s, x_n(s))|ds \leq M\mu \leq b.$$

Hence $x_0(t), x_1(t), \dots$ are defined and continuous on $[t_0, t_0 + \mu]$, and satisfy $|x_n(t) - x_0| \leq b$. It will be shown next by induction that

$$|x_{n+1}(t) - x_n(t)| \leq \frac{M\kappa^n(t - t_0)^{n+1}}{(n+1)!} \quad \text{for } t_0 \leq t \leq t_0 + \mu, \quad n = 0, 1, \dots, \quad (2.7)$$

where κ is a Lipschitz constant for G in its x component.

First it is straightforward to see that (2.7) holds for $n = 0$. Assume that (2.7) holds for $1, \dots, n - 1$. By (2.6), for $n \geq 1$:

$$x_{n+1}(t) - x_n(t) = \int_{t_0}^t [G(s, x_n(s)) - G(s, x_{n-1}(s))]ds.$$

Hence the Lipschitz condition of G implies that

$$\begin{aligned} |x_{n+1}(t) - x_n(t)| &\leq \kappa \int_{t_0}^t |x_n(s) - x_{n-1}(s)|ds \\ &\leq \frac{M\kappa^n}{n!} \int_{t_0}^t (s - t_0)^n ds = \frac{M\kappa^n(t - t_0)^{n+1}}{(n+1)!}, \end{aligned}$$

which proves (2.7).

Now consider

$$x(t) = x_0 + \sum_{n=0}^{\infty} [x_{n+1}(t) - x_n(t)].$$

It follows from (2.7) that $x(t)$ is uniformly convergent on $[t_0, t_0 + \mu]$, i.e.,

$$\lim_{n \rightarrow \infty} x_n(t) = x^*(t) \quad \text{exists uniformly.}$$

Since $G(t, x)$ is uniformly continuous on R due to its continuity and boundedness it follows that $G(t, x_n(t))$ converges to $G(t, x^*(t))$ uniformly on $[t_0, t_0 + \mu]$ as $n \rightarrow \infty$. Thus (2.6) can be integrated term by term to give

$$x^*(t) = x_0 + \int_{t_0}^t G(s, x^*(s))ds,$$

i.e., $x^*(t) = \lim_{n \rightarrow \infty} x_n(t)$ is a solution to (2.4).

To prove uniqueness, let $y(t)$ be any other solution to (2.4) on $[t_0, t_0 + \mu]$, then

$$y(t) = x_0 + \int_{t_0}^t G(s, y(s))ds,$$

and it follows from induction that

$$|x_n(t) - y(t)| \leq \frac{M\kappa^n(t - t_0)^{n+1}}{(n + 1)!} \quad \text{for } t_0 \leq t \leq t_0 + \mu, \quad n = 0, 1, \dots \quad (2.8)$$

Letting $n \rightarrow \infty$ in (2.8) gives immediately that $|x^*(t) - y(t)| = 0$, so $y(t) \equiv x^*(t)$.

□

Remark 2.1 If the vector field satisfies a global Lipschitz condition, then existence on the entire time interval is obtained. The above proof can be used in this case but requires the solutions on sufficiently small subintervals to be patched together. An alternative proof uses the Banach contraction mapping theorem on the Banach space $\mathcal{C}([t_0, T], \mathbb{R}^d)$ of continuous functions $x : [t_0, T] \rightarrow \mathbb{R}^d$ with the supremum norm

$$\|x\|_\infty = \max_{t_0 \leq t \leq T} |x(t)|.$$

The existence can be obtained on the entire interval in one step using the exponential norm

$$\|x\|_{\exp} = \max_{t_0 \leq t \leq T} \{|x(t)| e^{-2\kappa t}\},$$

which is equivalent to the supremum norm on the space $\mathcal{C}([t_0, T], \mathbb{R}^d)$. (Here κ is the Lipschitz constant).

The next theorem drops the Lipschitz assumption and sacrifices the uniqueness of solutions.

Theorem 2.2 (Peano's Existence Theorem) *Let $G : [t_0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ be continuous on a parallelepiped $R := \{(t, x) : t_0 \leq t \leq t_0 + a, |x - x_0| \leq b\}$. In addition, let M be an upper bound for $|G(t, x)|$ on R and denote by $\mu := \min\{a, b/M\}$. Then the IVP (2.4) has at least one solution $x = x(t)$ on $[t_0, t_0 + \mu]$.*

Proof Let $\delta > 0$ and $x_0(t)$ be a continuously differentiable d -dimensional vector-valued function on $[t_0 - \delta, t_0]$ satisfying $x_0(t_0) = x_0$, $|x_0(t) - x_0| \leq b$ and $|x_0(t) - x_0(s)| \leq M|t - s|$ for all $t, s \in [t_0 - \delta, t_0]$.

For any $0 < \varepsilon \leq \delta$, define a function $x_\varepsilon(t)$ on $[t_0 - \delta, t_0 + \mu]$ by

$$x_\varepsilon(t) = \begin{cases} x_0(t), & t \in [t_0 - \delta, t_0], \\ x_0 + \int_{t_0}^t G(s, x_\varepsilon(s - \varepsilon))ds, & t \in [t_0, t_0 + \mu]. \end{cases} \quad (2.9)$$

Note that (2.9) defines an extension $x_\varepsilon(t)$ of $x_0(t)$ from $[t_0 - \delta, t_0]$ to $[t_0 - \delta, t_0 + \min\{\mu, \varepsilon\}]$ and satisfies on this interval

$$|x_\varepsilon(t) - x_0| \leq b, \quad |x_\varepsilon(t) - x_\varepsilon(s)| \leq M|t - s|. \quad (2.10)$$

By (2.9), $x_\varepsilon(t)$ can be extended as a \mathcal{C}^0 function over $[t_0 - \delta, t_0 + \min\{\mu, 2\varepsilon\}]$ so it satisfies (2.10). Continuing in this manner, (2.9) serves to define $x_\varepsilon(t)$ on $[t_0 - \delta, t_0 + \mu]$ such that $x_\varepsilon(t)$ is a \mathcal{C}^0 function on $[t_0 - \delta, t_0 + \mu]$ and satisfies (2.10).

It follows that the family of functions, $\{x_\varepsilon(t)\}_{0 < \varepsilon \leq \delta}$ is equicontinuous. Hence by the Arzelà Selection Theorem, there exists a sequence $\varepsilon_1 > \varepsilon_2 > \dots$, such that $\varepsilon_n \rightarrow 0$ as $n \rightarrow \infty$ and

$$\lim_{n \rightarrow \infty} x_{\varepsilon_n}(t) = x(t) \text{ exists uniformly}$$

on $[t_0 - \delta, t_0 + \mu]$. By the uniform continuity of G , it follows that $G(t, x_{\varepsilon_n}(t - \varepsilon_n))$ converges to $G(t, x(t))$ uniformly as $n \rightarrow \infty$. Hence, integrating (2.9) with $\varepsilon = \varepsilon_n$ term by term gives

$$x(t) = x_0 + \int_{t_0}^t G(s, x(s))ds,$$

i.e., $x(t)$ is a solution to the IVP (2.4). □

2.1.2 Measurability of Solutions

Let (Ω, \mathcal{F}) be a measurable space and let $\mathcal{B}(\mathbb{R}^d)$ denote the σ -algebra of the Borel subsets of \mathbb{R}^d . A mapping $\phi : \Omega \rightarrow \mathbb{R}^d$ is said to be measurable if for any $B \in \mathcal{B}(\mathbb{R}^d)$,

$$\phi^{-1}(B) = \{\omega \in \Omega | \phi(\omega) \in B\} \in \mathcal{F}.$$

The vector field G in the IVP (2.4), in fact, depends on ω . Then the IVP can be rewritten as

$$\frac{dx}{dt} = G(\omega, t, x), \quad x(t_0) = x_0, \quad x \in \mathbb{R}^d, \quad (2.11)$$

where G is measurable in ω and continuous in (t, x) , since $G(\omega, t, x) := g(x, \eta_t(\omega))$ and g is at least continuous in both variables, and $\eta_t(\omega)$ is measurable in ω and has continuous sample paths.

Lemma 2.2 Suppose that the IVP (2.11) has a unique solution, denoted by $x(t, \omega)$. Then the mapping $\Omega \rightarrow \mathbb{R}^d$ defined by $\omega \mapsto x(t, \omega)$ is measurable for each t .

Proof Choose and fix a function $x_0(t) \in \mathcal{C}([t_0, T], \mathbb{R}^d)$ such that $x_0(t_0) = x_0$. Define a sequence of functions $x_n : [t_0, T] \times \Omega \rightarrow \mathbb{R}^d$ by

$$x_{n+1}(t, \omega) = x_0 + \int_{t_0}^t G(\omega, s, x_n(s, \omega))ds.$$

Similar to the proof of Theorem 2.2, it follows that

$$x(t, \omega) = \lim_{n \rightarrow \infty} x_n(t, \omega), \quad \text{for all } \omega \in \Omega.$$

Hence it is sufficient to prove that the mappings $x_n(t, \cdot) : \Omega \rightarrow \mathbb{R}^d$ are measurable for all $t \in [t_0, T]$ and $n \in \mathbb{N}$, which can be done by induction.

First, the statement holds for $n = 0$. Next suppose that for some $n \in \mathbb{N}$ and all $t \in [t_0, T]$, the function $x_n(t, \cdot) : \Omega \rightarrow \mathbb{R}^d$ is measurable. Define $x_n^{(k)} : [t_0, t] \times \Omega \rightarrow \mathbb{R}^d$ by

$$x_n^{(k)}(s, \omega) = \sum_{i=0}^{k-1} \chi_{\left[\frac{it}{k}, \frac{(i+1)t}{k}\right)}(s) \cdot x_n\left(\frac{it}{k}, \omega\right) \quad \text{for all } (s, \omega) \in [t_0, t] \times \Omega,$$

where χ_I is an indicator function with value 1 for $x \in I$ and 0 otherwise.

Using the fact that $G : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is a continuous function,

$$\lim_{k \rightarrow \infty} \int_{t_0}^t G(\omega, s, x_n^{(k)}(s, \omega))ds = \int_{t_0}^t G(\omega, s, x_n(s, \omega))ds.$$

Therefore the mapping $\omega \mapsto \int_{t_0}^t G(s, x_n(s, \omega))ds$ is $\mathcal{B}(\mathbb{R}^d)$ -measurable for each $t \in [t_0, T]$, which implies that the function $x_{n+1}(t, \cdot)$ is $\mathcal{B}(\mathbb{R}^d)$ -measurable for each $t \in [t_0, T]$. By induction, the mapping $x_n(t, \cdot)$ is $\mathcal{B}(\mathbb{R}^d)$ -measurable for all $t \in [t_0, T]$ and $n \in \mathbb{N}$. \square

The above result holds also if the initial value is measurable, i.e., a random variable with values $x_0(\omega)$.

2.1.3 Carathéodory Assumptions

The equivalence of the IVP (2.4) and the integral equation (2.5) for a RODE whose vector field is continuous in both variables does not hold when its vector field function is only measurable, but not continuous in time. In fact, the concept of a solution also needs to be modified in this case to one defined in the sense of Carathéodory. It is based on absolutely continuous functions.

Definition 2.1 A function $x : [t_0, T] \rightarrow \mathbb{R}^d$ is said to be *absolutely continuous* if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\sum_{i=1}^n |x(t_i) - x(s_i)| < \varepsilon$ whenever $\{[s_i, t_i] : 1 \leq i \leq n\}$ is a finite collection of non-overlapping intervals in $[t_0, T]$ that satisfy $\sum_{i=1}^n (t_i - s_i) < \delta$.

An absolutely continuous function is uniformly continuous. In fact, it is also of bounded variation, hence the difference of two monotone functions, from which it follows that it is weakly differentiable almost everywhere in $[t_0, T]$ and its derivative is Lebesgue integrable. Recall that the weak derivative here is defined by

$$\int_{t_0}^T x'(s)\phi(s) ds = - \int_{t_0}^T x(s)\phi'(s) ds$$

for all $\phi \in \mathcal{C}_0^\infty((t_0, T), \mathbb{R})$, i.e., \mathcal{C}^∞ functions with compact support [146].

Lemma 2.3 (Lemma 4.11 and Theorem 4.12 [52]) *Let $f : [t_0, T] \rightarrow \mathbb{R}$ be bounded and Lebesgue measurable, resp., Lebesgue integrable. If $x(t) = \int_{t_0}^t f(s) ds$ for each $t \in [t_0, T]$, then x is absolutely continuous in $[t_0, T]$ and its weak derivative $x'(t) = f(t)$ for almost all $t \in [t_0, T]$.*

Lemma 2.3 can be used to show that the solution of Eq. (2.4) satisfies the integral equation (2.5) and vice versa, i.e., a counterpart of Lemma 2.1 holds here. Note that if $x : [t_0, T] \rightarrow \mathbb{R}^d$ is continuous then the mapping $t \mapsto G(t, x(t))$ is Lebesgue integrable on $[t_0, T]$. The right-hand side of Eq. (2.4) can be integrated in the sense of Lebesgue with the given initial condition to give the right-hand side of the integral equation (2.5). By Lemma 2.3 the function

$$t \mapsto x_0 + \int_{t_0}^t G(s, x(s)) ds$$

is absolutely continuous on $[t_0, T]$ and its weak derivative is equal to $G(t, x(t))$ for Lebesgue almost all $t \in [t_0, T]$.

Definition 2.2 A function $G : [t_0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ of the IVP (2.4) is said to satisfy the *Carathéodory conditions* if

- (C1) (continuity) $G(t, x)$ is continuous in x for almost every $t \in [t_0, T]$;
- (C2) (measurability) $G(t, x)$ is Lebesgue measurable in t for each $x \in \mathbb{R}^d$;
- (C3) (boundedness) $|G(t, x)| \leq m(t)$ for each $x \in \mathbb{R}^d$ and almost every $t \in [t_0, T]$ for some absolutely continuous function $m(t)$.

Let $G : [t_0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ satisfy the Carathéodory conditions. Then a solution (in the extended sense) of the IVP (2.4) with vector field G is an absolutely continuous function $x : [t_0, T] \rightarrow \mathbb{R}^d$ with $x(t_0) = x_0$ for which the weak derivative satisfies

$$\frac{d}{dt} x(t) = G(t, x(t)) \quad \text{for Lebesgue almost all } t \in [t_0, T]. \quad (2.12)$$

The counterpart of the Picard–Lindelöf Theorem 2.1 holds for initial value problems under the Carathéodory conditions with solutions defined in this way. In 1918 Carathéodory [29] established a local existence result under assumptions (C1)–(C3), stated as follows.

Theorem 2.3 (Theorem 1.1, Chap. 2 [32]) *Let G be defined on $R := \{(t, x) : t_0 \leq t \leq T, |x - x_0| < b\}$ and satisfy the Carathéodory conditions. Then the IVP (2.4) has a solution $x^* : [t_0, t_0 + \delta] \rightarrow \mathbb{R}^d$ in the extended sense of Eq. (2.12).*

Proof For any $t \in [t_0, T]$, define $M(t)$ by

$$M(t) := \int_{t_0}^t m(s)ds. \quad (2.13)$$

Then $M(t)$ is a continuous nondecreasing function satisfying $M(t_0) = 0$. Therefore $(t, x_0 \pm M(t)) \in R$ on some interval $t_0 \leq t \leq t_0 + \delta \leq T$, where δ is some positive constant. For this $\delta > 0$, define function $x_n(t)$, $n = 1, 2, \dots$, by

$$x_n(t) := x_0, \quad t_0 \leq t \leq t_0 + \frac{\delta}{n}, \quad (2.14)$$

$$x_n(t) := x_0 + \int_{t_0}^{t-\delta/n} G(s, x_n(s))ds, \quad t_0 + \frac{\delta}{n} < t \leq t_0 + \delta. \quad (2.15)$$

It is clear that $x_1(t) = x_0$ is defined on $[t_0, t_0 + \delta]$. For any $n \geq 1$, formula (2.14) defines x_n on $[t_0, t_0 + \delta/n]$ and since $(t, x_0) \in R$ for $t \in [t_0, t_0 + \delta/n]$, formula (2.15) defines x_n as a continuous function on $(t_0 + \delta/n, t_0 + 2\delta/n]$. Furthermore, due to (C3) and (2.14) for any $t \in (t_0 + \delta/n, t_0 + 2\delta/n]$,

$$|x_n(t) - x_0| \leq M(t - \delta/n). \quad (2.16)$$

Assume that x_n is defined on $[t_0, t_0 + k \cdot \delta/n]$ for $1 < k < n$. Since the measurability of the integrand in (2.15) is only required on $[t_0, t_0 + k \cdot \delta/n]$, formula (2.15) defines x_n for $t \in (t_0 + k \cdot \delta/n, t_0 + (k+1) \cdot \delta/n]$. In addition, due to (C3) and (2.13), $x_n(t)$ satisfies (2.16) on $(t_0 + k \cdot \delta/n, t_0 + (k+1) \cdot \delta/n]$. Therefore by induction, (2.14) and (2.15) define all $x_n(t)$, $n = 1, 2, \dots$, as continuous functions on $t \in [t_0, t_0 + \delta]$ satisfying

$$x_n(t) = x_0, \quad t_0 \leq t \leq t_0 + \frac{\delta}{n}, \quad (2.17)$$

$$|x_n(t) - x_0| \leq M \left(t - \frac{\delta}{n} \right), \quad t_0 + \frac{\delta}{n} < t \leq t_0 + \delta. \quad (2.18)$$

For any $t_1, t_2 \in [t_0, t_0 + \delta]$, according to (C3), (2.13)–(2.15),

$$|x_n(t_1) - x_n(t_2)| \leq \left| M \left(t_1 - \frac{\delta}{n} \right) - M \left(t_2 - \frac{\delta}{n} \right) \right|. \quad (2.19)$$

Since M is continuous on $[t_0, t_0 + \delta]$, it is uniformly continuous on $[t_0, t_0 + \delta]$. Thus by (2.17)–(2.19), the sequence $\{x_n\}$ is equicontinuous and uniformly bounded on $[t_0, t_0 + \delta]$. It then follows from the Ascoli Lemma that there exists a subsequence $\{x_{n_j}\}$ of $\{x_n\}$ such that $\{x_{n_j}\}$ converges uniformly to a continuous limit function, namely, $x(t)$, on $[t_0, t_0 + \delta]$ as $j \rightarrow \infty$.

By (C3), $|G(t, x_{n_j}(t))| \leq m(t)$ for any $t \in [t_0, t_0 + \delta]$. Also, by (C1), G is continuous in x for almost every fixed t , so

$$\lim_{j \rightarrow \infty} G(t, x_{n_j}(t)) = G(t, x(t)) \quad \text{for almost every fixed } t \in [t_0, t_0 + \delta].$$

Therefore, by the Lebesgue dominated convergence theorem,

$$\lim_{j \rightarrow \infty} \int_{t_0}^t G(s, x_{n_j}(s))ds = \int_{t_0}^t G(s, x(s))ds, \quad \text{for all } t \in [t_0, t_0 + \delta]. \quad (2.20)$$

Noticing that

$$x_{n_j}(t) = x_0 + \int_{t_0}^t G(s, x_{n_j}(s))ds - \int_{t-\delta/n_j}^{t_0} G(s, x_{n_j}(s))ds,$$

and that

$$\lim_{j \rightarrow \infty} \int_{t-\delta/n_j}^{t_0} G(s, x_{n_j}(s))ds = 0,$$

it follows from (2.20) that

$$x(t) = x_0 + \int_{t_0}^t G(s, x(s))ds, \quad \text{for all } t \in [t_0, t_0 + \delta]. \quad \square$$

2.1.4 Positivity of Solutions

A solution $x(t) = (x_1(t), x_2(t), \dots, x_d(t))^T$ of (2.12) is called *positive* (*strongly positive*, resp.) if

$$x_i(t) \geq 0 \quad (> 0, \text{ resp.}) \quad \text{for all } t \text{ and } i = 1, 2, \dots, d.$$

The positivity of solutions is important in biological models like those that will be considered in Part IV (see, in particular, Chap. 18) as well as in physics, chemistry and engineering. The following conditions guarantee the positivity of solutions to (2.4).

Definition 2.3 A function $u = (u_1, u_2, \dots, u_d)^T : \mathbb{R}^1 \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ is called *quasi-positive*, or *off-diagonal positive*, if, for each $i = 1, 2, \dots, d$,

$$u_i(t, x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_d) \geq 0,$$

whenever $x_j \geq 0$ for $j \neq i$.

A function $u = (u_1, u_2, \dots, u_d)^\top$ is called *strongly quasipositive*, or *strongly off-diagonal positive*, if, for each $i = 1, 2, \dots, d$,

$$u_i(t, x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_d) > 0,$$

whenever $x_j \geq 0$ for all j and $\sum_j x_j > 0$.

Theorem 2.4 *If the vector field G of (2.4) is quasipositive, then the solution $x(t)$ of (2.4) satisfying the initial condition $x(0) = x_0 = (x_{0,1}, x_{0,2}, \dots, x_{0,d})$ is positive for $t \geq 0$ whenever $x_{0,i} \geq 0$ for $i = 1, 2, \dots, d$.*

If the vector field G of (2.4) is strongly quasipositive, then the solution $x(t)$ of (2.4) satisfying the initial condition $x(0) = x_0 = (x_{0,1}, x_{0,2}, \dots, x_{0,d})$ is strongly positive for $t \geq 0$ whenever $x_{0,i} > 0$ for $i = 1, 2, \dots, d$.

The proof of Theorem 2.4 for the case that $G(t, x)$ is continuous in t and x can be found, e.g., in Krasnosel'skii [95, Lemma 4.1]. It was noted in Szarski [131] that similar statements are also valid when $G(t, x)$ satisfies the Carathéodory conditions.

2.2 RODEs with Canonical Noise

RODEs typically involve given stochastic processes in their vector fields which can differ from example to example. The theory of random dynamical systems, in contrast, is formulated abstractly in terms of a canonical noise process. This allows greater generality and is, in particular, independent of the dimension of the driving noise process. The canonical noise process is represented by a measurable theoretical autonomous dynamical system θ on the sample space Ω of some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Specifically, it is a group under composition of measure preserving transformations $\theta_t : \Omega \rightarrow \Omega$, $t \in \mathbb{R}$, i.e., satisfying

- (i) $\theta_0 = Id$ (identity) and $\theta_t \circ \theta_s = \theta_{t+s}$ for all $t, s \in \mathbb{R}$,
- (ii) the map $(t, \omega) \mapsto \theta_t(\omega)$ is measurable and invariant with respect to \mathbb{P} in the sense that $\theta_t(\mathbb{P}) = \mathbb{P}$ for all $t \in \mathbb{R}$.

The notation $\theta_t(\mathbb{P}) = \mathbb{P}$ for the measure preserving property of θ_t with respect to \mathbb{P} is just a compact way of writing $\mathbb{P}(\theta_t(A)) = \mathbb{P}(A)$ for all $A \in \mathcal{F}$ and $t \in \mathbb{R}$.

In this context RODEs have the form

$$\frac{dx}{dt} = g(x, \theta_t(\omega)), \quad (2.21)$$

where the vector field function $g : \mathbb{R}^d \times \Omega \rightarrow \mathbb{R}^d$ is assumed to be suitably smooth in its first variable and measurable in the second.

Consider the simple scalar RODE

$$\frac{dx}{dt} = -x + W_t(\omega),$$

where the driving noise process is a two-sided¹ Wiener process, i.e., defined for $t \in \mathbb{R}$. The canonical noise system θ is not given directly in terms of the Wiener process W_t , but is defined in terms of shift operators θ_t on the canonical sample space $\Omega := \mathcal{C}_0(\mathbb{R}, \mathbb{R})$ of continuous functions $\omega : \mathbb{R} \rightarrow \mathbb{R}$ with $\omega(0) = 0$, i.e., with

$$\theta_t(\omega(\cdot)) := \omega(t + \cdot) - \omega(\cdot) \quad \text{for all } t \in \mathbb{R}.$$

The σ -algebra of Borel subsets of $\mathcal{C}_0(\mathbb{R}, \mathbb{R})$ is taken as the σ -algebra of events \mathcal{F} , while \mathbb{P} is the corresponding Wiener measure.² Essentially, the canonical noise system here is represented by the sample paths of the Wiener process.

2.3 Endnotes

Bunke's monograph [20], which is in German, is a classical reference on RODEs and contains most of the relevant literature before the 1970s. Bobrowski [18], which is in Polish, is similar in content. Sufficient conditions guaranteeing the existence and uniqueness of solutions of RODEs can be found in Arnold [4] and Bunke [20].

Properties of absolutely continuous and weakly differentiable functions are discussed in Evans and Gariepy [47], Gordon [52], Leoni [98] and Ziemer [146]. Existence and uniqueness theorems under classical and Carathéodory conditions are discussed extensively in Coddington and Levinson [32]. See also Carathéodory [29] and, for more general conditions, Goodman [51] and Biles and Binding [17]. See also Jentzen and Neuenkirch [79].

Monotonicity conditions on differential equations and the positivity of solutions are discussed in [95, 121, 131, 137].

Kac and Krasovski [143] considered RODEs driven by a finite Markov chain, while Arnold and Kloeden [5] analysed 2-dimensional RODEs driven by telegraphic noise. RODEs with fractional Brownian motion were investigated in Garrido-Atienza, Kloeden and Neuenkirch [49]. The recent book [108] by Neckel and Rupp focuses on modeling with RODEs.

¹Essentially, $\{W_t, t \geq 0\}$, and $\{W_{-t}, t \leq 0\}$ are two independent Wiener processes.

²No other topological properties of the space $\mathcal{C}_0(\mathbb{R}, \mathbb{R})$ are used here apart from those defining the Borel sets.

A systematic treatment of the random dynamical system theory and RODEs in the form (2.21) is given in Arnold [4]. This theory will be briefly reviewed in Chap. 4.

The discretisation of RODEs near a saddle point is investigated in Arnold and Kloeden [6]. A delay differential equation with the randomness in the delay under discretisation was investigated by Caraballo, Kloeden and Real [26].

Chapter 3

Stochastic Differential Equations

Stochastic differential equations (SODEs) are motivated by noisy differential equations with an affine structure

$$\frac{dx}{dt} = f(t, x) + \sigma(t, x)\eta_t, \quad (3.1)$$

where η_t is a Gaussian white noise, i.e., a stochastic process consisting of $\mathcal{N}(0, 1)$ -distributed random variables for each t that are independent for different time instants.

Physicists often consider Gaussian white noise to be the derivative of a Wiener process W_t ,¹ the stochastic process describing Brownian motion, and write the noisy differential equation (3.1) as

$$\frac{dx}{dt} = f(t, x) + \sigma(t, x) \dot{W}_t$$

or in integral form as

$$x(t) = x(t_0) + \int_{t_0}^t f(s, x(s)) ds + \int_{t_0}^t \sigma(s, x(s)) dW_s,$$

where the first integral is pathwise a deterministic Riemann integral, while the second looks like a pathwise Riemann–Stieltjes integral.

The above interpretation has, however, many mathematical deficiencies that need to be clarified and overcome. To begin, Gaussian white noise does not exist as a function in the conventional sense, so the noisy differential equation (3.1) is not even a RODE. Moreover, sample paths of a Wiener process are not differentiable² and still worse they do not have bounded variation on any time interval, however small.

¹A Wiener process is often called a Brownian motion in the literature.

²The sample paths of a Wiener process are, however, weakly differentiable.

It took a long time for mathematicians to develop an appropriate mathematical theory for stochastic differential equations, with the breakthrough being made by K. Itô in the 1940s. Itô introduced what is now called the Itô stochastic integral and developed the stochastic calculus for such integrals. The essential point is that an Itô stochastic differential equation (SODE) is not a differential equation at all, but an integral equation. The differential version of an SODE is just a symbolic shorthand representation.

The basic ideas and results on Itô SODEs will be sketched in this chapter for the scalar case. Vector valued analogues will be introduced later in the book as they are required.

3.1 Wiener Processes and Itô Integrals

A Wiener Process $\{W_t, t \geq 0\}$ is one of the most fundamental stochastic processes. It is defined by the properties:

- (i) $W_t \sim \mathcal{N}(0, t)$ for each $t \geq 0$, i.e., W_t is Gaussian distributed and satisfies
 - (a) $W_0 = 0$ with probability 1;
 - (b) $\mathbb{E}[W_t] = 0$ for each $t \geq 0$;
 - (c) $\mathbb{E}[W_t^2] = t$ for each $t \geq 0$;
- (ii) the nonoverlapping increments of W_t are independent, i.e.,

$$W_{t_2} - W_{t_1} \text{ and } W_{t_4} - W_{t_3}$$

are independent random variables for all $0 \leq t_1 < t_2 \leq t_3 < t_4$.

It follows that $W_t - W_s \sim \mathcal{N}(0, t - s)$ for all $0 \leq s < t$, and hence

$$\mathbb{E}[W_t - W_s] = 0, \quad \mathbb{E}[(W_t - W_s)^2] = t - s.$$

In addition, $\mathbb{E}[(W_t - W_s)^4] = 3(t - s)^2$. Hence, by the Kolmogorov criterion, the sample paths of the Wiener process are continuous. However, these sample paths are nowhere differentiable and do not have bounded variation on any bounded time interval. This means, in particular, that a stochastic integral

$$\int_0^T \sigma(s, \omega) dW_s(\omega)$$

cannot be defined pathwise as a Riemann–Stieltjes integral.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and consider a family of random real-valued functions $\sigma : [0, T] \times \Omega \rightarrow \mathbb{R}$ with the properties:

- (i) $\sigma_t(\cdot) \equiv \sigma(t, \cdot)$ is a random variable, i.e., is \mathcal{F} -measurable with $\mathbb{E}[\sigma_t^2] < \infty$ for each $t \in [0, T]$;

- (ii) the path $t \mapsto \sigma_t(\omega)$ is continuous on $[0, T]$ for almost all ω ;
- (iii) the stochastic process $\{\sigma_t, t \in [0, T]\}$ is non-anticipative w.r.t. the Wiener process $\{W_t, t \geq 0\}$, i.e., σ_t and $W_{t+h} - W_t$ are independent random variables for every $h > 0$ and $t \in [0, T]$.

Considering a partition of the interval $[0, T]$, $0 = t_0^{(N)} < t_1^{(N)} < \dots < t_N^{(N)} = T$, the Itô stochastic integral of such a function σ in the above class is defined as the mean-square limit

$$\int_0^T \sigma(t, \omega) dW_t(\omega) = \text{ms-} \lim_{N \rightarrow \infty} \sum_{j=0}^{N-1} \sigma(t_j^{(N)}, \omega) \left(W_{t_{j+1}^{(N)}}(\omega) - W_{t_j^{(N)}}(\omega) \right). \quad (3.2)$$

The limit in (3.2) does not depend on the sequence of partitions of the interval $[0, T]$ used, just the fact that

$$h^{(N)} = \max_{j=0, \dots, N-1} (t_{j+1}^{(N)} - t_j^{(N)}) \rightarrow 0 \quad \text{as } N \rightarrow \infty.$$

An Itô stochastic integral $\int_0^T \sigma_t dW_t$ has the important properties :

$$\mathbb{E} \left[\int_0^T \sigma_t dW_t \right] = 0, \quad \mathbb{E} \left[\left(\int_0^T \sigma_t dW_t \right)^2 \right] = \int_0^T \mathbb{E} [\sigma_t^2] dt.$$

These follow from the choice of evaluation point of the integrand function at the start of each subinterval, which means that the above Itô sums involve products of independent random variables. The expectation of a product of independent random variables is the product of their expectations and the variance of a sum of independent random variables is equal to the sum of their variances.

The second property here is called the *Itô isometry*. It plays an important role in estimating expressions involving Itô stochastic integrals, in particular, in the proof of existence and uniqueness of solutions of Itô SODEs.

3.2 Itô Stochastic Differential Equations

An Itô stochastic differential equation

$$dX_t = f(t, X_t) dt + \sigma(t, X_t) dW_t$$

is, in fact, a stochastic integral equation

$$X_t = X_{t_0} + \int_{t_0}^t f(s, X_s) ds + \int_{t_0}^t \sigma(s, X_s) dW_s. \quad (3.3)$$

A solution X_t of the SODE (3.3) is a stochastic process which satisfies the stochastic integral equation and has the properties:

- (i) X_t is a random variable with $\mathbb{E}[X_t^2] < \infty$ for each $t \in [t_0, T]$;
- (ii) X_t is non-anticipative w.r.t. the Wiener process W_t ;
- (iii) the sample path $t \mapsto X_t(\omega)$ is continuous for almost all ω .

The existence and uniqueness theorem, Theorem 2.1, for RODEs has the following counterpart for the SODE (3.3) in the scalar case.

Theorem 3.1 Suppose that $f, \sigma : [t_0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ are continuous in both variables and satisfy a global Lipschitz condition with respect to the variable $x \in \mathbb{R}$ uniformly, i.e., there exists a constant $\kappa > 0$ such that

$$|f(t, x_1) - f(t, x_2)|, |\sigma(t, x_1) - \sigma(t, x_2)| \leq \kappa |x_1 - x_2|$$

for all $x_1, x_2 \in \mathbb{R}$ and all $t \in [t_0, T]$. In addition, suppose that the initial condition $X_{t_0}^2$ is non-anticipative w.r.t. the Wiener process W_t with $\mathbb{E}[X_{t_0}^2] < \infty$.

Then the SODE (3.3) has a unique solution $X_t : [t_0, T] \rightarrow \mathbb{R}$.

As for ODEs and RODEs, the proof is by either successive approximations or a fixed point argument in an appropriate function space. A major difference is the Lipschitz condition is used in the term with the Itô integral. Unlike using directly

$$\begin{aligned} \left| \int_{t_0}^t f(s, X_s) ds - \int_{t_0}^t f(s, Y_s) ds \right| &\leq \int_{t_0}^t |f(s, X_s) - f(s, Y_s)| ds \\ &\leq \int_{t_0}^t \kappa |X_s - Y_s| ds \end{aligned}$$

in the drift term, one must first use the Itô isometry for the diffusion term:

$$\begin{aligned} &\mathbb{E} \left[\left| \int_{t_0}^t \sigma(s, X_s) dW_s - \int_{t_0}^t \sigma(s, Y_s) dW_s \right|^2 \right] \\ &= \mathbb{E} \left[\left| \int_{t_0}^t (\sigma(s, X_s) - \sigma(s, Y_s)) dW_s \right|^2 \right] \quad (\text{linearity}) \\ &= \int_{t_0}^t \mathbb{E} [|\sigma(s, X_s) - \sigma(s, Y_s)|^2] ds \quad (\text{Itô isometry}) \\ &\leq \kappa^2 \int_{t_0}^t \mathbb{E} [|X_s - Y_s|^2] ds \quad (\text{Lipschitz condition}) \end{aligned}$$

The above assumptions are rather strong and can be weakened in various ways, e.g., to a local Lipschitz condition, but then with an additional linear growth condition. Alternatively, some kind of dissipativity condition could be assumed.

Remark 3.1 The non-anticipative property of a solution X_t is often expressed in terms of the \mathcal{F}_t -measurability of X_t , where \mathcal{F}_t is a sub- σ -algebra of the σ -algebra \mathcal{F} in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. It belongs to a filtration $\{\mathcal{F}_t, t \geq 0\}$, i.e., a family of increasing sub- σ -algebras of \mathcal{F} with \mathcal{F}_t representing the information provided by the Wiener process W_t up to time t .

3.3 The Itô Formula: The Stochastic Chain Rule

Stochastic Taylor expansions are used to derive systematically higher order numerical schemes for SODE. This is done with an iterated application of the stochastic chain rule, which is usually called the Itô formula. It is often written in differential form, but, as with SODEs, this is only symbolical for its correct representation as an integral equation.

Let $U : [t_0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ be two times continuously differentiable and define $Z_t = U(t, X_t)$ for $t \in [t_0, T]$, where X_t is a solution of the Itô SODE (3.3). Then

$$dZ_t = L^0 U(t, X_t) dt + L^1 U(t, X_t) dW_t$$

or in the integral form

$$Z_t = U(t_0, X_{t_0}) + \int_{t_0}^t L^0 U(s, X_s) ds + \int_{t_0}^t L^1 U(s, X_s) dW_s$$

with the differential operators

$$L^0 U = \frac{\partial U}{\partial t} + f \frac{\partial U}{\partial x} + \frac{1}{2} \sigma^2 \frac{\partial^2 U}{\partial x^2}, \quad L^1 U = \sigma \frac{\partial U}{\partial x}.$$

These expressions are known as the *Itô formula*.

The operator L^0 here has an additional term $\frac{1}{2} \sigma^2 \frac{\partial^2 U}{\partial x^2}$ compared with the deterministic total derivative

$$\frac{d}{dt} U(t, x) = \frac{\partial U}{\partial t}(t, x) + f \frac{\partial U}{\partial x}(t, x)$$

w.r.t. an ODE with vector field f obtained through the deterministic chain rule. The origin of this term lies in the fact that $\mathbb{E}[(\Delta W_t)^2] = \Delta t$, i.e., a second order noise term is (in expectation) equivalent to a first order time term.

3.4 Stratonovich SODEs

A Stratonovich stochastic differential equation

$$dX_t = f(t, X_t) dt + \sigma(t, X_t) \circ dW_t,$$

is a stochastic integral equation

$$X_t = X_{t_0} + \int_{t_0}^t f(s, X_s) ds + \int_{t_0}^t \sigma(s, X_s) \circ dW_s \quad (3.4)$$

with a Stratonovich stochastic integral, which is distinguished from an Itô stochastic integral by the “ \circ ” symbol. It is defined similarly to (3.2), but uses the midpoint $\tau_j^{(N)} := \frac{1}{2}(t_j^{(N)} + t_{j+1}^{(N)})$ of each partition subinterval $(t_j^{(N)}, t_{j+1}^{(N)})$ instead of the lower end point $t_j^{(N)}$, i.e.,

$$\int_0^T \sigma(t, \omega) \circ dW_t(\omega) = \text{ms-} \lim_{N \rightarrow \infty} \sum_{j=0}^{N-1} \sigma(\tau_j^{(N)}, \omega) (W_{t_{j+1}^{(N)}}(\omega) - W_{t_j^{(N)}}(\omega)).$$

The Stratonovich integral does not have vanishing expectation and an analogous isometry property to the Itô integral. The chain rule for Stratonovich stochastic calculus is, however, the same as for deterministic calculus.

Let $U : [t_0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ be continuously differentiable and define $Z_t = U(t, X_t)$ for $t \in [t_0, T]$, where X_t is a solution of the Stratonovich SODE (3.4). Then

$$dZ_t = \tilde{L}^0 U(t, X_t) dt + \tilde{L}^1 U(t, X_t) \circ dW_t$$

or in the integral form

$$Z_t = U(t_0, X_{t_0}) + \int_{t_0}^t \tilde{L}^0 U(s, X_s) ds + \int_{t_0}^t \tilde{L}^1 U(s, X_s) \circ dW_s$$

with the differential operators

$$\tilde{L}^0 U = \frac{\partial U}{\partial t} + f \frac{\partial U}{\partial x}, \quad \tilde{L}^1 U = \sigma \frac{\partial U}{\partial x}.$$

Mathematically speaking, both the Itô and Stratonovich stochastic calculus are correct. Which one should be used in a given context is essentially a modeling issue, understood more heuristically than rigorously.

3.5 Relationship Between RODEs and SODEs

A RODE driven by an Itô process, i.e., the solution of an Itô SODE can be rewritten as a higher dimensional SODE, so results for one can be applied to the other. For example, a scalar RODE

$$\frac{dx}{dt} = g(x, Y_t)$$

driven by the solution Y_t of the scalar SODE

$$dY_t = f(Y_t) dt + \sigma(Y_t) dW_t$$

can be rewritten as the 2-dimensional SODE

$$d \begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} g(X_t, Y_t) \\ f(Y_t) \end{pmatrix} dt + \begin{pmatrix} 0 \\ \sigma(Y_t) \end{pmatrix} dW_t$$

3.5.1 Doss–Sussmann Transformation

Any finite dimensional SODE with regular coefficients can, moreover, be transformed to a RODE. In the case of commutative noise this is the famous Doss–Sussmann result [43, 129], which was generalised to all SODEs in recent years by Imkeller and Schmalfuß [72]. It is easily illustrated for a scalar SODE with additive noise. The equation

$$dX_t = f(X_t) dt + dW_t$$

is equivalent to the RODE

$$\frac{dz}{dt} = f(z + O_t) + O_t, \quad (3.5)$$

where $z(t) := X_t - O_t$ and O_t is the stochastic stationary Ornstein–Uhlenbeck process satisfying the linear SODE

$$dO_t = -O_t dt + dW_t. \quad (3.6)$$

To see this, subtract integral versions of both SODEs (3.5) and (3.6) and substitute to obtain

$$z(t) = z(0) + \int_0^t [f(z(s) + O_s) + O_s] ds.$$

It then follows by continuity and the fundamental theorem of integral and differential calculus that z is pathwise differentiable.

Similarly, for a scalar Itô SODE with linear multiplicative noise

$$dX_t = f(t, X_t) dt + b(t) X_t dW_t$$

the random transformation

$$z(t) = T(t)X_t, \quad T(t) := \exp\left(\frac{1}{2} \int_0^t b^2(s) ds - \int_0^t b(s) dW_s\right)$$

leads to the RODE

$$\frac{dz}{dt} = T(t)f\left(t, T^{-1}(t)z(t)\right),$$

or, more specifically,

$$\frac{d}{dt}z(t) = f\left(t, e^{-\frac{1}{2} \int_0^t b(s)^2 ds + \int_0^t b(s) dW_s} z(t)\right) e^{\frac{1}{2} \int_0^t b(s)^2 ds - \int_0^t b(s) dW_s}.$$

This can be confirmed by the Itô formula.

Remark 3.2 The Doss–Sussmann result and its generalisations are remarkable. They allow solutions of Itô SODEs, which are defined in terms of mean-square calculus, to be transformed to solutions of RODEs, which are defined in the pathwise sense. The theory of rough paths provides a deeper explanation of the relationship between ODEs, RODEs and SODEs.

3.6 Endnotes

There are many books on SODEs [1, 3, 46, 82, 91, 105, 116]. See Duan [44] for further specific examples.

The Doss–Sussmann result [43, 129] was generalised to all SODEs in recent years by Imkeller and Schmalfuß [72]. See also Imkeller and Lederer [70, 71].

Friz and Hairer [48] is a readable introduction to the theory of rough paths.

Chapter 4

Random Dynamical Systems

A random ordinary differential equation (RODE) in \mathbb{R}^d ,

$$\frac{dx}{dt} = g(x, \eta_t(\omega)), \quad x \in \mathbb{R}^d, \quad (4.1)$$

is essentially a nonautonomous ordinary differential equation (ODE), so many results on stability and boundedness for nonautonomous ODEs, in particular those involving Lyapunov functions, can be used for RODEs.

There is an extensive literature on stability results when the vector field satisfies $g(0, y) = 0$ for all y , i.e., when the RODE has a constant solution $\bar{x} = 0$. A constant solution is often called a steady state or equilibrium solution.

The situation can be more complicated. Such constant equilibrium solutions need not exist, but there is a generalisation of an equilibrium solution in random systems which plays an important role in random dynamics. A simple example is given in the next section. Such examples often involve a two-sided Wiener process as their driving noise. A two-sided Wiener process W_t is defined for all $t \in \mathbb{R}$, not just \mathbb{R}^+ . Essentially, W_t and W_{-t} for $t \in \mathbb{R}^+$ are two independent conventional Wiener processes.

4.1 Nontrivial Equilibrium Solutions

Let W_t be a two-sided Wiener process and consider the simple scalar RODE with bounded noise

$$\frac{dx}{dt} = -x + \cos W_t(\omega), \quad (4.2)$$

which obviously has no constant solution. The RODE (4.2) with an initial value $x_0 \in \mathbb{R}^1$ at the initial time $t_0 \in \mathbb{R}$ has the explicit solution

$$x(t, t_0, x_0, \omega) = x_0 e^{-(t-t_0)} + e^{-t} \int_{t_0}^t e^s \cos W_s(\omega) ds, \quad (4.3)$$

which has no pathwise limit as $t \rightarrow \infty$. On the other hand, the difference between any two solutions $x(t, t_0, x_0, \omega)$ and $y(t, t_0, y_0, \omega)$ satisfies

$$|x(t, t_0, x_0, \omega) - y(t, t_0, y_0, \omega)| \leq |x_0 - y_0| e^{-(t-t_0)} \rightarrow 0 \quad \text{as } t \rightarrow \infty, \quad (4.4)$$

so all solutions pathwise converge to each other. What do they converge to?

The pullback limit, i.e., for $t_0 \rightarrow -\infty$ with t held fixed, of the solution (4.3) exists and is given by

$$\lim_{t_0 \rightarrow -\infty} x(t, \omega) = \bar{x}(t, \omega) := e^{-t} \int_{-\infty}^t e^s \cos W_s(\omega) ds. \quad (4.5)$$

Moreover, $\bar{x}(t, \omega)$ is itself a solution of the RODE (4.2), so the inequality (4.4) applies and gives

$$|x(t, t_0, x_0, \omega) - \bar{x}(t, \omega)| \leq |x_0 - \bar{x}(t_0, \omega)| e^{-(t-t_0)} \rightarrow 0 \quad \text{as } t \rightarrow \infty,$$

which means that the pullback limit solution (4.5) pathwise attracts all other solutions. It is, in fact, pathwise globally asymptotically stable with an exponential rate of attraction. It is the counterpart of a steady state or equilibrium solution in autonomous ODE and is often called a *random equilibrium*.

The difference between forward and pullback attraction is illustrated in Figs. 4.1, 4.2. In both figures the solutions converge to the solid (red) solution. In Fig. 4.1 the convergence is in the forward sense as $t \rightarrow \infty$ with all solutions starting at a fixed time t_0 , whereas in Fig. 4.2 the convergence is in the pullback sense at a fixed time t as the starting time $t_0 \rightarrow -\infty$.

Essentially, pullback convergence uses information about the system in the past, whereas forward convergence uses information about it in the future. The two concepts are independent of each other.

Fig. 4.1 Forward attraction
of solutions with t_0 fixed and
 $t \rightarrow \infty$

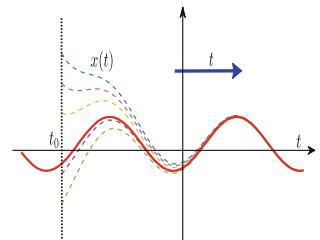
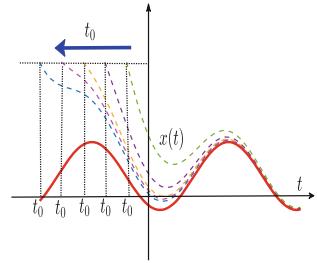


Fig. 4.2 Pullback attraction of solutions with t fixed and $t_0 \rightarrow -\infty$



The Equilibrium Solution of a Nonlinear RODE

Consider now a scalar nonlinear RODE with bounded noise

$$\frac{dx}{dt} = f(x) + \cos W_t(\omega), \quad (4.6)$$

where W_t is a two-sided Wiener process and suppose that f is continuously differentiable and satisfies the *dissipative one-sided Lipschitz condition*,¹ i.e.,

$$\langle x - y, f(x) - f(y) \rangle \leq -\kappa |x - y|^2 \quad \text{for all } x, y \in \mathbb{R}^1. \quad (4.7)$$

Examples of such a function are $-x$, $-x - x^3$ and $-x - x^3 + 1$. The corresponding autonomous ODE, i.e., (4.6), has a unique steady state solution, which is globally asymptotically stable.

The simple analysis for the linear RODE (4.2) cannot be used here since the explicit solution is not known. Nevertheless, a similar situation holds. In view of condition (4.7), the difference between any two solutions of the nonlinear RODE (4.6) satisfies

$$\begin{aligned} \frac{d}{dt} |x(t, \omega) - y(t, \omega)|^2 &= 2 \langle x(t, \omega) - y(t, \omega), \frac{dx}{dt} - \frac{dy}{dt} \rangle \\ &= 2 \langle x(t, \omega) - y(t, \omega), f(x(t, \omega)) - f(y(t, \omega)) \rangle \\ &\leq -2\kappa |x(t, \omega) - y(t, \omega)|^2, \end{aligned}$$

from which it follows that

$$|x(t, \omega) - y(t, \omega)| \leq |x(t_0, \omega) - y(t_0, \omega)| e^{-\kappa(t-t_0)} \rightarrow 0 \quad \text{as } t \rightarrow \infty. \quad (4.8)$$

Hence all solutions converge pathwise to each other as $t \rightarrow \infty$.

As in the simple case above, the pullback limit will be used to construct a limiting solution to which all other solutions converge. This requires a compactness argument,

¹The concept is useful in higher dimensions and is needed later, so the general scalar product $\langle \cdot, \cdot \rangle$ on \mathbb{R}^d is used here.

in particular that the solutions are absorbed (in the pullback sense) into a compact set in a finite time.

Write the pullback limiting solution of the linear RODE (4.2) as $\bar{x}(t, \omega)$ and subtract it from a solution $x(t, \omega)$ of the nonlinear RODE (4.6). Define $z(t, \omega) := x(t, \omega) - \bar{x}(t, \omega)$. Then

$$\frac{dz}{dt} = f(z + \bar{x}(t, \omega)) + \bar{x}(t, \omega). \quad (4.9)$$

Now take the inner product of (4.9) with $z(t)$ and apply the one-sided Lipschitz condition to the resulting RODE to obtain

$$\begin{aligned} \frac{d}{dt}|z(t)|^2 &= 2\langle z(t), \frac{dz}{dt} \rangle = 2\langle z(t), f(z + \bar{x}(t, \omega)) + \bar{x}(t, \omega) \rangle \\ &= 2\langle z(t), f(z(t) + \bar{x}(t, \omega)) - f(\bar{x}(t, \omega)) \rangle + 2\langle z(t), f(\bar{x}(t, \omega)) + \bar{x}(t, \omega) \rangle \\ &\leq -2\kappa|z(t)|^2 + \kappa|z(t)|^2 + \frac{4}{\kappa}|f(\bar{x}(t, \omega)) + \bar{x}(t, \omega)|^2. \\ &\leq -\kappa|z(t)|^2 + \frac{4}{\kappa}|f(\bar{x}(t, \omega)) + \bar{x}(t, \omega)|^2. \end{aligned}$$

Hence,

$$|z(t)|^2 \leq |z(t_0)|^2 e^{-\kappa(t-t_0)} + \frac{4e^{-\kappa t}}{\kappa} \int_{t_0}^t e^{\kappa s} |f(\bar{x}(s, \omega)) + \bar{x}(s, \omega)|^2 ds. \quad (4.10)$$

Taking $t_0 \rightarrow -\infty$ with t held fixed, inequality (4.10) then gives the pathwise pullback absorption

$$|z(t, \omega)|^2 \leq r^2(t, \omega) := 1 + \frac{4e^{-\kappa t}}{\kappa} \int_{-\infty}^t e^{\kappa s} |f(\bar{x}(s, \omega)) + \bar{x}(s, \omega)|^2 ds \quad (4.11)$$

for all $t_0 \leq t - T_D$, where $T_D > 0$ depends on the arbitrary bounded subsets D of initial conditions. Thus,

$$|x(t, \omega) - \bar{x}(t, \omega)| \leq r(t, \omega) \quad \text{for all } t_0 \leq t - T_D,$$

which implies that

$$|x(t, \omega)| \leq |\bar{x}(t, \omega)| + r(t, \omega) \quad \text{for all } t_0 \leq t - T_D.$$

The family of compact balls $B(t, \omega)$ centered on $\bar{x}(t, \omega)$ with radius $r(t, \omega)$ is thus pullback absorbing. This means that it absorbs the dynamics in a finite time provided that the starting time is sufficiently far in the past, see Definition 4.5. Moreover, it follows from (4.5) that the balls $B(t, \omega)$ are positively invariant in the sense that

$$x(t, t_0, B(t_0, \omega), \omega) \subset B(t, \omega) \quad \text{for all } t \geq t_0.$$

Note that $|\bar{x}(t, \omega)| = \left| e^{-t} \int_{-\infty}^t e^s \cos W_s(\omega) ds \right| \leq 1$ and f is continuous, so there exists $\hat{r} \geq 0$ such that $|f(\bar{x}(t, \omega))| \leq \hat{r}$ for all $t \in \mathbb{R}$. And therefore by (4.11) there exists an $r_\infty(\omega)$ such that $r(t, \omega) \leq r_\infty(\omega)$ for all $t \in \mathbb{R}$, i.e., the balls $B(t, \omega)$ are uniformly bounded.

Theorem 4.1 *For all $t \in \mathbb{R}$ and $\omega \in \Omega$, $\cap_{t_0 \leq t} x(t, t_0, B(t_0, \omega), \omega)$ is a singleton set, i.e.,*

$$\bigcap_{t_0 \leq t} x(t, t_0, B(t_0, \omega), \omega) = \{\hat{x}(t, \omega)\},$$

where, $\hat{x}(t, \omega)$ is an entire solution of the RODE (4.6), i.e., a solution defined for all $t \in \mathbb{R}$ for each $\omega \in \Omega$.

Proof The sets $x(t, t_0, B(t_0, \omega), \omega)$ are nonempty and compact so their intersection is nonempty. Suppose (for contradiction) that this intersection contains two distinct points $\hat{x}(t, \omega)$ and $\hat{y}(t, \omega)$ with

$$|\hat{x}(t, \omega) - \hat{y}(t, \omega)| \geq \varepsilon_0 \tag{4.12}$$

for some $\varepsilon_0 > 0$. By the positive invariance of the absorbing sets, for any sequence $t_n \rightarrow -\infty$ there exist $x_n, y_n \in B(t_n, \omega)$ and corresponding solutions $x(t, t_n, x_n, \omega)$, $y(t, t_n, y_n, \omega)$ of the RODE (4.6) such that

$$x(t, t_n, x_n, \omega) = \hat{x}(t, \omega), \quad y(t, t_n, y_n, \omega) = \hat{y}(t, \omega).$$

By the contraction property (4.8) (which is a consequence of the one-sided Lipschitz condition)

$$\begin{aligned} |\hat{x}(t, \omega) - \hat{y}(t, \omega)| &= |x(t, t_n, x_n, \omega) - y(t, t_n, y_n, \omega)| \\ &\leq |x_n - y_n| e^{-\kappa(t-t_n)} \\ &\leq 2r_\infty(\omega) e^{-\kappa(t-t_n)} \leq \frac{1}{2}\varepsilon_0 \end{aligned}$$

for n large enough, i.e., t_n negative enough. This contradicts the assumption (4.12), and hence the intersection consists of a single point.

The proof that $\hat{x}(t, \omega)$ is a solution of the RODE (4.6) is left to the reader. \square

Since $\hat{x}(t, \omega)$ is a solution of the RODE (4.6), then by using the contraction property (4.8) again we have

$$\begin{aligned} |x(t, t_0, x_0, \omega) - \hat{x}(t, \omega)| &\leq |x_0 - \hat{x}(t_0, \omega)| e^{-\kappa(t-t_0)} \\ &\leq 2r_\infty(\omega) e^{-\kappa(t-t_0)} \rightarrow 0 \quad \text{as } t \rightarrow \infty. \end{aligned}$$

This means that $\hat{x}(t, \omega)$ is pathwise globally asymptotically stable. It is the random equilibrium of this RODE and a simple example of a random attractor.

4.2 Random Dynamical Systems

RODEs often generate random dynamical systems. These and their attractors, called random attractors, are briefly introduced here for the state space \mathbb{R}^d . For the general theory on more general state spaces see Arnold [4]. In this theory the RODE is written in the canonical form

$$\frac{dx}{dt} = g(x, \theta_t(\omega)), \quad x \in \mathbb{R}^d, \quad (4.13)$$

with a noise represented by a measure-preserving dynamical system $\theta = \{\theta_t\}_{t \in \mathbb{R}}$ acting on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ rather than as a specific noise process η_t as in (4.1).

Definition 4.1 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A random dynamical system (θ, φ) on \mathbb{R}^d consists of an autonomous measurable and measure-preserving dynamical system $\theta = \{\theta_t\}_{t \in \mathbb{R}}$ acting on $(\Omega, \mathcal{F}, \mathbb{P})$ satisfying

- (i) $\theta_0(\omega) = \omega$ for all $\omega \in \Omega$;
- (ii) $\theta_{s+t}(\omega) = \theta_s \circ \theta_t(\omega)$ for all $\omega \in \Omega$ and any $s, t \in \mathbb{R}$;
- (iii) $(t, \omega) \mapsto \theta_t(\omega)$ is measurable for all $\omega \in \Omega$ and any $s, t \in \mathbb{R}$;
- (iv) $\theta_t \mathbb{P} = \mathbb{P}$ for every $t \in \mathbb{R}$,

and a cocycle mapping $\varphi : \mathbb{R}^+ \times \Omega \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ satisfying

- (i) initial condition: $\varphi(0, \omega, x) = x$ for all $\omega \in \Omega$ and $x \in \mathbb{R}^d$,
- (ii) cocycle property: $\varphi(s + t, \omega, x) = \varphi(s, \theta_t(\omega), \varphi(t, \omega, x))$ for all $s, t \in \mathbb{R}^+$, $\omega \in \Omega$ and $x \in \mathbb{R}^d$,
- (iii) measurability: $(t, \omega, x) \mapsto \varphi(t, \omega, x)$ is measurable,
- (iv) continuity: $x \mapsto \varphi(t, \omega, x)$ is continuous for all $(t, \omega) \in \mathbb{R} \times \Omega$.

Measurability refers to joint measurability with respect to the Borel σ -algebras on \mathbb{R} or \mathbb{R}^+ and \mathbb{R}^d , and \mathcal{F} . The notation $\theta_t(\mathbb{P}) = \mathbb{P}$ for the measure preserving property of θ_t with respect to \mathbb{P} is just a compact way of writing

$$\mathbb{P}(\theta_t(A)) = \mathbb{P}(A) \quad \text{for all } t \in \mathbb{R}, A \in \mathcal{F}.$$

If the random dynamical system is generated by a RODE, the cocycle mapping is the solution mapping corresponding to the initial value $x_0 \in \mathbb{R}^d$ at time $t = 0$ with the noise in the state $\omega \in \Omega$.

4.2.1 Random Attractors

Random attractors and their deterministic nonautonomous counterparts consist of families of sets that are mapped onto each other by the cocycle mapping. For random attractors the component sets are labeled by $\omega \in \Omega$. Another way of expressing this is in terms of random sets.

Definition 4.2 A family $\mathcal{D} = \{D(\omega) : \omega \in \Omega\}$ of nonempty subsets of \mathbb{R}^d is called a random set if for each $x \in \mathbb{R}^d$ the mapping $\omega \mapsto \text{dist}_{\mathbb{R}^d}(x, D(\omega))$ is \mathcal{F} -measurable, where

$$\text{dist}_{\mathbb{R}^d}(x, Y) = \inf_{y \in Y} |x - y|$$

is the distance between the point x and the subset Y of \mathbb{R}^d .

Definition 4.3 A random set $\mathcal{D} = \{D(\omega) : \omega \in \Omega\}$ is called a random closed set if $D(\omega)$ is closed for each $\omega \in \Omega$. A random set $\mathcal{D} = \{D(\omega) : \omega \in \Omega\}$ is called a random compact set if $D(\omega)$ is compact for each $\omega \in \Omega$. A random bounded set $\mathcal{D} = \{D(\omega) : \omega \in \Omega\}$ is called tempered if its growth with respect to the driving system θ is sub-exponential, i.e., there exists an $x_0 \in \mathbb{R}^d$ such that

$$D(\omega) \subset \{x \in \mathbb{R}^d : |x - x_0| \leq r(\omega)\}, \quad \omega \in \Omega,$$

where the random variable $r(\omega) > 0$ is tempered, i.e.,

$$\lim_{t \rightarrow \infty} e^{-ct} \sup_{t \in \mathbb{R}} |r(\theta_{-t}(\omega))| = 0, \quad \omega \in \Omega, \quad \forall c > 0.$$

Denote by \mathbb{T} (*tiān*) the collection of all tempered random sets in \mathbb{R}^d .

A random attractor of a random dynamical system is a random set which is a pullback attractor in the pathwise sense with respect to the attracting basin of tempered random sets, which is defined as follows.

Definition 4.4 A random set $\mathcal{A} = (A(\omega))_{\omega \in \Omega}$ from \mathbb{T} is called a random \mathbb{T} attractor (pullback \mathbb{T} attractor or random attractor) for a random dynamical system (θ, φ) on \mathbb{R}^d if

- (i) \mathcal{A} is a random compact set;
- (ii) \mathcal{A} is a φ -invariant set, i.e., $\varphi(t, \omega, A(\omega)) = A(\theta_t(\omega))$ for a.e. $\omega \in \Omega$ and all $t \geq 0$;
- (iii) \mathcal{A} is pathwise pullback attracting in \mathbb{T} , i.e.,

$$\lim_{t \rightarrow \infty} \text{dist}_{\mathbb{R}^d} (\varphi(t, \theta_{-t}(\omega), D(\theta_{-t}(\omega))), A(\omega)) = 0$$

for a.e. $\omega \in \Omega$ and all $\mathcal{D} \in \mathbb{T}$.

The existence of a random attractor is ensured by the existence of a pullback absorbing set, which is defined as follows.

Definition 4.5 A tempered random set $\mathcal{K} = \{K(\omega) : \omega \in \Omega\}$ is called an absorbing set in \mathbb{X} , if for a.e. $\omega \in \Omega$ and all $D \in \mathbb{X}$, there exists $T_D(\omega) > 0$ such that

$$\varphi(t, \theta_{-t}(\omega), D(\theta_{-t}(\omega))) \subset K(\omega), \quad \text{for all } t \geq T_D(\omega).$$

Theorem 4.2 Let (θ, φ) be a random dynamical system on \mathbb{R}^d . If there exists an absorbing set $\mathcal{K} = \{K(\omega) : \omega \in \Omega\}$ with compact component sets, then the random dynamical system (θ, φ) has a random attractor $\mathcal{A} = \{A(\omega) : \omega \in \Omega\}$ with component sets defined by

$$A(\omega) = \overline{\bigcup_{\tau > T_D(\omega)} \varphi(t, \theta_{-t}(\omega), K(\theta_{-t}(\omega)))}.$$

Note that there is no analogous result for the existence of a random attractor with pathwise forward convergence. However, the θ_t -invariance of the probability measure \mathbb{P} implies that

$$\begin{aligned} \mathbb{P}\{\omega \in \Omega : \text{dist}(\varphi(t, \theta_{-t}(\omega), K(\theta_{-t}(\omega))), A(\omega)) \geq \varepsilon\} \\ = \mathbb{P}\{\omega \in \Omega : \text{dist}(\varphi(t, \omega, K(\omega)), A(\theta_t(\omega))) \geq \varepsilon\} \end{aligned}$$

for any $\varepsilon > 0$. Since \mathbb{P} -almost sure convergence implies convergence in probability (see Appendix A), a random pullback attractor also converges in the forwards sense, but only in the weaker sense of convergence in probability. This allows individual sample path to have large deviations from the attractor, but still to converge in this probabilistic sense.

An Example

As an illustration of Theorem 4.2 consider a particular class of RODE (4.13) on \mathbb{R}^d for which the vector field has the form $g(x, \theta_t(\omega)) = \mu(x) + \zeta(\theta_t(\omega))$:

$$\frac{dx}{dt} = \mu(x) + \zeta(\theta_t(\omega)) \quad x(0) = x_0. \quad (4.14)$$

Assume that

- (i) the mapping $\mu: \mathbb{R}^d \rightarrow \mathbb{R}^d$ is continuously differentiable and satisfies the generalized linear growth bound

$$\langle \mu(x), x \rangle \leq -M\|x\|^2 + N^2, \quad x \in \mathbb{R}^d \quad (4.15)$$

for some $M > 0$ and $N \in \mathbb{R}$;

- (ii) the stationary random process $t \mapsto \zeta(\theta_t(\omega))$ is tempered, i.e., its paths satisfy a sub-exponential growth condition

$$\lim_{|t| \rightarrow \infty} e^{-c|t|} \|\zeta(\theta_t(\omega))\| = 0 \quad (4.16)$$

for any $\omega \in \Omega$ and $c > 0$.

Remark 4.1 A sufficient condition for (4.16) is given by [3]

$$\mathbb{E} \left[\sup_{t \in [0,1]} \log^+ \|\zeta(\theta_t(\omega))\| \right] < \infty.$$

These structural assumptions ensure that the RODE (4.14) has a unique global solution $\varphi(t, \omega, x_0)$ and that (θ, φ) generates an RDS on \mathbb{R}^d .

Theorem 4.3 *The RDS (θ, φ) generated by the RODE (4.14) has a global random attractor.*

Proof Taking inner product of equation (4.14) with $x(t)$ and using assumption (4.15) gives

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|x(t)\|^2 &= \langle \mu(x), x \rangle + \langle \zeta(\theta_t(\omega)), x \rangle \\ &\leq -M\|x\|^2 + N^2 + \frac{M}{2}\|x\|^2 + \frac{1}{2M}\|\zeta(\theta_t(\omega))\|^2, \end{aligned}$$

or equivalently,

$$\frac{d}{dt} \|x(t)\|^2 \leq -M\|x\|^2 + 2N^2 + \frac{1}{M}\|\zeta(\theta_t(\omega))\|^2. \quad (4.17)$$

Integrating (4.17) from 0 to t gives

$$\|\varphi(t, \omega, x_0)\|^2 \leq e^{-Mt} \|x_0\|^2 + \frac{2N^2}{M} + \frac{1}{M} \int_0^t e^{-M(t-s)} \|\zeta(\theta_s(\omega))\|^2 ds.$$

Starting the noise at $\theta_{-t}(\omega)$ instead of ω results in

$$\begin{aligned} \|\varphi(t, \theta_{-t}(\omega), x_0)\|^2 &\leq e^{-Mt} \|x_0\|^2 + \frac{2N^2}{M} + \frac{1}{M} \int_0^t e^{-M(t-s)} \|\zeta(\theta_{s-t}(\omega))\|^2 ds \\ &= e^{-Mt} \|x_0\|^2 + \frac{2N^2}{M} + \frac{1}{M} \int_{-t}^0 e^{M\tau} \|\zeta(\theta_\tau(\omega))\|^2 d\tau. \end{aligned}$$

Let

$$r(\omega) := \left(\frac{2N^2}{M} + \frac{1}{M} \int_{-\infty}^0 e^{M\tau} \|\zeta(\theta_\tau(\omega))\|^2 d\tau \right)^{1/2}.$$

Then for any $x_0 \in D(\theta_{-t}(\omega))$ with $\mathcal{D} = \{D(\theta_{-t}\omega) : \omega \in \Omega\} \in \mathcal{X}$,

$$\|\varphi(t, \theta_{-t}(\omega), x_0)\|^2 \leq e^{-Mt} \sup_{x \in D(\theta_{-t}(\omega))} \|x\|^2 + r^2(\omega),$$

and

$$\lim_{t \rightarrow \infty} e^{-Mt} \sup_{x \in D(\theta_{-t}(\omega))} \|x\|^2 = 0.$$

In addition, due to assumption (4.16), for any $c > 0$

$$\begin{aligned} (e^{-ct} r(\theta_{-t}(\omega)))^2 &= e^{-2ct} \frac{2N^2}{M} + \frac{e^{-2ct}}{M} \int_{-\infty}^0 e^{M\tau} \|\zeta(\theta_{\tau-t}(\omega))\|^2 d\tau \\ &= e^{-2ct} \frac{2N^2}{M} + \frac{e^{-2ct}}{M} \int_{-\infty}^{-t} \|\zeta(\theta_s(\omega))\|^2 ds \\ &\rightarrow 0 \quad \text{as } t \rightarrow \infty, \end{aligned}$$

i.e., $r(\omega)$ is tempered. Define

$$K(\omega) := \{x \in \mathbb{R}^d : \|x\|^2 \leq 1 + r^2(\omega)\}.$$

Then $\mathcal{K} = \{K(\omega) : \omega \in \Omega\}$ is a compact \mathcal{F} -absorbing set for the RDS (θ, φ) . The existence of a random attractor $\mathcal{A} = (A(\omega))_{\omega \in \Omega}$ with $A(\omega) \subset K(\omega)$ for each $\omega \in \Omega$ follows immediately by Theorem 4.2. \square

4.2.2 Contractive Cocycles

Theorem 4.2 can be strengthened when the cocycle mapping is strictly contracting, such as in the nonlinear RODE (4.9) that satisfies a one-sided dissipative Lipschitz condition. Theorem 4.1 implies that the resulting random attractor consists of singleton sets, i.e., $A(\omega) = \{x^*(\omega)\}$ for some random variable x^* with $x^*(\omega) \in \mathbb{R}^d$. Essentially, it corresponds to a stochastic process $\bar{x}_t(\omega) := x^*(\theta_t(\omega))$.

More precisely, when the cocycle mapping is strictly uniformly contracting [25], i.e., there exists $\kappa > 0$ such that

$$|\varphi(t, \omega, x_0) - \varphi(t, \omega, y_0)| \leq e^{-\kappa t} |x_0 - y_0|, \quad \forall t \geq 0, \omega \in \Omega, x_0, y_0 \in \mathbb{R}^d, \tag{4.18}$$

the random attractor \mathcal{A} consists of a singleton subsets $A(\omega) = \{x^*(\omega)\}$, i.e., is a single stochastic process. This can be proved using a Cauchy sequence rather than compactness arguments.

In this case, the random attractor is pathwise attracting in both the pullback sense and the forward sense. The examples in Sect. 4.1 generate contractive cocycles. A more general situation in the context of random Markov chains will be considered in Chap. 18.

4.3 Endnotes

Lyapunov stability, asymptotic stability and attractors for autonomous ODEs are discussed in many books on differential equations or dynamical systems. Two of particular interest here are Stuart and Humphries [128], which investigates the effects of discretisation on such properties, and Kloeden and Rasmussen [92], which generalises the concepts to nonautonomous and random dynamical systems.

Lyapunov stability results for RODEs are given in Bunke [20] and Neckel and Rupp [108]. See Arnold [4] for linear systems and Lyapunov exponents, and Kloeden and Rasmussen [92] for deterministic nonautonomous dynamical systems, where pullback convergence is used extensively.

Khasminskii [83] discussed the stability of RODEs and provided references on them to the Russian language up to 1967, when the first edition of his book appeared. Kac and Krasovski [143] investigated the stability of RODEs driven by a finite Markov chain, while Arnold and Kloeden [5] determined explicit expressions for Lyapunov exponents of 2-dimensional RODEs driven by telegraphic noise.

The monograph Arnold [4] is devoted to random dynamical systems and random attractors. See Crauel and Kloeden [34] for a recent review article. Contractive cocycles satisfying (4.18) were considered in Kloeden and Kozyakin [85, 88]. Random periodic solutions were studied by Zhao and Zheng [144]. See also Caraballo and Han [24], Caraballo, Kloeden and Schmalfuß [25] and Kloeden and Lorenz [86].

The Chinese character 天 used for the collection of all tempered random sets means “sky” and is pronounced *tiān*. The connection is phonetical, since it is pronounced something like the first part of tempered, by Chinese at least.

The effects of discretisation on the dynamical behaviour of random dynamical systems generated by RODEs were investigated by Arnold and Kloeden [5] and Caraballo, Kloeden and Real [26]. See also Chap. 18.

Chapter 5

Numerical Dynamics

Numerical dynamics is concerned with the relationship between the dynamical behavior of the solutions of a differential equation and that of the solutions of a numerical scheme. Two major issues, the preservation of an attractor and a hyperbolic neighbourhood under discretisation are considered here in the context of RODEs and the random dynamical systems.

RODEs formulated with a noise process often generate random dynamical systems, see Arnold [4]. In this theory the RODE is written in the form

$$\frac{dx}{dt} = g(x, \theta_t(\omega)), \quad x \in \mathbb{R}^d,$$

with a canonical noise represented by a measure-preserving dynamical system $\theta = \{\theta_t\}_{t \in \mathbb{R}}$ acting on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ rather than as a specific noise process η_t as in (4.1). Throughout this chapter it will be assumed that the sample paths of the driving noise process are Hölder continuous.

5.1 Discretisation of Random Attractors

In Chap. 4 the existence of a random attractor $\mathcal{A} = \{A(\omega) : \omega \in \Omega\}$ with respect to the attracting universe \mathbb{X} consisting of families $\mathcal{D} = \{D(\omega) : \omega \in \Omega\}$ of nonempty closed and bounded measurable sets \mathbb{R}^d with sub-exponential growth was established for the random dynamical system (θ, φ) generated by the RODE with the structure

$$\frac{dx}{dt} = g(x, \theta_t(\omega)) := \mu(x) + \zeta(\theta_t(\omega)), \quad (5.1)$$

where the mapping $\mu : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is continuously differentiable and satisfies the generalised linear growth bound

$$\langle \mu(x), x \rangle \leq -M\|x\|^2 + N^2, \quad x \in \mathbb{R}^d$$

for some $M > 0$. In addition, the paths $t \mapsto \zeta(\theta_t(\omega))$ of the stationary stochastic process $\zeta : \Omega \rightarrow \mathbb{R}^d$ satisfy a sub-exponential growth condition, i.e.,

$$\lim_{|t| \rightarrow \infty} e^{-c|t|} \|\zeta(\theta_t(\omega))\| = 0$$

for any $\omega \in \Omega$ and $c > 0$.

The implicit Euler numerical scheme with constant step size $h > 0$ applied to the RODE (5.1) is given by

$$x_{n+1} = x_n + h [\mu(x_{n+1}) + \zeta(\theta_{nh}(\omega))]. \quad (5.2)$$

It can be shown for sufficiently small step size h that the implicit equation (5.2) can be solved uniquely to give an explicit expression

$$x_{n+1} = G_h(x_n, \theta_{nh}(\omega)).$$

This generates a discrete time RDS (θ, ψ_h) on $\Omega \times \mathbb{R}^d$ with cocycle mapping ψ_h defined by

$$\psi_h(n, \omega, x_0) := G_h(\cdot, \theta_{(n-1)h}(\omega)) \circ \cdots \circ G_h(x_0, \omega)$$

for $n \geq 2$ with

$$\psi_h(0, \omega, x_0) := x_0, \quad \psi_h(1, \omega, x_0) := G_h(x_0, \omega).$$

In much the same way as in Theorem 4.3 of Chap. 4 for the RODE (5.1) it can be shown that the numerical RDS (θ, ψ_h) also has a random attractor. In particular,

$$\begin{aligned} \|x_{n+1}\|^2 &= \langle x_{n+1}, x_n + h\mu(x_{n+1}) + h\zeta(\theta_{nh}(\omega)) \rangle \\ &= \langle x_{n+1}, x_n \rangle + h\langle x_{n+1}, \mu(x_{n+1}) \rangle + h\langle x_{n+1}, \zeta(\theta_{nh}(\omega)) \rangle \\ &\leq \frac{1}{2} \|x_{n+1}\|^2 + \frac{1}{2} \|x_n\|^2 + hN^2 - hM \|x_{n+1}\|^2 \\ &\quad + \frac{1}{2} hM \|x_{n+1}\|^2 + \frac{2h}{M} \|\zeta(\theta_{nh}(\omega))\|^2, \end{aligned}$$

so

$$\|x_{n+1}\|^2 \leq \frac{1}{1+hM} \|x_n\|^2 + \frac{2hN^2}{1+hM} + \frac{4h}{M(1+hM)} \|\zeta(\theta_{nh}(\omega))\|^2.$$

Writing $\lambda := \frac{1}{1+hM}$ and iterating from $\theta_{-nh}(\omega)$ instead of from ω gives

$$\begin{aligned}\|x_n\|^2 &\leq \lambda^n \|x_0\|^2 + \frac{2hN^2\lambda}{1-\lambda} + \frac{4h\lambda}{M} \sum_{j=0}^{\infty} \lambda^{n-1-j} \|\zeta(\theta_{-(j+1)h}(\omega))\|^2, \\ &\leq \lambda^n \|x_0\|^2 + \frac{2N^2}{M} + \frac{4h}{M} \sum_{j=0}^{\infty} \lambda^{n-1-j} \|\zeta(\theta_{-(j+1)h}(\omega))\|^2.\end{aligned}$$

Thus the family \mathcal{K}_h of closed balls $K_h(\omega)$ in \mathbb{R}^d with center 0 and radii

$$r_h(\omega) = \left(1 + \frac{2N^2}{M} + \frac{4h}{M} \sum_{j=0}^{\infty} \lambda^{n-1-j} \|\zeta(\theta_{-(j+1)h}(\omega))\|^2 \right)^{\frac{1}{2}},$$

is a pullback absorbing set for the numerical RDS (θ, ψ_h) with respect to the families $\mathcal{D} = \{D(\omega) : \omega \in \Omega\}$ of nonempty closed and bounded tempered measurable sets \mathbb{R}^d . Moreover, $\mathcal{K}_h \in \mathcal{D}$, i.e., is also tempered, since $\lambda^n \approx e^{-nhM}$ for very small h .

Recall that $\text{dist}_{\mathbb{R}^d}$ is the Hausdorff semi-distance between subsets of \mathbb{R}^d (e.g., Definition 4.2) is defined by

$$\text{dist}_{\mathbb{R}^d}(A, B) := \sup_{a \in A} \inf_{b \in B} \|a - b\| = \sup_{a \in A} \text{dist}_{\mathbb{R}^d}(a, B),$$

where $\text{dist}_{\mathbb{R}^d}(a, B) := \inf_{b \in B} \|a - b\|$. An application of Theorem 4.2 then yields

Theorem 5.1 *Under the assumptions of Theorem 4.2, the numerical RDS (θ, ψ_h) generated by the implicit Euler scheme applied to the RODE (5.1) with step size $h > 0$ has a random attractor $\mathcal{A}_h = \{A_h(\omega) : \omega \in \Omega\}$ for h sufficiently small.*

Moreover, the component subsets of the numerical random attractor converges upper semi continuously to those of the random attractor generated by the RODE, i.e.,

$$\lim_{h \rightarrow 0} \text{dist}_{\mathbb{R}^d}(A_h(\omega), A(\omega)) = 0 \tag{5.3}$$

for each $\omega \in \Omega$.

Proof From the construction of random attractor component sets $A_h(\omega) \subset K_h(\omega)$, it can be shown that $A_h(\omega)$ is contained in a ball $B(0, \rho(\omega))$ in \mathbb{R}^d with center 0 and a tempered radius $\rho(\omega)$.

Suppose (for contradiction) that the upper semi continuous convergence assertion (5.3) of the theorem is not true. Then there exist an $\varepsilon_0 > 0$, an $\omega \in \Omega$, a subsequence $h_n \rightarrow 0$ as $n \rightarrow \infty$ and points $a_n \in A_{h_n}(\omega)$ such that

$$\text{dist}_{\mathbb{R}^d}(A_{h_n}(\omega), A(\omega)) = \sup_{a_n \in A_{h_n}} \text{dist}_{\mathbb{R}^d}(a_n, A(\omega)) > \varepsilon_0 \quad (5.4)$$

for all n .

The family of balls $\{B(0, \rho(\omega)) : \omega \in \Omega\}$ is in the attracting universe \mathcal{T} of \mathcal{A} , so there is a T_0 such that

$$\text{dist}_{\mathbb{R}^d}(\varphi(t, \theta_{-t}(\omega), B(0, \rho(\theta_{-t}(\omega))), A(\omega)) < \frac{1}{4}\varepsilon_0, \quad t \geq T_0.$$

The global discretisation error of the implicit Euler scheme (5.2) on an interval of length T_0 for the driving system starting at $\xi \in B(0, \rho(\theta_{-T_0}(\omega)))$ is

$$\|\psi_h(j, \theta_{-T_0}(\omega), \xi) - \varphi(jh, \theta_{-T_0}(\omega), \xi)\| \leq C_{T_0}(\omega)h^q, \quad 0 \leq j \leq T_0/h,$$

where q is determined by the Hölder continuity exponent of the driving noise. Define $h_0^* = [\varepsilon_0/(4C_{T_0}(\omega))]^{1/q}$ and pick $h \leq h_0^*$ and N_h such that $N_h h \leq T_0$. Then $C_{T_0}(\omega)h^q \leq \varepsilon_0/4$ and

$$\|\psi_h(j, \theta_{-T_0}(\omega), \xi) - \varphi(jh, \theta_{-T_0}(\omega), \xi)\| \leq C_{T_0}h^q \leq \varepsilon_0/4, \quad 0 \leq j \leq T_0/h.$$

Pick and fix $h_n \leq h_0^*$ and suppose for convenience that $N_{h_n}h_n = T_0$. By the invariance property of the random attractor \mathcal{A}_{h_n} there is a $\xi_n \in A_{h_n}(\theta_{-T_0}(\omega)) \subset B(0, \rho(\theta_{-T_0}(\omega)))$ such that $\psi_{h_n}(N_{h_n}, \theta_{-T_0}(\omega), \xi_n) = a_n$. Thus for any $a_n \in A_{h_n}$,

$$\begin{aligned} \text{dist}_{\mathbb{R}^d}(a_n, A(\omega)) &= \text{dist}_{\mathbb{R}^d}(\psi_{h_n}(N_{h_n}, \theta_{-T_0}(\omega), \xi_n), A(\omega)) \\ &\leq \|\psi_{h_n}(N_{h_n}, \theta_{-T_0}(\omega), \xi_n) - \varphi(T_0, \theta_{-T_0}(\omega), \xi_n)\| \\ &\quad + \text{dist}_{\mathbb{R}^d}(\varphi(T_0, \theta_{-T_0}(\omega), \xi_n), A(\omega)) \\ &< \frac{1}{4}\varepsilon_0 + \frac{1}{4}\varepsilon_0 = \frac{1}{2}\varepsilon_0, \end{aligned}$$

which contradicts (5.4). The proof is complete. \square

Remark 5.1 A similar result holds when variable step sizes are used, but this results in a more complicated kind of random dynamical system and random attractor. This is because the shift operator on the sequence of time steps becomes a component of the driving system in addition to the noise.

For $0 < a < b$, let $\mathbb{H}^{a,b}$ be the space of positive valued bi-infinite sequences $\mathbf{h} = \{h_j\}_{j \in \mathbb{Z}}$ with $a \leq h_j \leq b$ for $j \in \mathbb{Z}$. It forms a compact metric space with the metric

$$d_{\mathbb{H}}(\mathbf{h}, \mathbf{h}') = \sum_{j \in \mathbb{Z}} b^{-|j|}|h_j - h'_j|.$$

The shift operator ϑ on $\mathbb{H}^{a,b}$, which is defined by $\vartheta(\mathbf{h}) := \mathbf{h}'$ with $h'_j = h_{j+1}$ for $j \in \mathbb{Z}$ forms a discrete time autonomous dynamical system on $\mathbb{H}^{a,b}$.

The implicit Euler scheme (5.2) for a step size sequences $\mathbf{h} \in \mathbb{H}^{a,b}$ takes the form

$$x_{n+1} = x_n + h_n \mu(x_{n+1}) + h_n \zeta(\theta_{h_0+\dots+h_n}(\omega))$$

which can be solved uniquely for sufficiently small step size bound b to give an explicit expression

$$x_{n+1} = G_{h_n}(x_n, \theta_{h_0+\dots+h_n}(\omega)).$$

This generates a discrete time nonautonomous dynamical system (skew product flow) (Θ, ψ) on $\Omega \times \mathbb{H}^{a,b} \times \mathbb{R}^d$ with the autonomous driving system $\Theta := (\theta, \vartheta)$ on $\Omega \times \mathbb{H}^{a,b}$ and cocycle mapping ψ defined by

$$\psi(n, (\mathbf{h}, \omega), \xi) := G_{h_{n-1}}(\cdot, \theta_{h_0+\dots+h_{n-2}}(\omega)) \circ \dots \circ G_{h_0}(\xi, \omega)$$

for $n \geq 2$ with

$$\psi(0, (\mathbf{h}, \omega), \xi) := \xi, \quad \psi(1, (\mathbf{h}, \omega), \xi) := G_{h_0}(\xi, \omega).$$

It does not form a random dynamical system as defined in Chap. 4, but what was called a *partial random dynamical system* and has a *partial random attractor* as in Crauel, Kloeden and Yang [35] with component sets $A_{\mathbf{h}(\omega)}$ defined by pullback attraction. It was shown in Kloeden, Keller and Schmalfuß [84] for a related system that they converge upper semi continuously to the corresponding component sets $A(\omega)$ in the sense that

$$\lim_{b \rightarrow 0} \sup_{\mathbf{h} \in \mathbb{H}^{a,b}} \text{dist}_{\mathbb{R}^d}(A_{\mathbf{h}}(\omega), A(\omega)) = 0.$$

5.2 Discretisation of a Random Hyperbolic Point

A steady state $\bar{x} = 0 = f(0)$ of a deterministic ODE $\dot{x} = f(x)$ is said to be *hyperbolic* if the real parts of the eigenvalues of the matrix in the differential equation obtained by linearisation about the steady state are strictly negative or positive. Beyn [15] showed that the phase portrait of the deterministic dynamical system generated by such an ODE near a hyperbolic steady state $\bar{x} = 0 = f(0)$ is correctly replicated by a one-step numerical method with equal step size $h > 0$, provided some obvious consistency and smoothness assumptions are satisfied.

In particular, the stable and unstable manifolds M_h^s, M_h^u of the discretised system converge to their counterparts M^s, M^u of the original system as $h \rightarrow 0$ with the same order of convergence as the global discretisation error of the numerical method (see Fig. 5.1).

Moreover, the corresponding exact and numerical solutions starting at the same point off the stable and unstable manifolds can diverge. This is because they lie

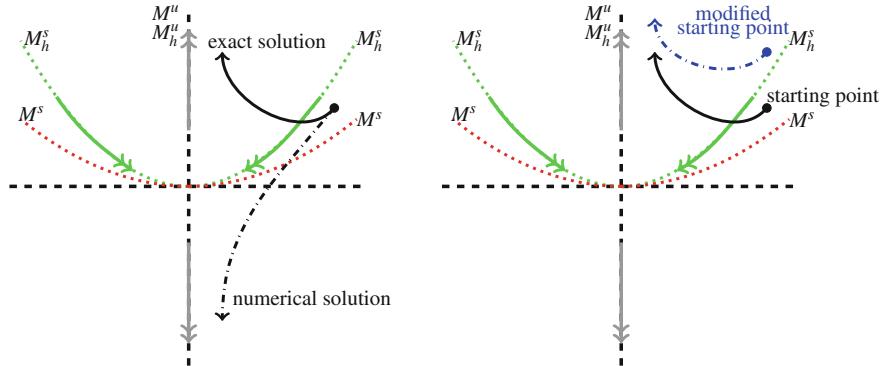


Fig. 5.1 The stable and unstable manifolds of the discretised system

on opposite sides of their corresponding stable or unstable manifolds and hence ultimately move in opposite directions. The starting point of one of them may thus have to be modified to find a numerical solution that approximates an exact solution and vice versa. This is a shadowing result, which says, essentially, that one should compare the totality of solutions rather than individual ones starting at the same point.

A counterpart of the result of Beyn [15] for RODEs will be summarised here following Arnold and Kloeden [6], where further details and proofs can be found. Let

$$\dot{x} = g(x, \theta_t(\omega)) \quad (5.5)$$

be a nonlinear RODE on \mathbb{R}^d driven by noise on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ represented by a metric dynamical system, i.e., group, $\{\theta_t\}_{t \in \mathbb{R}}$ of $(\mathcal{B}(\mathbb{R}) \otimes \mathcal{F}, \mathcal{F})$ measurable mappings with respect to which \mathbb{P} is ergodic. Recall that \mathbb{P} is said to be ergodic with respect to a measure preserving transformation T on $(\Omega, \mathcal{F}, \mathbb{P})$ (or alternatively that T is ergodic with respect to \mathbb{P}), if $\mathbb{P}(A) = 0$ or 1 for every measurable set A satisfying $T^{-1}(A) = A$.

Under appropriate regularity assumptions on g , the pair (θ, φ) forms a nonlinear RDS, where φ is the solution mapping of the RODE.

Random Hyperbolic Points

Let $f(0, \omega) = 0$, then $\bar{x}_0(\omega) \equiv 0$ is a random equilibrium point of (5.5) about which (5.5) can be linearised to yield the linear RODE of variational equations

$$\dot{x} = J(\theta_t(\omega))x, \quad J(\omega) := \left(\frac{\partial g^i(x, \omega)}{\partial x_j} \right) \Big|_{x=0}, \quad (5.6)$$

where g^i is the i th component of g . This linear RODE generates a linear RDS (θ, Φ) with a linear $d \times d$ matrix valued cocycle mapping Φ which is the linearisation of φ about $x = 0$, i.e.,

$$\Phi(t, \omega) := \left(\frac{\partial \varphi^i(t, x, \omega)}{\partial x_j} \right) \Big|_{x=0}, \quad \Phi(0, \omega) = Id.$$

If $\int_{\Omega} \|A(\omega)\| d\mathbb{P}(\omega) < \infty$, then Oseledets' Multiplicative Ergodic Theorem [3] assures the existence of p ($\leq d$) nonrandom numbers $\lambda_1 > \dots > \lambda_p$ (Lyapunov exponents) and a random splitting of $\mathbb{R}^d = E_1(\omega) \oplus \dots \oplus E_p(\omega)$ into Φ -invariant measurable linear subspaces (Oseledets' spaces) such that

$$x \in E_i(\omega) \setminus \{0\} \iff \lim_{t \rightarrow \pm\infty} \frac{1}{t} \log \|\Phi(t, \omega)x\| = \lambda_i, \quad i = 1, \dots, p.$$

Here $\dim E_i(\omega) = d_i$ is nonrandom and $d_1 + \dots + d_p = d$.

Remark 5.2 These assertions hold just for \mathbb{P} -almost all ω , but this will not be repeated in what follows.

The random equilibrium point $\bar{x}_0(\omega) = 0$ is called a *random hyperbolic point* if all of the Lyapunov exponents satisfy $\lambda_i \neq 0$. The corresponding stable and unstable spaces are then defined, respectively, by

$$E^s(\omega) := \bigoplus_{\lambda_i < 0} E_i(\omega), \quad E^u(\omega) := \bigoplus_{\lambda_i > 0} E_i(\omega),$$

for which $E^s(\omega) \oplus E^u(\omega) = \mathbb{R}^d$.

Moreover, there exists a random norm $\|\cdot\|_{\omega}$ and a nonrandom constant $c > 0$ such that with the corresponding random operator norms

$$\|\Phi(t, \omega)|_{E^s(\omega)}\|_{\omega, \theta_t \omega} \leq e^{-ct}, \quad \|\Phi^{-1}(t, \omega)|_{E^u(\omega)}\|_{\theta_t \omega, \omega} \leq e^{-ct}$$

for all $t \geq 0$. The random norm $\|\cdot\|_{\omega}$ is equivalent to the Euclidean norm $\|\cdot\|$ on \mathbb{R}^d with random constants. Specifically, for each $\varepsilon > 0$ there exists a random variable $v_{\varepsilon} : \mathbb{R} \rightarrow [0, \infty)$ with

$$e^{-|t|} v_{\varepsilon}(\omega) \leq v_{\varepsilon}(\theta_t(\omega)) \leq e^{|t|} v_{\varepsilon}(\omega), \quad t \in \mathbb{R},$$

such that

$$\frac{1}{v_{\varepsilon}(\omega)} \|\cdot\| \leq \|\cdot\|_{\omega} \leq v_{\varepsilon}(\omega) \|\cdot\|.$$

Essentially, these random norms absorb the nonuniformities in Φ and allow non-random constants to be used in various estimates.

Numerical RDS

Consider a one-step numerical scheme with constant nonrandom step size $h > 0$,

$$x_{n+1} = G_h(x_n, \theta_{nh}(\omega)) := x_n + hg_h(x_n, \theta_{nh}(\omega)), \quad g_h(0, \omega) = 0, \quad (5.7)$$

corresponding to the RODE (5.5).

Under appropriate regularity assumptions on g_h (see [6]) the numerical scheme (5.7) generates a discrete time RDS (θ_h, ψ_h) with nonlinear cocycle mapping ψ_h . Similarly, its linearisation about the equilibrium point $\bar{x}_0(\omega) = 0$,

$$x_{n+1} = x_n + hJ_h(\theta_{nh}(\omega))x_n, \quad J_h(\omega) := \left(\frac{\partial g_h^i(x, \omega)}{\partial x_j} \right) \Big|_{x=0}$$

generates a discrete time RDS (θ_h, Ψ_h) with a linear $d \times d$ matrix valued cocycle mapping Ψ_h .

The following local discretisation error estimates and consistency assumptions comparing the above continuous and discrete time cocycle mappings are required.

Assumption A (*Linear discretisation error*) There exist nonrandom positive constants ε_0, h_0, C_0 such that

$$\|\Phi(h, \omega) - \Psi_h(1, \omega)\|_{\omega, \theta_h \omega} \leq C_0 h^{1+\varepsilon_0}$$

for all $h \in (0, h_0]$.

Assumption B (*Nonlinear discretisation error*) There exist nonrandom positive constants ε_1, h_1, C_1 and r such that

$$\|\varphi(h, x, \omega) - \psi_h(1, x, \omega)\|_{\theta_h \omega} \leq C_1 h^{1+\varepsilon_1} \|x\|_\omega$$

for all $h \in (0, h_1]$ and $\|x\|_\omega \leq r$.

Assumption C (*Consistency of linear and nonlinear schemes*) There exists a non-random positive constant h^* such that for any $\varepsilon > 0$ there exists $\rho(h, \varepsilon) > 0$ such that the mapping $(\mathbb{R}^d, \|\cdot\|_\omega) \rightarrow (\mathbb{R}^d, \|\cdot\|_{\theta_h \omega})$ determined by $x \mapsto J_h(\omega)x - g_h(x, \omega)$ is Lipschitz on $\{x \in \mathbb{R}^d : \|x\|_\omega \leq \rho(h, \varepsilon)\}$ with Lipschitz constant ε , in the sense that

$$\|J_h(\omega)x - g_h(x, \omega) - J_h(\omega)y + g_h(y, \omega)\|_{\theta_h \omega} \leq \varepsilon \|x - y\|_\omega$$

for all $h \in (0, h^*]$, and all x, y with $\|x - y\|_\omega \leq \rho(h, \varepsilon)$.

The following theorem says, essentially, that such discretisation replicates the phase portrait of the (linear and) nonlinear RDS in a random neighbourhood of a random hyperbolic point. It involves the stable and unstable subspace balls

$$\begin{aligned} B_\rho^s(\omega) &:= \{x \in E^s(\omega) : \|x\|_\omega \leq \rho\}, \\ B_\rho^u(\omega) &:= \{x \in E^u(\omega) : \|x\|_\omega \leq \rho\} \end{aligned}$$

of radius $\rho > 0$.

Theorem 5.2 Let $x = 0$ be a hyperbolic fixed point of the linear cocycle Φ generated by the linear RODE (5.6) and let φ be the nonlinear cocycle generated by the nonlinear RODE (5.5). Suppose that Assumptions A, B and C hold for the numerical scheme (5.7) and its associated linear and nonlinear cocycles Ψ_h and ψ_h .

Then there exists an $h^* > 0$ such that $x = 0$ is hyperbolic for Ψ_h and ψ_h for all $h \in (0, h^*]$. Also for each $h \in (0, h^*]$ there exists $\rho(h) > 0$ and random continuous mappings

$$\phi_{h,\omega}^s : B_{\rho(h)}^s(\omega) \rightarrow E^u(\omega), \quad \phi_{h,\omega}^u : B_{\rho(h)}^u(\omega) \rightarrow E^s(\omega)$$

with $\phi_{h,\omega}^s(0) = \phi_{h,\omega}^u(0) = 0$ such that the sets

$$M_h^s(\omega) := \{(x(\omega), \phi_{h,\omega}^s(x(\omega))) : x(\omega) \in B_{\rho(h)}^s(\omega)\}$$

and

$$M_h^u(\omega) := \{(\phi_{h,\omega}^u(x(\omega)), x(\omega)) : x(\omega) \in B_{\rho(h)}^u(\omega)\}$$

are the local stable and unstable invariant manifolds for the RDS with cocycle mapping ψ_h , respectively.

Moreover, if the mappings

$$\phi_\omega^s : B_{\rho(h)}^s(\omega) \rightarrow E^u(\omega), \quad \phi_\omega^u : B_{\rho(h)}^u(\omega) \rightarrow E^s(\omega)$$

provide similar graph representations of the local stable and unstable manifolds $M^s(\omega)$ and $M^u(\omega)$ corresponding to the cocycle mapping φ , then

$$\begin{aligned} \|\phi_{h,\omega}^s(x(\omega)) - \phi_\omega^s(x(\omega))\|_\omega &\leq \kappa h^a, & x(\omega) \in B_{\rho(h)}^s(\omega), \\ \|\phi_{h,\omega}^u(x(\omega)) - \phi_\omega^u(x(\omega))\|_\omega &\leq \kappa h^a, & x(\omega) \in B_{\rho(h)}^u(\omega) \end{aligned}$$

for some constants $a > 0$ and κ .

Finally, given $\xi(\omega) \notin M^s(\omega)$ with $\|\xi(\omega)\|_\omega \leq \rho(h)$ there exists a $y_h(\omega)$ with $\|y_h(\omega)\|_\omega \leq \rho(h)$ and a positive integer $N(h, \omega)$ such that

$$\|\varphi(jh, \xi(\omega), \omega)\|_{\theta_{jh}(\omega)} \leq \rho(h), \quad \|\psi_h(j, y_h(\omega), \omega)\|_{\theta_{jh}(\omega)} \leq \rho(h)$$

and

$$\|\varphi(jh, \xi(\omega), \omega) - \psi_h(j, y_h(\omega), \omega)\|_{\theta_{jh}(\omega)} \leq \kappa h^a$$

for $j = 0, 1, \dots, N(h, \omega)$; and vice versa.

Since the negativity of all Lyapunov exponents implies pathwise exponential asymptotic stability of the equilibrium point, the following corollary is an immediate consequence of Theorem 5.2.

Corollary 5.1 *If the null solution of the linear RODE (5.6) is exponentially asymptotic stable, then it is also exponentially asymptotically stable for the nonlinear RODE (5.5) and for the numerical scheme for $h \in (0, h^*]$ with a random neighbourhood of 0 as the domain of attraction.*

Assumptions A and B are quite demanding as they require the stable and unstable manifolds of the linear cocycles to be locked in close together, whereas rotation with albeit small probability is possible in general [111]. Examples where they are satisfied are given in [6], including the small noisy perturbation of a deterministic linear hyperbolic ODE.

5.3 Endnotes

This chapter is based on an updated and reworked version of the 1999 article *Towards a theory of random numerical dynamics* by Kloeden, Keller and Schmalfuß [84].

See Stuart and Humphries [128] for an introduction to deterministic numerical dynamics. Kloeden and Schmalfuß [94] investigate variable time step discretisation of deterministic nonautonomous systems and their attractors.

The results on the discretisation of a random dynamical system near a hyperbolic point, follow Arnold and Kloeden [6], which was motivated by Beyn [15], where deterministic systems were considered.

Part II

Taylor Expansions

Chapter 6

Taylor Expansions for Ordinary and Stochastic Differential Equations

Taylor expansions are a very basic tool in numerical analysis and other areas of mathematics which require approximations. In particular, they allow one to derive one-step numerical schemes for ordinary differential equations (ODEs) of arbitrarily high order, although in practice such Taylor schemes are rarely implemented but are used instead as a theoretical comparison for determining the convergence orders of other schemes that have been derived by more heuristic methods. On the other hand, in view of the less robust nature of the Itô stochastic integral, stochastic Taylor expansions and the corresponding stochastic Taylor schemes are the essential starting point for the derivation of consistent higher order numerical schemes for stochastic differential equations (SODEs).

6.1 Taylor Approximations for ODEs

The Taylor expansion of a $p + 1$ times continuously differentiable function $x : [0, T] \rightarrow \mathbb{R}$ about $t_0 \in [0, T]$ is given by

$$x(t_0 + h) = x(t_0) + x'(t_0) h + \cdots + \frac{1}{p!} x^{(p)}(t_0) h^p + \frac{1}{(p+1)!} x^{(p+1)}(\tau_h) h^{p+1} \quad (6.1)$$

with an intermediate value (usually unknown) $\tau_h \in [t_0, t_0 + h] \subset [0, T]$ in the last term.

This expression applies to the solution $x(t) = x(t, t_0, x_0)$ of the initial value problem for a scalar ordinary differential equation

$$\frac{dx}{dt} = f(t, x), \quad x(t_0) = x_0 \quad (6.2)$$

on an interval $[t_0, T]$ provided the vector field is at least p times continuously differentiable in its variables. In this case the derivatives $x^{(j)}(t)$ are given by

$$x^{(j)}(t) = D^{j-1} f(t, x(t)), \quad j = 1, 2, \dots, p,$$

where Du is the total derivative of a function u along a solution of the ODE (6.2), i.e.,

$$Du(t, x) := \frac{\partial u}{\partial t}(t, x) + f(t, x) \frac{\partial u}{\partial x}(t, x).$$

It follows by the chain rule that

$$\frac{d}{dt} u(t, x(t)) = \frac{\partial u}{\partial t}(t, x(t)) + \frac{\partial u}{\partial x}(t, x(t)) \cdot \frac{d}{dt} x(t) = Du(t, x(t)),$$

and $D^{j+1} = DD^j$ with $D^0 \equiv Id$.

This gives the p -Taylor approximation

$$\Phi_p(t_0, x_0, h) = x_0 + \sum_{j=1}^p \frac{h^j}{j!} D^{j-1} f(t_0, x_0),$$

of the solution value $x(t_0 + h, x_0)$. It has a truncation error of order $p + 1$, since

$$\begin{aligned} \mathcal{E}_p(x_0, t_0, h) &:= |x(t_0 + h, x_0) - \Phi_p(x_0, t_0, h)| \\ &= \frac{h^{p+1}}{(p+1)!} |D^p f(\tau_h, x(\tau_h, t_0, x_0))| \leq C_{p,T,B} h^{p+1}, \end{aligned}$$

where

$$C_{p,T,B} := \frac{1}{(p+1)!} \max_{\substack{t_0 \leq t \leq T \\ x \in B}} |D^p f(t, x)|$$

and B is a large compact set which contains the solution $x(t)$ on the time interval $[0, T]$.

A similar situation holds for vector ODEs, but with more complicated terminology. They are essentially a special case of the following expansions for Itô SODEs with the noise coefficients set to zero.

6.2 Taylor Approximations of Itô SODEs

First consider for simplicity a scalar Itô SODE

$$dX_t = f(t, X_t) dt + \sigma(t, X_t) dW_t,$$

with drift and diffusion coefficients $f, \sigma : [t_0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ and a scalar Wiener process W_t . Such an Itô SODE is only symbolic for the Itô stochastic integral equation

$$X_t = X_{t_0} + \int_{t_0}^t f(s, X_s) ds + \int_{t_0}^t \sigma(s, X_s) dW_s, \quad (6.3)$$

where the drift integral is pathwise a Riemann integral and the diffusion integral is an Itô stochastic integral.

The sample paths of the solution process of the SODE (6.3) are only Hölder continuous, but not differentiable and hence do not have a Taylor expansion like (6.1) for the solutions of an ODE. Instead, a stochastic Taylor expansion can be derived through an iterated application of the Itô formula, i.e., stochastic chain rule. Specifically, for a function $U : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$, the Itô formula reads

$$U(t, X_t) = U(t_0, X_{t_0}) + \int_{t_0}^t L^0 U(s, X_s) ds + \int_{t_0}^t L^1 U(s, X_s) dW_s$$

with the differential operators

$$L^0 = \frac{\partial}{\partial t} + f \frac{\partial}{\partial x} + \frac{1}{2} \sigma^2 \frac{\partial^2}{\partial x^2}, \quad L^1 = \sigma \frac{\partial}{\partial x}.$$

Applying this to $U = f$ and $U = \sigma$ in (6.3) respectively gives the simple stochastic Taylor expansion

$$\begin{aligned} X_t &= X_{t_0} + f(t_0, X_{t_0}) \int_{t_0}^t dt + \sigma(t_0, X_{t_0}) \int_{t_0}^t dW_s \\ &\quad + \int_{t_0}^t \int_{t_0}^s L^0 f(\tau, X_\tau) d\tau ds + \int_{t_0}^t \int_{t_0}^s L^1 f(\tau, X_\tau) dW_\tau ds \\ &\quad + \int_{t_0}^t \int_{t_0}^s L^0 \sigma(\tau, X_\tau) d\tau dW_s + \int_{t_0}^t \int_{t_0}^s L^1 \sigma(\tau, X_\tau) dW_\tau dW_s. \end{aligned}$$

Now consider a d -dimensional Itô SODE with m mutually independent scalar Wiener processes W_t^1, \dots, W_t^m in the integral form

$$X_t = X_{t_0} + \int_{t_0}^t f(s, X_s) ds + \sum_{j=1}^m \int_{t_0}^t \sigma_j(s, X_s) dW_s^j. \quad (6.4)$$

Defining $f_0 = f$ and $f_j = \sigma_j$, $j = 1, \dots, m$, (6.4) can be written more compactly as

$$X_t = X_{t_0} + \sum_{j=0}^m \int_{t_0}^t f_j(s, X_s) dW_s^j, \quad (6.5)$$

with a fictitious “noise” component $W_t^0 \equiv t$, so the first integral term can be included in the summation, which will be notationally very convenient in what follows. The notation of hierarchical sets of multi-indices, iterated integrals and iterated differential operators from Kloeden and Platen [91] allows expansions of arbitrarily high order to be written out very compactly and transparently.

6.2.1 Multi-indices

Let $m \geq 0$ correspond to the number of components of the Wiener process in the SODE (6.4). A row vector

$$\alpha = (j_1, j_2, \dots, j_l),$$

where $j_i \in \{0, 1, \dots, m\}$ for $i = 1, \dots, l$, is called a *multi-index* of length $\ell(\alpha) = l \geq 1$. For completeness \emptyset will denote the multi-index of length zero, i.e., $\ell(\emptyset) = 0$. Let \mathcal{J}_m be the set of all such multi-indices, i.e.,

$$\mathcal{J}_m = \{\alpha : j_i \in \{0, 1, \dots, m\}, i \in \{1, \dots, l\} \text{ for } l = 1, 2, 3, \dots\} \bigcup \{\emptyset\}.$$

For any $\alpha = (j_1, j_2, \dots, j_l) \in \mathcal{J}_m$ with $\ell(\alpha) \geq 2$, denote by $-\alpha$ and $\alpha-$ for the multi-indices in \mathcal{J}_m obtained by deleting the first and the last component of α , respectively, i.e.,

$$-\alpha = (j_2, \dots, j_l), \quad \alpha- = (j_1, \dots, j_{l-1}),$$

with $-(j_1) = (j_1)- = \emptyset$ when $\ell(\alpha) = 1$.

Finally, define $n(\alpha)$ to be the number of components of a multi-index $\alpha \in \mathcal{J}_m$ that are equal to 0.

6.2.2 Multiple Integrals of Stochastic Processes

For a multi-index $\alpha = (j_1, j_2, \dots, j_l) \in \mathcal{J}_m$ and an integrable function $u : [t_0, T] \rightarrow \mathbb{R}$ the *multiple integral* $I_\alpha[u(\cdot)]_{t_0, t}$ is defined recursively by

$$I_\alpha[u(\cdot)]_{t_0, t} := \begin{cases} u(t), & l = 0, \\ \int_{t_0}^t I_{\alpha-[u(\cdot)]_{t_0, s}} dW_s^{j_l}, & l \geq 1. \end{cases}$$

For example,

$$I_\emptyset[u(\cdot)]_{t_0, t} = u(t), \quad I_{(0)}[u(\cdot)]_{t_0, t} = \int_{t_0}^t u(s) ds, \quad I_{(1)}[u(\cdot)]_{t_0, t} = \int_{t_0}^t u(s) dW_s^1,$$

$$\begin{aligned} I_{(0,1)}[u(\cdot)]_{0,t} &= \int_0^t \int_0^{s_2} u(s_1) ds_1 dW_{s_2}^1 = \int_0^t \left[\int_0^{s_2} u(s_1) ds_1 \right] dW_{s_2}^1, \\ I_{(0,2,1)}[u(\cdot)]_{0,t} &= \int_0^t \int_0^{s_3} \int_0^{s_2} u(s_1) ds_1 dW_{s_2}^2 dW_{s_3}^1. \end{aligned}$$

For simpler notation, $I_\alpha[u(\cdot)]_{t_0,t}$ will often be abbreviated to $I_{\alpha,t_0,t}$ or just I_α when $u(t) \equiv 1$.

6.2.3 Coefficient Functions

Assume that the functions f_0, f_1, \dots, f_m are sufficiently smooth and define the partial differential operators L^0, L^1, \dots, L^m by

$$\begin{aligned} L^0 &= \frac{\partial}{\partial t} + \sum_{k=1}^d f_0^k \frac{\partial}{\partial x^k} + \frac{1}{2} \sum_{k,l=1}^d \sum_{j=1}^m f_j^k f_j^l \frac{\partial^2}{\partial x^k \partial x^l}, \\ L^j &= \sum_{k=1}^d f_j^k \frac{\partial}{\partial x^k}, \quad j = 1, \dots, m, \end{aligned}$$

where f_j^k is the k th component of the vector valued function f_j for $j = 0, 1, \dots, m$, and x^k is the k th component of x , for $k = 1, \dots, d$.

For each $\alpha = (j_1, \dots, j_l) \in \mathcal{J}_m$ and a sufficiently smooth function $F : [t_0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$, the *coefficient function* F_α is defined recursively by

$$F_\alpha = \begin{cases} F, & l = 0 \\ L^{j_1} F_{-\alpha}, & l \geq 1 \end{cases}. \quad (6.6)$$

Example 6.1 For an autonomous scalar SODE (6.5) with scalar noise where $d = m = 1$ and $f_j = f_j(x)$, the coefficient functions corresponding to the identity function $F(t, x) \equiv x$ are

$$\begin{aligned} F_{(0)} &= f_0, \quad F_{(j_1)} = f_{j_1}, & F_{(0,0)} &= f_0 f_0' + \frac{1}{2} (f_1)^2 f_0'', \\ F_{(0,j_2)} &= f_0 f_{j_2}' + \frac{1}{2} (f_1)^2 f_{j_2}'', & F_{(j_1,0)} &= f_0' f_{j_1}, \quad F_{(j_1,j_2)} = f_{j_1} f_{j_2}'. \end{aligned}$$

where the dash ' denotes differentiation with respect to x .

When the function F is not stated explicitly in the text it is always taken to be the identity function $F(t, x) \equiv x$.

6.2.4 Hierarchical and Remainder Sets

Since different integrals can be expanded in forming an Itô–Taylor expansion, the terms with constant integrands cannot be written down completely arbitrarily. Rather, the set of corresponding multi-indices must form an hierarchical set.

Definition 6.1 A subset $\mathcal{H} \subset \mathcal{J}_m$ is called an hierarchical set if

- (i) \mathcal{H} is nonempty;
- (ii) the multi-indices in \mathcal{H} are uniformly bounded in length, i.e., $\sup_{\alpha \in \mathcal{H}} l(\alpha) < \infty$;
- (iii) $-\alpha \in \mathcal{H}$ for each $\alpha \in \mathcal{H} \setminus \{\emptyset\}$, where \emptyset is the multi-index of length zero.

Thus, if a multi-index α belongs to an hierarchical set, then so does the multi-index $-\alpha$ obtained by deleting the first component of α .

The remainder term of a Taylor expansion constructed with a given hierarchical set \mathcal{H} involves only those multiple integrals of stochastic processes with multi-indices belonging to the corresponding *remainder set* $\mathcal{R}(\mathcal{H})$, which is defined by

$$\mathcal{R}(\mathcal{H}) = \{\alpha \in \mathcal{J}_m \setminus \mathcal{H} : -\alpha \in \mathcal{H}\}.$$

It thus consists of all of the next following multi-indices with respect to the given hierarchical set that do not already belong to the hierarchical set and is formed simply by adding a further component taking all possible values at the beginning of the “maximal” multi-indices in the hierarchical set.

6.3 General Itô–Taylor Expansions

The above terminology allows general Itô–Taylor expansions for an Itô SODE to be derived and expressed in a compact way.

Theorem 6.1 Let $F : \mathbb{R}^+ \times \mathbb{R}^d \rightarrow \mathbb{R}$ and let $\mathcal{H} \subset \mathcal{J}_m$ be an hierarchical set with remainder set $\mathcal{R}(\mathcal{H})$. Then the Itô–Taylor expansion corresponding to the hierarchical set \mathcal{H} is given by

$$F(t, X_t) = \sum_{\alpha \in \mathcal{H}} F_\alpha(t_0, X_{t_0}) I_\alpha[1]_{t_0, t} + \sum_{\alpha \in \mathcal{R}(\mathcal{H})} I_\alpha [F_\alpha(\cdot, X_\cdot)]_{t_0, t}, \quad (6.7)$$

provided all of the derivatives of F , f_0, f_1, \dots, f_m and all of the multiple integrals of stochastic processes appearing here exist and are continuous.

Proof The proof follows that of the Itô–Taylor expansion for SODE [91, Theorem 5.5.1]. A version restricted to simple types of hierarchical sets is sketched here.

First recall the integral version of the chain rule working on a function $F(t, X_t)$ of a solution to the SODE (6.5),

$$F(t, X_t) = F(t_0, X_{t_0}) + \int_{t_0}^t L^0 F(s, X_s) ds + \sum_{j=1}^m \int_{t_0}^t L^j F(s, X_s) dW_s^j. \quad (6.8)$$

Notice that

$$\int_{t_0}^t L^0 F(s, X_s) ds = I_{(0)} [L^0 F(\cdot, X_\cdot)]_{t_0, t}$$

and with $(j)- = \emptyset$

$$\int_{t_0}^t L^j F(s, X_s) dW_s^j = \int_{t_0}^t I_{(j)-} [L^j F(s, X_s)]_{t_0, s} dW_s^j = I_{(j)} [L^j F(\cdot, X_\cdot)]_{t_0, t}.$$

Hence (6.8) can be written as

$$F(t, X_t) = F(t_0, X_{t_0}) + \sum_{j=0}^m I_{(j)} [L^{(j)} F(\cdot, X_\cdot)]_{t_0, t}. \quad (6.9)$$

Apply (6.9) to the function F_α for some multi-index $\alpha = (j_1, \dots, j_l) \in \mathcal{H}$ to obtain

$$\begin{aligned} I_\alpha [F_\alpha(\cdot, X_\cdot)]_{t_0, t} &= F(t_0, X_{t_0}) I_\alpha [1]_{t_0, t} + I_\alpha \left[\sum_{j=0}^m I_{(j)} [L^j F_\alpha(\cdot, X_\cdot)]_{t_0, \cdot} \right]_{t_0, t} \\ &= F_\alpha(t_0, X_{t_0}) I_\alpha [1]_{t_0, t} + \sum_{j=0}^m I_{(j)*\alpha} [F_{(j)*\alpha}(\cdot, X_\cdot)]_{t_0, t}, \end{aligned}$$

where $(j)*\alpha$ is the concatenated multi-index for (j) and $\alpha = (j_1, \dots, j_l)$, i.e., $(j)*\alpha = (j, j_1, \dots, j_l)$.

The expression (6.7) will be verified by induction over $\ell := \max\{\mathbf{l}(\alpha) \mid \alpha \in \mathcal{H}\}$.

First for $\ell = 0$, the hierarchical set is simply $\mathcal{H} = \{\emptyset\}$, so the assertion follows directly from (6.9).

For $\ell \geq 1$, consider the hierarchical set $\mathcal{E} := \{\alpha \in \mathcal{H} \mid \mathbf{l}(\alpha) \leq \ell - 1\}$ and assume (for induction) that

$$F(t, X_t) = \sum_{\alpha \in \mathcal{E}} F(t_0, X_{t_0}) I_\alpha [1]_{t_0, t} + \sum_{\alpha \in \mathcal{R}(\mathcal{E})} I_\alpha [F(\cdot, X_\cdot)]_{t_0, t}.$$

Since $\mathcal{H} \setminus \mathcal{E} \subseteq \mathcal{R}(\mathcal{E})$ by the definition of a remainder set, it can be concluded that

$$\begin{aligned} F(t, X_t) &= \sum_{\alpha \in \mathcal{E}} F(t_0, X_{t_0}) I_\alpha[1]_{t_0, t} + \sum_{\alpha \in \mathcal{H} \setminus \mathcal{E}} I_\alpha[F(\cdot, X_\cdot)]_{t_0, t} \\ &\quad + \sum_{\alpha \in \mathcal{R}(\mathcal{H}) \setminus (\mathcal{H} \setminus \mathcal{E})} I_\alpha[F(\cdot, X_\cdot)]_{t_0, t} \\ &= \sum_{\alpha \in \mathcal{E}} F(t_0, X_{t_0}) I_\alpha[1]_{t_0, t} + \sum_{\alpha \in \tilde{\mathcal{R}}} I_\alpha[F(\cdot, X_\cdot)]_{t_0, t}, \end{aligned}$$

where the last equality follows from (6.6).

The desired expression (6.7) then follows since the definition of a remainder set implies that $\tilde{\mathcal{R}} = \mathcal{R}(\mathcal{H})$. \square

Example 6.2 In the general case with the hierarchical and remainder sets

$$\mathcal{H} = \{\emptyset\}, \quad \mathcal{R}(\{\emptyset\}) = \{(0), \dots, (m)\},$$

the Itô–Taylor expansion is

$$\begin{aligned} F(t, X_t) &= F_\emptyset(t_0, X_{t_0}) I_\emptyset[1]_{t_0, t} + \sum_{\alpha \in \mathcal{R}(\{\emptyset\})} I_\alpha[F_\alpha(\cdot, X_\cdot)]_{t_0, t} \\ &= F(t_0, X_{t_0}) + \int_{t_0}^t L^0 F(s, X_s) ds + \sum_{j=1}^m \int_{t_0}^t L^j F(s, X_s) dW_s^j. \end{aligned}$$

This is just the chain rule (6.8) for a function $F(t, X_t)$ of a solution of SODE (6.5).

Example 6.3 Consider the autonomous scalar case $d = m = 1$ with $F(t, x) \equiv x$ and the hierarchical and remainder sets

$$\mathcal{H} = \{\alpha \in \mathcal{J}_1 : l(\alpha) \leq 2\}, \quad \mathcal{R}(\mathcal{H}) = \{\alpha \in \mathcal{J}_1 : l(\alpha) = 3\}.$$

Using the coefficient functions calculated in Example 6.1, the Itô–Taylor expansion is

$$\begin{aligned} X_t &= X_{t_0} + f_0 I_{(0)} + f_1 I_{(1)} + \left(f_0 f_0' + \frac{1}{2} (f_1)^2 f_0'' \right) I_{(0,0)} \\ &\quad + \left(f_0 f_1' + \frac{1}{2} (f_1)^2 f_1'' \right) I_{(0,1)} + f_1 f_0' I_{(1,0)} + f_1 f_1' I_{(1,1)} + R_3(t, t_0). \end{aligned}$$

Here the integrals are over the interval $[t_0, t]$, the coefficient functions are all evaluated at (t_0, x_0) , the dash ' denotes differentiation with respect to x , and $R_3(t, t_0)$ is the corresponding remainder term.

6.4 Strong Itô–Taylor Approximations

Recall that the d -dimensional Itô SODE (6.5) with m mutually independent scalar Wiener processes W_t^1, \dots, W_t^m can be written in the convenient compact form

$$dX_t = \sum_{j=0}^m f_j(t, X_t) dW_t^j, \quad (6.10)$$

with the convention that $W_t^0 = t$.

For $\gamma = \frac{1}{2}, 1, \frac{3}{2}, \dots$, let Λ_γ be the hierarchical set of multi-indices given by

$$\Lambda_\gamma = \left\{ \alpha \in \mathcal{J}_m : l(\alpha) + n(\alpha) \leq 2\gamma \text{ or } l(\alpha) = n(\alpha) = \gamma + \frac{1}{2} \right\},$$

where $l(\alpha)$ is the length of α and $n(\alpha)$ is the number of zero entries of α .

With this notation, omitting the remainder terms in general Itô–Taylor expansion in Theorem 6.1 applied to the identity function leads to the following strong Itô–Taylor approximation.

Definition 6.2 The order γ strong Itô–Taylor approximation for a solution of the SODE (6.10) is defined componentwise as

$$\Phi_\gamma^k(t_0, X_{t_0}, h) = \sum_{\alpha \in \Lambda_\gamma} L^{j_1} L^{j_2} \cdots L^{j_{i-1}} f_{j_i}^k(t_0, X_{t_0}) I_\alpha[1]_{t_0, t_0+h}, \quad (6.11)$$

where Φ_γ^k is the k th component of the vector Φ_γ and $f_{j_i}^k$ is the k th component of the vector f_{j_i} .

By Theorem 10.6.4 in [91] the order γ strong Taylor approximation (6.11) converges *strongly* with order γ , i.e.,

$$\mathbb{E} [| \Phi_\gamma(t_0, X_{t_0}, h) - X_{t_0+h} |] \leq C_{\gamma, T} h^\gamma$$

provided that the coefficients of the scheme and its remainder satisfy global Lipschitz conditions. This follows if the coefficient functions in the SODE (6.10) are $2\gamma + 1$ times continuously differentiable with all of the partial derivatives being bounded uniformly. These are called the *standard assumptions*, but are not satisfied by many SODEs in important applications (see Hutzenthaler and Jentzen [67]). One way to overcome this problem is to use a stopping time to restrict to a bounded set.

6.4.1 Examples

Consider a scalar Itô SODE with $d = m = 1$:

$$dX_t = f(t, X_t) dt + \sigma(t, X_t) dW_t.$$

For the multi-indices $\alpha = (0), (1), (1, 1)$, the coefficient functions F_α are

$$F_{(0)} = f, \quad F_{(1)} = \sigma, \quad F_{(1,1)} = \sigma \frac{\partial \sigma}{\partial x}$$

and the iterated integrals are

$$\begin{aligned} I_{(0), t_0, t_0+h} &= \int_{t_0}^{t_0+h} dW_s^0 = h, & I_{(1), t_0, t_0+h} &= \int_{t_0}^{t_0+h} dW_s^1 = \Delta W_0, \\ I_{(1,1), t_0, t_0+h} &= \int_{t_0}^{t_0+h} \int_{t_0}^s dW_\tau^1 dW_s^1 = \frac{1}{2} [(\Delta W_0)^2 - h], \end{aligned}$$

where

$$\Delta W_0 = W_{t_0+h} - W_{t_0}.$$

The Itô–Taylor approximation of strong order $\gamma = \frac{1}{2}$ corresponding to the hierarchical set $\Lambda_{\frac{1}{2}} = \{(0), (1)\}$ is thus

$$\Phi_{\frac{1}{2}}(t_0, X_{t_0}, h) = X_{t_0} + f(t_0, X_{t_0}) h + \sigma(t_0, X_{t_0}) \Delta W_0,$$

while the Itô–Taylor approximation of strong order $\gamma = 1$ corresponding to the hierarchical set $\Lambda_1 = \{(0), (1), (1, 1)\}$ is

$$\Phi_1(t_0, X_{t_0}, h) = \Phi_{\frac{1}{2}}(t_0, X_{t_0}, h) + \frac{1}{2} \sigma(t_0, X_{t_0}) \frac{\partial \sigma}{\partial x}(t_0, X_{t_0}) [(\Delta W_0)^2 - h].$$

Note that the index (0) appears in $\Lambda_{\frac{1}{2}} = 1$ due to the condition $l((0)) = n((0)) = \gamma + \frac{1}{2} = 1$, while it appears in Λ_1 due to the condition $l((0)) \leq \gamma = 1$.

6.4.2 Pathwise Convergence

A stochastic approximation is said to converge *pathwise* if

$$|\Phi_\gamma(t_0, X_{t_0}, h, \omega) - X_{t_0+h}(\omega)| \rightarrow 0 \quad \text{as } h \rightarrow 0$$

for (at least) almost all $\omega \in \Omega$, where Ω is the sample space of the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$. This is interesting because numerical calculations of the approximating random variables are carried out path by path. In addition, the solutions of some SODEs are non-integrable, i.e., $\mathbb{E}[|X_t|] = \infty$ for some $t \geq 0$, so a strongly convergent approximation is not always possible.

Itô stochastic calculus is a mean-square calculus but not a pathwise calculus. Nevertheless some results for the pathwise approximation of solutions of SODEs are known.

Theorem 6.2 *Under the standard assumptions an Itô–Taylor approximation of strong order $\gamma > 0$ converges pathwise with order $\gamma - \varepsilon$ for all $\varepsilon > 0$, i.e.,*

$$|\Phi_\gamma(t_0, X_{t_0}, h, \omega) - X_{t_0+h}(\omega)| \leq C_{\varepsilon, T}^\gamma(\omega) \cdot h^{\gamma-\varepsilon}$$

for almost all $\omega \in \Omega$.

Note that the error constant in Theorem 6.2 depends on ω , so is, in fact, a random variable. The nature of its statistical properties is an interesting question, but so far only some empirical distributions of the random error have been found.

The proof of Theorem 6.2 is based on a stronger form of strong convergence,

$$(\mathbb{E} [| \Phi_\gamma(t_0, X_{t_0}, h) - X_{t_0+h} |^p])^{1/p} \leq C(p, \gamma, T) h^\gamma, \quad p \in \mathbb{N},$$

which can be established using the Burkholder–Davis–Gundy inequality (see Appendix A). The following link between the common convergence rate of the p -th means and the pathwise convergence rate is then applied to the random variables

$$Z_n = |\Phi_\gamma(t_0, X_{t_0}, 1/n) - X_{t_0+1/n}|$$

in the following lemma.

Lemma 6.1 *Let $a > 0$ and $C(p) \in [0, \infty)$ for $p \geq 1$. Moreover let $Z_n, n \in \mathbb{N}$ be a sequence of random variables such that*

$$(\mathbb{E} [| Z_n |^p])^{1/p} \leq C(p) \cdot n^{-a}$$

for all $p \in \mathbb{N}$ and all $n \in \mathbb{N}$.

Then for all $\varepsilon > 0$ there exists a random variable η_ε such that

$$|Z_n| \leq \eta_\varepsilon \cdot n^{-a+\varepsilon} \quad a.s.$$

for all $n \in \mathbb{N}$. Moreover, $\mathbb{E} [|\eta_\varepsilon|^p] < \infty$ for all $p \in \mathbb{N}$.

Proof Fix $\varepsilon > 0$ and $p > 1/\varepsilon$. Then for all $\delta > 0$ it follows from the Chebyshev–Markov inequality and the assumptions of the lemma that

$$\mathbb{P}(n^{a-\varepsilon}|Z_n| > \delta) \leq \frac{\mathbb{E}[|Z_n|^p]}{\delta^p} n^{(a-\varepsilon)p} \leq \frac{C(p)^p}{\delta^p} n^{-p\varepsilon}.$$

Since $p > 1/\varepsilon$,

$$\sum_{n=1}^{\infty} \mathbb{P}(n^{a-\varepsilon}|Z_n| > \delta) < \infty$$

for all $\delta > 0$. Thus the Borel–Cantelli Lemma (see Appendix A) implies that $Z_n \rightarrow 0$ a.s. for $n \rightarrow \infty$. Now set $\eta_\varepsilon = \sup_{n \in \mathbb{N}} n^{a-\varepsilon}|Z_n|$. It follows that

$$\mathbb{E}[|\eta_\varepsilon|^p] \leq \mathbb{E}\left[\sup_{n \in \mathbb{N}} n^{(a-\varepsilon)p} |Z_n|^p\right] \leq \sum_{n=1}^{\infty} n^{(a-\varepsilon)p} \mathbb{E}[|Z_n|^p] \leq C(p)^p \sum_{n=1}^{\infty} n^{-p\varepsilon} < \infty$$

for $p > 1/\varepsilon$. Jensen’s inequality then gives $\mathbb{E}|\eta_\varepsilon|^p < \infty$ for all $p \geq 1$. The assertion of the lemma now follows by

$$|Z_n| \leq \left(\sup_{n \in \mathbb{N}} n^{a-\varepsilon} |Z_n|\right) \cdot n^{-a+\varepsilon} = \eta_\varepsilon \cdot n^{-a+\varepsilon}.$$

□

6.5 Endnotes

For an alternative derivation of the p -Taylor approximation of the solution of an ODE using the integral equation representation of an initial value problem see Jentzen and Kloeden [75] and Chap. 1. See also Deuflhard and Bornemann [38], Hairer and Wanner [59, 60] and Hairer, Nørsett and Wanner [61].

See Kloeden and Platen [91] and Milstein [105] for detailed discussions of stochastic Taylor expansions for Itô SODEs. See also Milstein and Tretyakov [106].

Theorem 6.2 is taken from Jentzen, Kloeden and Neuenkirch [80] with the proof there being adapted from Gyöngy [58]. The Burkholder–Davis–Gundy inequality and the Borel–Cantelli Lemma are given in Appendix A.

Hutzenthaler and Jentzen [67] and Hutzenthaler, Jentzen and Kloeden [68] consider strong Itô–Taylor schemes for SODEs which do not satisfy the standard assumptions.

Taylor schemes for rough differential equations are derived in Hu, Li and Nualart [65].

Chapter 7

Taylor Expansions for RODEs with Affine Noise

A d -dimensional RODE with m -dimensional affine noise has the structure

$$\frac{dx}{dt} = f_0(t, x) + \sum_{j=1}^m f_j(t, x) \eta_t^j, \quad (7.1)$$

where $x = (x^1, \dots, x^d) \in \mathbb{R}^d$ and the noise process $\eta_t = (\eta_t^1, \dots, \eta_t^m)$ takes values in \mathbb{R}^m . The sample paths of noise processes are assumed to be Lebesgue measurable, so the differential equation (7.1) is interpreted in the sense of Carathéodory and has an equivalent integral equation representation

$$x(t) = x(t_0) + \int_{t_0}^t f_0(s, x(s)) ds + \sum_{j=1}^m \int_{t_0}^t f_j(s, x(s)) \eta_s^j ds. \quad (7.2)$$

Taylor expansions of a function $U(t, x(t))$ of a solution $x(t)$ of the RODE with affine noise (7.1) can be derived in a similar way to the Itô-Taylor expansions for solutions of SODEs in the previous chapter, but without stochastic integrals. It is based on an iterated application of the chain rule for the absolutely continuous solutions of (7.1) (see Theorem B.3 in Appendix B). For any continuously differentiable function $U : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ this chain rule in integral form reads

$$U(t, x(t)) = U(t_0, x(t_0)) + \int_{t_0}^t L^0 U(s, x(s)) ds + \sum_{j=1}^m \int_{t_0}^t L^j U(s, x(s)) \eta_s^j ds$$

for $t \in [t_0, T]$, with the partial differential operators L^0 and L^1, \dots, L^m defined by

$$L^0 = \frac{\partial}{\partial t} + \sum_{k=1}^d f_0^k \frac{\partial}{\partial x^k}, \quad L^j = \sum_{k=1}^d f_j^k \frac{\partial}{\partial x^k}, \quad j = 1, \dots, m,$$

where f_j^k is the k th component of the vector valued function f_j for $k = 1, \dots, d$ and $j = 0, 1, \dots, m$.

7.1 An Illustrative Example

Consider the solution $x(t)$ of the 1-dimensional “autonomous” RODE with scalar affine noise:

$$\frac{dx}{dt} = f(x) + \sigma(x) \eta_t$$

with the equivalent integral equation representation

$$x(t) = x(t_0) + \int_{t_0}^t f(x(s)) ds + \int_{t_0}^t \sigma(x(s)) \eta_s ds \quad (7.3)$$

for $t \in [t_0, T]$, where the coefficients f and σ in (7.3) are assumed to be sufficiently smooth real-valued functions and satisfy necessary properties to ensure the existence and uniqueness of solutions on the whole time interval $[t_0, T]$. In particular, the function $G(t, x) := f(x) + \sigma(x)\eta_t$ should satisfy the Carathéodory conditions 2.2, so that the existence and uniqueness theorem for Carathéodory ODEs, Theorem 2.3, can be applied.

For a continuously differentiable function $U : \mathbb{R} \rightarrow \mathbb{R}$ the chain rule for the absolutely continuous solutions of Eq. (7.3) reads

$$U(x(t)) = U(x(t_0)) + \int_{t_0}^t L^0 U(x(s)) ds + \int_{t_0}^t L^1 U(x(s)) \eta_s ds, \quad (7.4)$$

for $t \in [t_0, T]$, where the operators L^0 and L^1 are defined by

$$L^0 = f \frac{\partial}{\partial x}, \quad L^1 = \sigma \frac{\partial}{\partial x}.$$

If $U(x) \equiv x$, then $L^0 U = f$ and $L^1 U = \sigma$, in which case (7.4) reduces to the original RODE with affine noise (7.3), i.e., to

$$x(t) = x(t_0) + \int_{t_0}^t f(x(s)) ds + \int_{t_0}^t \sigma(x(s)) \eta_s ds.$$

Applying the chain rule (7.4) to each of the functions $U = f$ and $U = \sigma$, respectively, in (7.3) gives

$$x(t) = x(t_0) + \int_{t_0}^t \left(f(x(t_0)) + \int_{t_0}^s L^0 f(x(r)) dr + \int_{t_0}^s L^1 f(x(r)) \eta_r dr \right) ds$$

$$\begin{aligned}
& + \int_{t_0}^t \left(\sigma(x(t_0)) + \int_{t_0}^s L^0 \sigma(x(r)) dr + \int_{t_0}^s L^1 \sigma(x(r)) \eta_r dr \right) \eta_s ds \\
& = x(t_0) + f(x(t_0)) \int_{t_0}^t ds + \sigma(x(t_0)) \int_{t_0}^t \eta_s ds + R
\end{aligned} \tag{7.5}$$

with the remainder

$$\begin{aligned}
R &= \int_{t_0}^t \int_{t_0}^s L^0 f(x(r)) dr ds + \int_{t_0}^t \int_{t_0}^s L^1 f(x(r)) \eta_r dr ds \\
&+ \int_{t_0}^t \int_{t_0}^s L^0 \sigma(x(r)) \eta_s dr ds + \int_{t_0}^t \int_{t_0}^s L^1 \sigma(x(r)) \eta_r \eta_s dr ds.
\end{aligned}$$

This is the simplest nontrivial Taylor expansion for the RODE with affine noise (7.3).

Continuing this procedure, then applying the chain rule (7.4) to $U = L^1 \sigma$ in (7.5) gives

$$\begin{aligned}
x(t) &= x(t_0) + f(x(t_0)) \int_{t_0}^t ds + \sigma(x(t_0)) \int_{t_0}^t \eta_s ds \\
&+ L^1 \sigma(x(t_0)) \int_{t_0}^t \int_{t_0}^s \eta_r \eta_s dr ds + \bar{R}
\end{aligned}$$

with remainder

$$\begin{aligned}
\bar{R} &= \int_{t_0}^t \int_{t_0}^s L^0 f(x(r)) dr ds + \int_{t_0}^t \int_{t_0}^s L^1 f(x(r)) \eta_r dr ds \\
&+ \int_{t_0}^t \int_{t_0}^s L^0 \sigma(x(r)) \eta_s dr ds + \int_{t_0}^t \int_{t_0}^s \int_{t_0}^r L^0 L^1 \sigma(x(\tau)) \eta_r \eta_s d\tau dr ds \\
&+ \int_{t_0}^t \int_{t_0}^s \int_{t_0}^r L^1 L^1 \sigma(x(\tau)) \eta_\tau \eta_r \eta_s d\tau dr ds.
\end{aligned}$$

The main properties of the Taylor expansions are already apparent in the preceding examples. In particular, the expansion has terms evaluated at $x(t_0)$ with the multiple Lebesgue integrals

$$\int_{t_0}^t ds, \quad \int_{t_0}^t \eta_s ds, \quad \int_{t_0}^t \int_{t_0}^s \eta_r \eta_s dr ds$$

and a remainder term involving the next multiple integrals, but with nonconstant integrands.

7.2 Affine-RODE Taylor Expansions

The nonautonomous d -dimensional RODE with m -dimensional affine noise in the integral form (7.2) can be rewritten compactly as

$$x(t) = x(t_0) + \sum_{j=0}^m \int_{t_0}^t f_j(s, x(s)) \eta_s^j \, ds \quad (7.6)$$

by introducing a fictitious “noise” component $\eta_t^0 \equiv 1$. This allows the notation adapted from stochastic Itô-Taylor expansions in Sect. 6.4 to be used here with some modifications due to the fact that the multiple integrals are now pathwise Riemann or Lebesgue integrals. These are defined recursively by

$$I_\alpha[u(\cdot)]_{t_0,t} := \begin{cases} u(t), & l = 0 \\ \int_{t_0}^t I_{\alpha-[u(\cdot)]_{t_0,s}} \eta_s^{j_l} \, ds, & l \geq 1 \end{cases}$$

where $\alpha- = (j_1, \dots, j_l)- = (j_1, \dots, j_{l-1})$ if $l(\alpha) = l \geq 1$ and $(j_1)- = \emptyset$. For example,

$$\begin{aligned} I_\emptyset[u(\cdot)]_{t_0,t} &= u(t), & I_{(0)}[u(\cdot)]_{t_0,t} &= \int_{t_0}^t u(s) \, ds \\ I_{(1)}[u(\cdot)]_{t_0,t} &= \int_{t_0}^t u(s) \eta_s^1 \, ds, \\ I_{(0,1)}[u(\cdot)]_{t_0,t} &= \int_0^t \int_0^{s_2} u(s_1) \eta_{s_2}^1 \, ds_1 \, ds_2 = \int_0^t \left[\int_0^{s_2} u(s_1) \, ds_1 \right] \eta_{s_2}^1 \, ds_2 \\ I_{(0,2,1)}[u(\cdot)]_{t_0,t} &= \int_0^t \left[\int_0^{s_3} \int_0^{s_2} u(s_1) \eta_{s_2}^2 \, ds_1 \, ds_2 \right] \eta_{s_3}^1 \, ds_3. \end{aligned}$$

For simpler notation, $I_\alpha[u(\cdot)]_{t_0,t}$ will often be abbreviated to $I_{\alpha,t_0,t}$ or just I_α when $u(t) \equiv 1$.

In addition, the coefficient functions F_α 's are defined similarly to those in Sect. 6.2.3 except that the partial differential operator L^0 is now defined by

$$L^0 = \frac{\partial}{\partial t} + \sum_{k=1}^d f_0^k \frac{\partial}{\partial x^k},$$

where f_0^k is the k th component of vector function f_0 , as in deterministic calculus.

Example 7.1 For the autonomous scalar RODE with scalar affine noise (7.3), i.e., $d = m = 1$, the coefficient functions corresponding to the identity function $F(t, x) \equiv x$ are

$$\begin{aligned} F_{(0)} &= f_0, & F_{(j_1)} &= f_{j_1}, & F_{(0,0)} &= f_0 f_0', \\ F_{(0,j_2)} &= f_0 f_{j_2}', & F_{(j_1,0)} &= f_0' f_{j_1}, & F_{(j_1,j_2)} &= f_{j_1} f_{j_2}', \end{aligned}$$

where the dash ' denotes differentiation with respect to x .

7.3 General Affine-RODE-Taylor Approximations

The affine-Taylor expansion corresponding to the hierarchical set \mathcal{H} and its remainder set $\mathcal{R}(\mathcal{H})$ (see Chap. 6 for the definitions) is given by

$$F(t, x(t)) = \sum_{\alpha \in \mathcal{H}} F_\alpha(t_0, x(t_0)) I_\alpha[1]_{t_0, t} + \sum_{\alpha \in \mathcal{R}(\mathcal{H})} I_\alpha [F_\alpha(\cdot, x(\cdot))]_{t_0, t}. \quad (7.7)$$

Affine-RODE-Taylor approximations of arbitrary higher order can be constructed by including in an appropriate way more terms from the Taylor expansions that are then truncated.

For a given hierarchical set \mathcal{H} the affine-RODE-Taylor expansion (7.7) of a function F is truncated to give the affine-RODE-Taylor approximation

$$\Phi_{\mathcal{H}}(t, x(t); F) = \sum_{\alpha \in \mathcal{H}} F_\alpha(t_0, x(t_0)) I_\alpha[1]_{t_0, t}.$$

In order to obtain an order of such approximations more needs to be assumed about the nature of the driving noise processes in the RODE with affine noise.

RODEs with Affine Bounded Noise

It will be shown here that an affine-RODE-Taylor approximation of order $\gamma = 1, 2, \dots$, for bounded measurable noise corresponds to hierarchical set

$$\Lambda_\gamma = \{\alpha \in \mathcal{J}_m : l(\alpha) \leq \gamma\}$$

needs all of the multiple integral terms with the constant coefficients $F_\alpha(t_0, x(t_0))$ from the Taylor expansion of up to and including order γ . Recall that the corresponding remainder set here is

$$\mathcal{R}(\Lambda_\gamma) = \{\alpha \in \mathcal{J}_m : l(\alpha) = \gamma + 1\}$$

and that the multiple integral corresponding to the multi-index $\alpha = (j_1, \dots, j_l)$ is

$$I_{\alpha, t_0, t_0+h} = \int_{t_0}^{t_0+h} \int_{t_0}^{s_l} \cdots \int_{t_0}^{s_2} \eta_{s_1}^{j_1} \cdots \eta_{s_{l-1}}^{j_{l-1}} \eta_{s_l}^{j_l} ds_1 \cdots ds_l.$$

Thus, in the general multi-dimensional case with $d, m = 1, 2, \dots$, the affine-RODE-Taylor approximation for $\gamma = 1, 2, 3, \dots$ is given by

$$\begin{aligned}\Phi_\gamma(t_0, x(t_0), h; F) &:= \sum_{\alpha \in \Lambda_\gamma} F_\alpha(t_0, x(t_0)) I_{\alpha, t_0, t_0+h} \\ &= F(t_0, x(t_0)) + \sum_{\alpha \in \Lambda_\gamma \setminus \{\emptyset\}} F_\alpha(t_0, x(t_0)) I_{\alpha, t_0, t_0+h}\end{aligned}$$

with the coefficient functions F_α corresponding to the function $F(t, x)$.

Note that when the function $F(t, x)$ is $\gamma + 1$ times continuously differentiable and the drift and noise coefficients f_0, f_1, \dots, f_m of the RODE with affine noise (7.6) are γ times continuously differentiable, then each of the integrals $I_\alpha[(F_\alpha(\cdot, x(\cdot)))]_{t_0, t_0+h}$, i.e.,

$$\int_{t_0}^{t_0+h} \int_{t_0}^{s_l} \cdots \int_{t_0}^{s_2} F_\alpha(s_1, x(s_1)) \eta_{s_1}^{j_1} \cdots \eta_{s_{l-1}}^{j_{l-1}} \eta_{s_l}^{j_l} ds_1 \cdots ds_l,$$

for α in the remainder set $\mathcal{R}(\Lambda_\gamma)$ is of order $h^{\gamma+1}$. Since there are only finitely many, specifically $(m+1)!$ remainder integrals, the (local) truncation error here is

$$|\Phi_\gamma(t_0, x(t_0), h) - F(t_0 + h, x(t_0 + h))| \leq C h^{\gamma+1},$$

where the constant C depends on γ as well as on a compact set containing the initial value $(t_0, x(t_0))$ and the sample paths of the solution of the RODE with affine noise. See Theorem 11.1 in Chap. 11.

Example 7.2 For the function $F(t, x) \equiv x^k$, the k th component of the vector x , and $\gamma = 1, 2$ and 3 , respectively, the solution $x(t_0 + h)$ of the affine RODE (7.6) satisfies the componentwise approximations

$$\begin{aligned}x^k(t_0 + h) &= x^k(t_0) + \sum_{j=0}^m f_j^k(t_0, x(t_0)) I_{(j)} + \mathcal{O}(h^2), \\ x^k(t_0 + h) &= x^k(t_0) + \sum_{j=0}^m f_j^k(t_0, x(t_0)) I_{(j)} + \sum_{j_1, j_2=0}^m L^{j_1} f_{j_2}^k I_{(j_1, j_2)} + \mathcal{O}(h^3) \\ x^k(t_0 + h) &= x^k(t_0) + \sum_{j=0}^m f_j^k(t_0, x(t_0)) I_{(j)} + \sum_{j_1, j_2=0}^m L^{j_1} f_{j_2}^k I_{(j_1, j_2)} \\ &\quad + \sum_{j_1, j_2, j_3=0}^m L^{j_1} L^{j_2} f_{j_3}^k(t_0, x(t_0)) I_{(j_1, j_2, j_3)} + \mathcal{O}(h^4),\end{aligned}$$

for $k = 1, \dots, d$, where $I_{(j)}$, $I_{(j_1, j_2)}$, $I_{(j_1, j_2, j_3)}$ stand for $I_{(j), t_0, t_0+h}$, $I_{(j_1, j_2), t_0, t_0+h}$, $I_{(j_1, j_2, j_3), t_0, t_0+h}$, respectively.

7.4 Endnotes

The Taylor expansions above are based on those in [55] for d -dimensional affinely nonlinear systems with m -dimensional control functions of \mathbb{R}^m of the form

$$\frac{dx}{dt} = f_0(t, x) + \sum_{j=1}^m f_j(t, x) u^j(t), \quad (7.8)$$

where the control functions $u(t) = (u^1(t), \dots, u^m(t))$ are measurable and take values in a compact convex subset of \mathbb{R}^m , and the differential equation (7.8) is interpreted in the Carathéodory sense. These Taylor expansions were motivated by Stratonovich-Taylor expansions which are similar to the Itô-Taylor expansions in the previous chapter, but without stochastic integrals. The results were specialized to switching systems in [56] of the form (7.8) with $f_0 \equiv 0$ and the “control” functions taking piecewise constant with values in $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m\}$, the unit vectors in \mathbb{R}^m , which corresponds to just one of the vector fields f_1, f_2, \dots, f_d being active at any one time. The material in this chapter is based on the papers Asai and Kloeden [10, 11].

Chapter 8

Taylor-Like Expansions for General Random Ordinary Differential Equations

The solutions of RODEs for which the driving stochastic process has Hölder continuous sample paths have continuously differentiable sample paths, but the derivatives of the sample paths are Hölder continuous in time and are not further differentiable. Thus such solutions do not have Taylor expansions in the usual sense. Nevertheless integral versions of Taylor expansions can be derived using a chain rule for such RODEs in a similar way to stochastic Taylor expansions for SODEs.

8.1 Preliminaries and Notation

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space and let $(\eta_t)_{t \geq 0}$ be an \mathbb{R}^m -valued stochastic process with continuous sample paths. In addition, let $g : \mathbb{R}^d \times \mathbb{R}^m \rightarrow \mathbb{R}^d$ be a continuous function and consider the random ordinary differential equation (RODE) in \mathbb{R}^d ,

$$\frac{dx}{dt} = g(x, \eta_t(\omega)). \quad (8.1)$$

For convenience, assume that g is infinitely often continuously differentiable in its variables, although k times continuously differentiable with k sufficiently large would suffice. In particular, if g is locally Lipschitz in x then an initial value problem for the RODE (8.1) has a unique pathwise solution $x(t, \omega)$ for every $\omega \in \Omega$, which will be assumed to exist on the finite time interval $[0, T]$ under consideration.

Let \mathbb{N} be the set of natural numbers and let \mathbb{N}_0 denote the set of nonnegative integers. For any vector $v = (v_1, \dots, v_d) \in \mathbb{R}^d$ define the norms¹

$$\|v\|_1 = |v_1| + \dots + |v_d|, \quad \|v\|_2 = \sqrt{v_1^2 + \dots + v_d^2}.$$

¹ $\|\cdot\|_2$ is the Euclidean norm of \mathbb{R}^d . Throughout this book $|\cdot|$ denotes the absolute value and also the Euclidean norm when no confusion can arise.

For any vector-valued function $u(t) = (u^1(t), \dots, u^d(t)) : [0, T] \rightarrow \mathbb{R}^d$ define the supremum norm

$$\|u\|_\infty := \sup_{s \in [0, T]} \|u(s)\|_2.$$

Recall that for a scalar $\beta > 0$, a scalar function $u : [0, T] \rightarrow \mathbb{R}$ has the Hölder norm defined as

$$|u|_\beta := \sup_{s \in [0, T]} |u(s)| + \sup_{t \neq s \in [0, T]} \frac{|u(t) - u(s)|}{|t - s|^\beta},$$

when it exists. For a vector-valued function $u(t) = (u^1(t), \dots, u^d(t)) : [0, T] \rightarrow \mathbb{R}^d$ and a vector $\beta = (\beta_1, \dots, \beta_d) \in (0, 1]^d$ the Hölder norm is defined to be

$$\|u\|_\beta := \max_{j=1, \dots, d} |u^j|_{\beta_j}.$$

The following assumptions, notation and abbreviations will be used.

8.1.1 Regularity of the Driving Stochastic Process

The driving stochastic process $\eta_t = (\eta_t^1, \dots, \eta_t^m)$ is assumed to have continuous sample paths, so the maximum can be used in the supremum norm $\|\eta\|_\infty = \sup_{s \in [0, T]} \|\eta_s\|_2$, which is in fact a random variable. The Hölder norm of η_t with exponents $(\beta_1, \dots, \beta_d)$ is defined pathwise as

$$\|\eta\|_\beta = \max_{j=1, \dots, m} |\eta^j|_{\beta_j},$$

where

$$|\eta^j|_{\beta_j} := \sup_{s \in [0, T]} |\eta_s^j| + \sup_{t \neq s \in [0, T]} \frac{|\eta_t^j - \eta_s^j|}{|t - s|^{\beta_j}},$$

is also a random variable.

The following standing assumption about the Hölder continuity of the driving stochastic process will allow different Hölder exponents of different component processes to be better exploited.

Assumption 8.1.1 There is a vector $\Theta = (\theta_1, \dots, \theta_m) \in (0, 1]^m$ such that

$$\|\eta\|_\beta < \infty$$

holds pathwise for all $0 < \beta < \Theta$, i.e., $0 < \beta_1 < \theta_1, \dots, 0 < \beta_m < \theta_m$.

Denote by $\vartheta := \min(\theta_1, \dots, \theta_m)$ the minimum of $\theta_1, \dots, \theta_m$.

8.1.2 Multi-index Notation

A multi-index which uses matrix-valued indices rather than vector-valued indices as for SODEs in the previous chapters will be used in this chapter.

For any nonempty set S and $m, n \in \mathbb{N}$, the set of all $m \times n$ matrices with entries in S , respectively the set of all n -tuples of m -tuples of S , is denoted by

$$S^{m \times n} := (S^m)^n = \left\{ ((a_1^1, \dots, a_1^m)^\top, \dots, (a_n^1, \dots, a_n^m)^\top) \mid a_i^j \in S, \begin{array}{l} 1 \leq i \leq n \\ 1 \leq j \leq m \end{array} \right\}.$$

In particular, consider the set of all $m \times i$ -matrices of nonnegative integers, $\mathbb{N}_0^{m \times i}$, with $i \geq 1$. Denote an element $\alpha = (\alpha_1, \dots, \alpha_i) \in \mathbb{N}_0^{m \times i}$ with $\alpha_j = (\alpha_j^1, \dots, \alpha_j^m) \in \mathbb{N}_0^m$ for $j = 1, \dots, i$ as

$$\alpha = (\alpha_1, \dots, \alpha_i) = \begin{pmatrix} \alpha_1^1 & \dots & \alpha_i^1 \\ \vdots & & \vdots \\ \alpha_1^m & \dots & \alpha_i^m \end{pmatrix}.$$

For any $\alpha \in \mathbb{N}_0^{m \times i}$, write $\iota(\alpha) := i$. In addition, write $\mathbb{N}_0^{m \times 0} = \mathbb{N}_0^0 := \{\emptyset\}$, where \emptyset is the empty set symbol, which will be used to denote the “empty” index. This will be used to label the constant, i.e., the first term in the RODE-Taylor expansions.

Denote the set of all such matrix-valued multi-indices by

$$\mathcal{M}_m := \bigcup_{i=0}^{\infty} \mathbb{N}_0^{m \times i}.$$

For any $\alpha \in \mathcal{M}_m \setminus \{\emptyset\}$ with $\iota(\alpha) = i \geq 1$, define

$$\begin{aligned} \alpha! &:= (\alpha_1^1)! \cdot \dots \cdot (\alpha_1^m)! \cdot \dots \cdot (\alpha_i^1)! \cdot \dots \cdot (\alpha_i^m)! \\ |\alpha| &:= \alpha_1^1 + \dots + \alpha_1^m + \dots + \alpha_i^1 + \dots + \alpha_i^m. \end{aligned} \tag{8.2}$$

Since α^\top is an $i \times m$ -matrix for $\alpha \in \mathcal{M}_m$ with $\iota(\alpha) \geq 1$ and $\Theta = (\theta_1, \dots, \theta_m)$ in Assumption 8.1.1 is an m -dimensional vector, it follows that

$$\alpha^\top \Theta = \begin{pmatrix} \alpha_1^1 \theta_1 + \dots + \alpha_1^m \theta_m \\ \vdots \\ \alpha_i^1 \theta_1 + \dots + \alpha_i^m \theta_m \end{pmatrix} \in \mathbb{R}_{\geq 0}^m,$$

and

$$\|\alpha^\top \Theta\|_1 = \theta_1 \sum_{j=1}^i \alpha_j^1 + \dots + \theta_m \sum_{j=1}^i \alpha_j^m.$$

Finally, for $\alpha = \emptyset$ define

$$\iota(\emptyset) := 0, \quad \emptyset! := 1, \quad |\emptyset| := 0, \quad \|\emptyset^\top \Theta\|_1 := 0.$$

8.1.3 Iterated Integrals

For a function $u = (u^1, \dots, u^m) : [0, \infty) \rightarrow \mathbb{R}^m$ and for $0 \leq t \leq s < \infty$ define

$$\Delta u_{t,s} := u(s) - u(t) = (\Delta u_{t,s}^1, \dots, \Delta u_{t,s}^m),$$

where

$$\Delta u_{t,s}^k = u^k(s) - u^k(t), \quad k = 1, \dots, m.$$

For $u \equiv 1$ write $\Delta_{t,s} := s - t$.

Given such a function and a matrix index $\alpha = (\alpha_1, \dots, \alpha_i) \in \mathcal{M}_m$ with $\iota(\alpha) = i \geq 1$, define

$$(\Delta u_{t,s})^{\alpha_j} := \prod_{k=1}^m (\Delta u_{t,s}^k)^{\alpha_j^k}, \quad j = 1, \dots, i.$$

This notation will be used only for the driving stochastic process η_t in iterated integrals, which are defined pathwise by

$$I_{\alpha,t,s} := \int_t^s \int_t^{s_1} \cdots \int_t^{s_{i-1}} [(\Delta \eta_{t,s_1})^{\alpha_1} \cdots (\Delta \eta_{t,s_i})^{\alpha_i}] \, ds_i \cdots ds_1$$

for $i = \iota(\alpha) \geq 1$ with $I_{\emptyset,t,s} := 1$. Obviously, $I_{\alpha,t,s}$ is a random variable.

8.1.4 Function Spaces

Let $\mathcal{V}_0 = \mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R}^d)$ be the vector space over \mathbb{R} of smooth functions from \mathbb{R}^d to \mathbb{R}^d . For $i \in \mathbb{N}$ let $\mathcal{V}_i = \mathcal{C}^\infty(\mathbb{R}^d \times \mathbb{R}^{m \times i}, \mathbb{R}^d)$ be the vector space over \mathbb{R} of all smooth functions from $\mathbb{R}^d \times \mathbb{R}^{m \times i}$ to \mathbb{R}^d . For example, the vector field g of the RODE (8.1) is in \mathcal{V}_1 . Let

$$\mathcal{V} := \bigcup_{i=0}^{\infty} \mathcal{V}_i.$$

Write the components of the variables $(x, \mathbf{w}) \in \mathbb{R}^d \times \mathbb{R}^{m \times i}$ of a function in \mathcal{V}_i with $i \geq 1$ as

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix}, \quad \mathbf{w} = (w_1, \dots, w_i) = \begin{pmatrix} w_1^1 & \dots & w_i^1 \\ \vdots & & \vdots \\ w_1^m & \dots & w_i^m \end{pmatrix}.$$

For a matrix multi-index $\alpha \in \mathcal{M}_m$ with $i = \iota(\alpha) \geq 1$ and a function $U \in \mathcal{V}_i$, define the α -derivative of U with respect to $\mathbf{w} \in \mathbb{R}^{m \times i}$ as

$$\partial_{\mathbf{w}}^{\alpha} U := (\partial_{w_1^1})^{\alpha_1^1} \dots (\partial_{w_1^m})^{\alpha_1^m} \dots (\partial_{w_i^1})^{\alpha_i^1} \dots (\partial_{w_i^m})^{\alpha_i^m} U$$

with $\partial^0 U = U$ for $U \in \mathcal{V}_0$. Obviously, $\partial^{\alpha} U \in \mathcal{V}_i$ when $U \in \mathcal{V}_i$ for any $i \geq 0$.

8.1.5 Iterated Differential Operators

Let $(X_t)_{t \in [0, T]}$ be the stochastic process which is a pathwise solution of the RODE (8.1) with the driving stochastic process $(\eta_t)_{t \geq 0}$. The definition of iterated differential operators of functions of this solution process requires appropriate total derivatives.

For $i \in \mathbb{N}_0$, and $U \in \mathcal{V}_i$, define the linear differential operator $L_i: \mathcal{V}_i \rightarrow \mathcal{V}_{i+1}$ by

$$L_i U(x, w_1, \dots, w_{i+1}) := \underbrace{\partial_x U(x, w_1, \dots, w_i)}_{d \times d - \text{matrix}} \cdot \underbrace{g(x, w_{i+1})}_{\text{vector in } \mathbb{R}^d}$$

for $x \in \mathbb{R}^d$ and $w_1, \dots, w_{i+1} \in \mathbb{R}^m$ (for $i = 0$ omit the w_1, \dots, w_i variables). The right hand side here is the product of a $d \times d$ -matrix and a d -dimensional vector.

There is an analogue of the fundamental theorem of calculus for these differential operators and functions in \mathcal{V} .

Lemma 8.1 *Let $U \in \mathcal{V}_i$ for some $i \in \mathbb{N}_0$ and $0 \leq t_0 \leq t \leq T$. Then,*

$$U(X_t(\omega), w_1, \dots, w_i) = U(X_{t_0}(\omega), w_1, \dots, w_i) + \int_{t_0}^t L_i U(X_s(\omega), \eta_s(\omega), w_1, \dots, w_i) \, ds$$

for all $w_1, \dots, w_i \in \mathbb{R}^m$ and each $\omega \in \Omega$. For $i = 0$, omit the w_1, \dots, w_i variables.

Proof Fix any $\omega \in \Omega$, consider the function $u: [0, T] \rightarrow \mathbb{R}^d$ given by

$$u(s) := U(X_s(\omega), w_1, \dots, w_i), \quad s \in [0, T],$$

which is continuously differentiable with the derivative

$$\begin{aligned} u'(s) &= \partial_x U(X_s(\omega), w_1, \dots, w_i) \cdot g(X_s(\omega), \eta_s(\omega)) \\ &= L_i U(X_s(\omega), \eta_s(\omega), w_1, \dots, w_i). \end{aligned}$$

The assertion thus follows by applying the fundamental theorem of calculus to the function u . \square

More generally, define the differential operator $L:\mathcal{V} \rightarrow \mathcal{V}$ by

$$LU := L_i U \quad \text{if } U \in \mathcal{V}_i \text{ for some } i \in \mathbb{N}_0.$$

Note that if $U \in \mathcal{V}_i$ for some $i \in \mathbb{N}_0$, then $L(U) \in \mathcal{V}_{i+1}$ and $L(L(U)) \in \mathcal{V}_{i+2}$, and so on. Thus, the L -operator can be iterated:

$$L^j := \underbrace{L \circ L \circ \cdots \circ L}_{j \text{ times}} : \mathcal{V}_i \rightarrow \mathcal{V}_{i+j}$$

$$\text{as } \mathcal{V}_i \xrightarrow{L} \mathcal{V}_{i+1} \xrightarrow{L} \cdots \xrightarrow{L} \mathcal{V}_{i+j}.$$

8.2 Integral Equation Expansions

The above iterated integrals and differential operators can be used to derive an integral equation expansion of the solution of the RODE (8.1), which will be the key to developing the RODE-Taylor approximations in the next section. The underlying idea is to use Lemma 8.1 iteratively.

Instead of a general $U \in \mathcal{V}_0$, attention is restricted here to the identity function $\chi: \mathbb{R}^d \rightarrow \mathbb{R}^d$, i.e., $\chi(x) = x$ for all $x \in \mathbb{R}^d$. It is clear that $\chi \in \mathcal{V}_0$, so Lemma 8.1 can be applied with $U = \chi$ to obtain

$$x(t) = x_0 + \int_{t_0}^t L\chi(x(s), \eta_s) \, ds, \quad (8.3)$$

which is just an integral version of the RODE (8.1) with $x(t_0) = x_0$, since $L\chi = L_0\chi = g$.

Lemma 8.1 is then applied to the integrand $L\chi \in \mathcal{V}_1$ over the interval $[t_0, s_1]$ to obtain

$$L\chi(x(s_1), \eta_{s_1}) = L\chi(x_0, \eta_{s_1}) + \int_{t_0}^{s_1} L^2\chi(x(s_2), \eta_{s_1}, \eta_{s_2}) \, ds_2.$$

Inserting this into the Eq. (8.3) gives

$$x(t) = x_0 + \int_{t_0}^t L\chi(x_0, \eta_{s_1}) \, ds_1 + \int_{t_0}^t \int_{t_0}^{s_1} L^2\chi(x(s_2), \eta_{s_1}, \eta_{s_2}) \, ds_2 \, ds_1.$$

Iterating this process yields an *integral equation expansion* of the solution of the RODE (8.1) with $x(t_0) = x_0$:

$$\begin{aligned} x(t) = x_0 + \sum_{j=1}^p \int_{t_0}^t \dots \int_{t_0}^{s_{j-1}} L^j \chi(x_0, \eta_{s_1}, \dots, \eta_{s_j}) ds_j \dots ds_1 \\ + \int_{t_0}^t \dots \int_{t_0}^{s_p} L^{p+1} \chi(x(s_{p+1}), \eta_{s_1}, \dots, \eta_{s_{p+1}}) ds_{p+1} \dots ds_1, \end{aligned} \quad (8.4)$$

for each $p \in \mathbb{N}_0$. Here $t_0, t \in [0, T]$ and $t_0 \leq t$.

Example 8.1 For $p = 0$ the integral equation approximation (8.4) reduces to

$$x(t) = x_0 + \int_{t_0}^t g(x(s), \eta_s) ds,$$

i.e., the original RODE (8.1). For $p = 2$

$$\begin{aligned} x(t) = x_0 + \int_{t_0}^t g(x_0, \eta_{s_1}) ds_1 + \int_{t_0}^t \int_{t_0}^{s_1} g_x(x_0, \eta_{s_1}) g(x_0, \eta_{s_2}) ds_2 ds_1 \\ + \int_{t_0}^t \int_{t_0}^{s_1} \int_{t_0}^{s_2} g_{xx}(x(s_3), \eta_{s_1}) (g(x(s_3), \eta_{s_2}), g(x(s_3), \eta_{s_3})) ds_3 ds_2 ds_1 \\ + \int_{t_0}^t \int_{t_0}^{s_1} \int_{t_0}^{s_2} g_x(x(s_3), \eta_{s_1}) (g_x(x(s_3), \eta_{s_2}) g(x(s_3), \eta_{s_3})) ds_3 ds_2 ds_1, \end{aligned}$$

where g_x and g_{xx} are abbreviations for the partial derivatives $\partial_x g$ and $\partial_x^2 g$, respectively. Note that

$$(g(x(s_3), \eta_{s_2}), g(x(s_3), \eta_{s_3})) = \sum_{k=1}^d g^k(x(s_3), \eta_{s_2}) g^k(x(s_3), \eta_{s_3}),$$

and the expression

$$g_x(x(s_3), \eta_{s_1}) (g_x(x(s_3), \eta_{s_2}) g(x(s_3), \eta_{s_3}))$$

is equivalent to

$$\sum_{k=1}^d \sum_{l=1}^d \partial_{x_k} g(x(s_3), \eta_{s_1}) \cdot \partial_{x_l} g^k(x(s_3), \eta_{s_2}) \cdot g^l(x(s_3), \eta_{s_3}),$$

where g^k and g^l for $k, l \in \mathbb{N}$ are the k th and l th component of g , respectively.

8.3 RODE-Taylor Approximations

The implicit integral equation expansions (8.4) of the solution of the RODE (8.1) can be simplified further by using Assumption 8.1.1 that the driving stochastic process $(\eta_t)_{t \geq 0}$ is Hölder continuous.

Since $L^i \chi: \mathbb{R}^d \times \mathbb{R}^{m \times i} \rightarrow \mathbb{R}^d$ is in \mathcal{V}_i , it can be approximated by a Taylor approximation in its last $m \times i$ variables. These will be used in the integral equation expansion (8.4) to construct an explicit temporal approximation of the solution of the RODE. The resulting approximations will be called RODE-Taylor approximations.

For $\alpha \in \mathcal{M}_m$, use the abbreviation

$$F_\alpha = \frac{1}{\alpha!} (\partial^\alpha L^{\iota(\alpha)} \chi) \in \mathcal{V}_{\iota(\alpha)}.$$

For a vector $v = (v_1, \dots, v_d) \in \mathbb{R}^d$ with $d \in \mathbb{N}$, denote by $v^{\times n}$ for $n \in \mathbb{N}$ the $d \times n$ matrix

$$v^{\times n} := \underbrace{(v, \dots, v)}_{n \text{ times}} = \begin{pmatrix} v_1 & \dots & v_1 \\ \vdots & & \vdots \\ v_d & \dots & v_d \end{pmatrix}.$$

In particular, consider the specific subsets of matrix multi-indices of the form

$$\mathcal{M}_m^\gamma := \{ \alpha \in \mathcal{M}_m : \iota(\alpha) + \|\alpha^\top \Theta\|_1 < \gamma + 1 \},$$

where $\gamma > 0$. For such a γ define the function $\Phi_\gamma: \mathbb{R}^d \times [0, T] \times [0, 1] \times \Omega \rightarrow \mathbb{R}^d$ by

$$\Phi_\gamma(x, t, h, \omega) := \sum_{\alpha \in \mathcal{M}_m^\gamma} F_\alpha(x, \eta_t(\omega)^{\times \iota(\alpha)}) \cdot I_{\alpha, t, t+h}(\omega).$$

The pathwise γ -RODE-Taylor approximation about (t, x) is then defined by

$$x^{(h)}(\omega) = \Phi_\gamma(x, t, h, \omega).$$

To facilitate error estimates, $x^{(h)}(\omega)$ can be extended to a continuous time stochastic process $(\bar{x}_s^{(h)})_{s \in [t, t+h]}$ defined by

$$\bar{x}_s^{(h)}(\omega) = \Phi_\gamma(x, t, s - t, \omega), \quad s \in [t, t + h].$$

This process obviously has continuous sample paths and is also non-anticipative if the stochastic process η_t is non-anticipative.

The following pathwise local truncation error estimate will be proved later in Sect. 8.6. Here $\lceil \rho \rceil$ is the smallest integer $n \geq \rho$.

Theorem 8.1 (Truncation error of the RODE-Taylor expansions) *Let $\gamma > 0$ and let $\varepsilon = (\varepsilon_1, \dots, \varepsilon_m) \in [0, 1]^m$ with $\|\varepsilon\|_1 = \sum_{j=1}^m \varepsilon_j < \gamma$. Then,*

$$\sup_{0 \leq t \leq T} \|x(t, \omega) - \bar{x}_t^{(h)}(\omega)\|_2 \leq K_\gamma^\varepsilon(\omega) \cdot h^{(\gamma+1-\|\varepsilon\|_1)}$$

for all $0 < h \leq h_0(\omega)$ and all $\omega \in \Omega$ with the nonnegative random variables

$$K_\gamma^\varepsilon = (C_\gamma^\varepsilon T) e^{\kappa_\gamma T}, \quad h_0 = (K_\gamma^\varepsilon)^{1/(\|\varepsilon\|_1 - \gamma)} \wedge 1,$$

where $\beta := \Theta - \varepsilon / \lceil \frac{\gamma}{\vartheta} \rceil$, $i := \iota(\alpha)$,

$$C_\gamma^\varepsilon := \sum_{\substack{\alpha \in \mathcal{M}_m \setminus \mathcal{M}_m^\gamma \\ \iota(\alpha) \leq \lceil \gamma + 1 \rceil \\ |\alpha| \leq \lceil \frac{\gamma}{\vartheta} \rceil}} \frac{(\|\eta\|_\beta)^{|\alpha|}}{\iota(\alpha)!} \sup_{\substack{\|\omega_1\|_2, \dots, \|\omega_i\|_2 \leq \|\eta\|_\infty \\ \|y\|_2 \leq \|x\|_\infty}} \|F_\alpha(y, \omega_1, \dots, \omega_i)\|_2,$$

$$\kappa_\gamma := \sum_{\alpha \in \mathcal{M}_m^\gamma \setminus \{\emptyset\}} (2\|\eta\|_\infty)^{|\alpha|} \left(\sup_{0 \leq t \leq T} \sup_{\|y - x(t, \omega)\| \leq 1} \|\partial_x F_\alpha(y, \eta_t^{\times \iota(\alpha)})\|_2 \right),$$

and $|\alpha|$ is defined as in (8.2).

Suppose that a stochastic process $(\eta_t)_{t \geq 0}$ with $\Theta = (\theta_1, \dots, \theta_m)$ and the desired order $\gamma > 0$ are given. Then $\varepsilon \in [0, 1]^m$ should be chosen so that $\|\varepsilon\|_1 < \gamma$ and $\varepsilon_j = 0$ iff $|\eta^j|_{\theta_j} < \infty$ for $j = 1, \dots, m$. In this case, $K_\gamma^\varepsilon < \infty$.

8.4 Essential RODE-Taylor Approximations

The above γ -RODE-Taylor approximation works for every $\gamma > 0$, but in practice only some specific $\gamma > 0$ are important. Consider the set

$$\begin{aligned} \Lambda_\Theta &:= \mathbb{N}_0 + \theta_1 \mathbb{N}_0 + \cdots + \theta_m \mathbb{N}_0 \\ &= \{k + l_1 \theta_1 + \cdots + l_m \theta_m : k, l_1, \dots, l_m \in \mathbb{N}_0\}. \end{aligned}$$

It is clear that Λ_Θ coincides with a sequence $0 = \gamma_0 < \gamma_1 < \gamma_2 < \dots$

Lemma 8.2 below states that if $\gamma \in (\gamma_n, \gamma_{n+1}]$ for some $n \in \mathbb{N}_0$, then the γ -RODE-Taylor approximation is the same as the γ_{n+1} -RODE-Taylor approximation. Hence, only the γ_n -RODE-Taylor approximations for $n \in \mathbb{N}$ need to be considered. For $\gamma > 0$ define

$$[\gamma] = [\gamma]_\Theta := \min_{\substack{\gamma_n \geq \gamma \\ \gamma_n \in \Lambda_\Theta}} \gamma_n.$$

Clearly, $\gamma \leq [\gamma]$ and $[\gamma] \in \Lambda_\Theta$.

Lemma 8.2 Let $\gamma, \tilde{\gamma} > 0$. Then $\mathcal{M}_m^\gamma = \mathcal{M}_m^{\tilde{\gamma}}$ if and only if $[\gamma] = [\tilde{\gamma}]$. In particular, $\mathcal{M}_m^\gamma = \mathcal{M}_m^{[\gamma]}$ and $\Phi_\gamma = \Phi_{[\gamma]}$.

Proof Assume without loss of generality that $\tilde{\gamma} \geq \gamma$.

(\Leftarrow) Suppose (for contradiction) that there exists an $\alpha \in \mathcal{M}_m^{\tilde{\gamma}} \setminus \mathcal{M}_m^\gamma$ such that

$$\gamma + 1 \leq \iota(\alpha) + \|\alpha^\top \Theta\|_1 < \tilde{\gamma} + 1.$$

Obviously, $\iota(\alpha) \geq 1$, thus

$$\gamma \leq \underbrace{(\iota(\alpha) - 1) + \|\alpha^\top \Theta\|_1}_{\in \Lambda_\Theta} < \tilde{\gamma}.$$

This implies that

$$[\gamma] \leq \underbrace{(\iota(\alpha) - 1) + \|\alpha^\top \Theta\|_1}_{\in \Lambda_\Theta} < \tilde{\gamma} \leq [\tilde{\gamma}] = [\gamma].$$

which is a contradiction.

(\Rightarrow) Suppose (for contradiction) that $[\gamma] < [\tilde{\gamma}]$. Then there exist $a, b_1, \dots, b_m \in \mathbb{N}_0$ such that

$$\gamma \leq [\gamma] = a + b_1\theta_1 + \dots + b_m\theta_m < [\tilde{\gamma}]$$

and furthermore,

$$[\gamma] < \tilde{\gamma} \leq [\tilde{\gamma}].$$

Now define

$$\alpha = ((\alpha_1^1, \dots, \alpha_1^m), \dots, (\alpha_i^1, \dots, \alpha_i^m)) \in \mathbb{N}^{m \times i}$$

with $i = a + 1 \in \mathbb{N}$ and

$$\begin{aligned} \alpha_1^1 &= b_1, & \alpha_1^2 &= b_2, \dots, \alpha_1^m &= b_m, \\ \alpha_2^1 &= \dots = \alpha_2^m = 0, & \dots, & \alpha_i^1 &= \dots = \alpha_i^m = 0. \end{aligned}$$

Then

$$\begin{aligned} \gamma + 1 &\leq \underbrace{\iota(\alpha) + \|\alpha^\top \Theta\|_1}_{=1+a+b_1\theta_1+\dots+b_m\theta_m=1+[\gamma]} && < \tilde{\gamma} + 1, \end{aligned}$$

which implies that α is in $\mathcal{M}_m^{\tilde{\gamma}}$, but not in \mathcal{M}_m^γ . This is a contradiction. \square

8.5 Examples

Three examples of the RODE-Taylor approximations for the solutions to the RODE

$$\frac{dx}{dt} = g(x, \eta_t(\omega)), \quad x(t_0) = x_0,$$

are presented here. In particular, for $0 \leq t_0 < t_0 + h \leq T$ and $x_0 \in \mathbb{R}^d$, the function $\Phi_\gamma(x_0, t_0, h)$ of the γ -RODE-Taylor approximation with time step $h \in (0, 1]$ is determined for different choice of γ, Θ, d and m .

Recall that $\Delta\eta_{t_0,s} = \eta_s - \eta_{t_0}$, $\Delta_{t_0,s} = s - t_0$ and denote

$$I_a := I_{a,t_0,t_0+h} = \int_{t_0}^{t_0+h} \int_{t_0}^{s_1} \cdots \int_{t_0}^{s_{i-1}} [(\Delta\eta_{t_0,s_1})^{a_1} \cdots (\Delta\eta_{t_0,s_i})^{a_i}] ds_i \cdots ds_1,$$

for $a = (a_1, \dots, a_i)$. In addition, the a -derivative of function g with respect to $w \in \mathbb{R}^{m \times i}$ is given by

$$\partial_w^a g = (\partial_{w_1^1})^{a_1^1} \cdots (\partial_{w_1^m})^{a_1^m} \cdots (\partial_{w_i^1})^{a_i^1} \cdots (\partial_{w_i^m})^{a_i^m} g.$$

All coefficient functions in the examples below are evaluated at (x_0, η_{t_0}) .

Case I: $\gamma = 0.6, 1.6, 2.1$, $\Theta = 0.3$, $d = m = 1$

This noise process could be a fractional Brownian motion with Hurst parameter $H = 0.3$. Here

$$\begin{aligned} A_\Theta &= \mathbb{N}_0 + \theta_1 \mathbb{N}_0 \\ &= \{0, 0.3, 0.6, 0.9, 1, 1.2, 1.3, 1.5, 1.6, 1.8, 1.9, 2.0, 2.1, 2.2, 2.3, \dots\}, \end{aligned}$$

which implies that $\gamma_2 = 0.6$, $\gamma_8 = 1.6$ and $\gamma_{12} = 2.1$. The corresponding multi-index sets are

$$\begin{aligned} \mathcal{M}_1^{0.6} &= \{\emptyset, (0), (1)\}, \\ \mathcal{M}_1^{1.6} &= \{\emptyset, (0), (1), (2), (3), (4), (5), (0, 0), (1, 0), (0, 1)\} \\ \mathcal{M}_1^{2.1} &= \left\{ \begin{array}{l} \emptyset, (0), (1), (2), (3), (4), (5), (6), (0, 0), (1, 0), (0, 1), \\ (1, 1), (2, 0), (0, 2), (3, 0), (0, 3), (2, 1), (1, 2), (0, 0, 0) \end{array} \right\}. \end{aligned}$$

In this scalar context the 0.6-RODE-Taylor approximation reads

$$\Phi_{0.6}(x_0, t_0, h) = x_0 + hg + \partial_w g \int_{t_0}^t \Delta\eta_{t_0,s} ds,$$

the 1.6-RODE-Taylor approximation reads

$$\begin{aligned}\Phi_{1.6}(x_0, t_0, h) = & x_0 + hg + \sum_{i=1}^5 \frac{1}{i!} \partial_w^i g I_{(i)} + \frac{1}{2} (\partial_x g) g h^2 \\ & + (\partial_x \partial_w g) g I_{(1,0)} + (\partial_x g) (\partial_w g) I_{(0,1)},\end{aligned}$$

and the 2.1-RODE-Taylor approximation reads

$$\begin{aligned}\Phi_{2.1}(x_0, t_0, h) = & \Phi_{1.6}(x_0, t_0, h) + \frac{1}{6!} (\partial_w^6 g) g I_{(6)} + (\partial_x \partial_w g) (\partial_w g) I_{(1,1)} \\ & + \frac{1}{2} (\partial_x \partial_w^2 g) g I_{(2,0)} + \frac{1}{2} (\partial_x g) (\partial_w^2 g) I_{(0,2)} \\ & + \frac{1}{2} (\partial_x \partial_w^2 g) (\partial_w g) I_{(2,1)} + \frac{1}{2} (\partial_x \partial_w g) (\partial_w^2 g) I_{(1,2)} \\ & + \frac{1}{6} (\partial_x \partial_w^3 g) g I_{(3,0)} + \frac{1}{6} (\partial_x g) (\partial_w^3 g) I_{(0,3)} \\ & + \frac{1}{6} (\partial_x g)^2 g h^3 + \frac{1}{6} (\partial_x^2 g) g^2 h^3.\end{aligned}$$

Case II: $\gamma = 3$, $\Theta = \frac{1}{2}$, $d = m = 1$

The RODE and the driving stochastic process η_t are 1-dimensional and the noise process could be a Brownian motion, i.e., Wiener process. Here

$$\Lambda_\Theta = \mathbb{N}_0 + \frac{1}{2} \mathbb{N}_0 = \{0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, \dots\}.$$

Thus $\gamma = \gamma_6 = 3.0 \in \Lambda_\Theta$. The multi-index $\mathcal{M}_1^3 = \{2\iota(\alpha) + \|\alpha\|_1 \leq 7\}$ is given by

$$\mathcal{M}_1^3 = \left\{ \begin{array}{l} \emptyset, (0), (1), (2), (3), (4), (5)(0, 0), (1, 0), (0, 1), (1, 1), (2, 0), (0, 2), \\ (2, 1), (1, 2), (3, 0), (0, 3), (0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1). \end{array} \right\}.$$

The 3.0-RODE-Taylor approximation reads

$$\begin{aligned}\Phi_{3.0}(x_0, t_0, h) = & x_0 + hg + \sum_{i=1}^5 \frac{1}{i!} (\partial_w^i g) I_{(i)} + \frac{1}{2} (\partial_x g) gh^2 + (\partial_x \partial_w g) g I_{(1,0)} \\ & + (\partial_x g) (\partial_w g) I_{(0,1)} + (\partial_x \partial_w g) (\partial_w g) I_{(1,1)} + \frac{1}{2} (\partial_x \partial_w^2 g) g I_{(2,0)} \\ & + \frac{1}{2} (\partial_x g) (\partial_w^2 g) I_{(0,2)} + \frac{1}{2} (\partial_x \partial_w^2 g) (\partial_w g) I_{(2,1)} \\ & + \frac{1}{2} (\partial_x \partial_w g) (\partial_w^2 g) I_{(1,2)} + \frac{1}{6} (\partial_x \partial_w^3 g) g I_{(3,0)} \\ & + \frac{1}{6} (\partial_x g) (\partial_w^3 g) I_{(0,3)} + \frac{1}{6} (\partial_x g)^2 gh^3 + \frac{1}{6} (\partial_x^2 g) g^2 h^3\end{aligned}$$

$$\begin{aligned}
& + \left[(\partial_x \partial_w g) (\partial_x g) g + (\partial^2 x \partial_w g) g^2 \right] I_{(1,0,0)} \\
& + \left[(\partial_x g) (\partial_x \partial_w g) g + (\partial_x^2 g) (\partial_w g) g \right] I_{(0,1,0)} \\
& + \left[(\partial_x g)^2 (\partial_w g) + (\partial_x^2 g)^2 g (\partial_w g) \right] I_{(0,0,1)},
\end{aligned}$$

where the integrals over the interval $[t_0, t_0 + h]$ are given by

$$I_{(i)} = \int_{t_0}^{t_0+h} (\Delta \eta_{t_0,t})^i dt \quad \text{for } i = 1, \dots, 6$$

and

$$\begin{aligned}
I_{(1,0)} &= \int_{t_0}^{t_0+h} \Delta \eta_{t_0,t} \Delta_{t_0,t} dt, & I_{(0,1)} &= \int_{t_0}^{t_0+h} \int_{t_0}^s \Delta \eta_{t_0,t} dt ds, \\
I_{(1,1)} &= \int_{t_0}^{t_0+h} \int_{t_0}^s \Delta \eta_{t_0,s} \Delta \eta_{t_0,t} dt ds, & I_{(2,0)} &= \int_{t_0}^{t_0+h} (\Delta \eta_{t_0,t})^2 \Delta_{t_0,t} dt, \\
I_{(0,2)} &= \int_{t_0}^{t_0+h} \int_{t_0}^s (\Delta \eta_{t_0,t})^2 dt ds, & I_{(2,1)} &= \int_{t_0}^{t_0+h} \int_{t_0}^s (\Delta \eta_{t_0,s})^2 \Delta \eta_{t_0,t} dt ds, \\
I_{(1,2)} &= \int_{t_0}^{t_0+h} \int_{t_0}^s \Delta \eta_{t_0,s} (\Delta \eta_{t_0,t})^2 dt ds, & I_{(3,0)} &= \int_{t_0}^{t_0+h} (\Delta \eta_{t_0,t})^3 \Delta_{t_0,t} dt, \\
I_{(0,3)} &= \int_{t_0}^{t_0+h} \int_{t_0}^s (\Delta \eta_{t_0,t})^3 dt ds, & I_{(1,0,0)} &= \frac{1}{2} \int_{t_0}^{t_0+h} \Delta \eta_{t_0,t} (\Delta_{t_0,t})^2 dt, \\
I_{(0,1,0)} &= \int_{t_0}^{t_0+h} \int_{t_0}^s \Delta \eta_{t_0,t} \Delta_{t_0,t} dt ds, & I_{(0,0,1)} &= \int_{t_0}^{t_0+h} \int_{t_0}^s \int_{t_0}^t \Delta \eta_{t_0,\tau} d\tau dt ds.
\end{aligned}$$

Case III: $\gamma = 1.5$, $\Theta = (\frac{1}{2}, \frac{3}{4})$, $d = m = 2$

Here the RODE with $g = (g^1, g^2)$ and the driving stochastic process $\eta_t = (\eta_t^1, \eta_t^2)$ are both 2-dimensional.

In view of the choice of Θ , the first component of the noise process η_t^1 could be a Wiener process, while the second component η_t^2 could be a fractional Brownian motion with Hurst index $H = \frac{3}{4}$. Note that

$$\Lambda_\Theta = \mathbb{N}_0 + \frac{1}{2} \mathbb{N}_0 + \frac{3}{4} \mathbb{N}_0 = \left\{ 0, \frac{1}{2}, \frac{3}{4}, 1, \frac{5}{4}, \frac{3}{2}, \frac{7}{4}, 2, \frac{9}{4}, \frac{5}{2}, \dots \right\},$$

then a representative choice is $\gamma = \gamma_5 = 1.5$ with the multi-index set

$$\begin{aligned}
\mathcal{M}_2^{1.5} &= \left\{ \mathbf{a} \in \mathcal{M}_2 : \iota(\mathbf{a}) + \left\| \mathbf{a}^\top \binom{1/2}{3/4} \right\|_1 \leq \frac{9}{4} \right\} \\
&= \left\{ \emptyset, \binom{0}{0}, \binom{0}{1}, \binom{1}{0}, \binom{1}{1}, \binom{2}{0}, \binom{0 0}{0 0} \right\}.
\end{aligned}$$

For these multi-indices the coefficient functions are given by

$$F_\emptyset(x) = x, \quad F_{\binom{i}{j}}(x, \mathbf{w}) = \frac{1}{i!j!} \partial_{w_2}^j \partial_{w_1}^i g(x, \mathbf{w}),$$

$$F_{\binom{0}{0}}(x, \mathbf{w}) = \sum_{j=1}^2 \partial_{x_j} g(x, \mathbf{w}) g^j(x, \mathbf{w}),$$

and the integrals are given by

$$I_{\binom{i}{j}} = \int_{t_0}^{t_0+h} (\Delta \eta_{t_0,t}^1)^i (\Delta \eta_{t_0,t}^2)^j dt, \quad I_{\binom{0}{0}} = \frac{1}{2} h^2.$$

Hence the 1.5-RODE-Taylor approximation reads

$$\begin{aligned} \Phi_{1.5}(x_0, t_0, h) &= x_0 + hg + \partial_{w_1} g \int_{t_0}^{t_0+h} \Delta \eta_{t_0,t}^1 dt + \partial_{w_2} g \int_{t_0}^{t_0+h} \Delta \eta_{t_0,t}^2 dt \\ &\quad + \partial_{w_1} \partial_{w_2} g \int_{t_0}^{t_0+h} \Delta \eta_{t_0,t}^1 \Delta \eta_{t_0,t}^2 dt + \frac{1}{2} \partial_{w_1}^2 g \int_{t_0}^{t_0+h} (\Delta \eta_{t_0,t}^1)^2 dt \\ &\quad + \frac{1}{2} ((\partial_{x_1} g) g^1 + (\partial_{x_2} g) g^2) h^2. \end{aligned}$$

8.6 Proof of Theorem 8.1

The proof of Theorem 8.1 will be done via the following three lemmata.

Lemma 8.3 *Let $p \in \mathbb{N}_0$, $i \in \mathbb{N}$, and $\mathfrak{a} = (\mathfrak{a}_1, \dots, \mathfrak{a}_i)$. Then*

$$\begin{aligned} L^i \chi(x_0, \eta_{s_1}, \dots, \eta_{s_i}) &= \sum_{\substack{|\mathfrak{a}| \leq p \\ \mathfrak{a} \in \mathbb{N}_0^{m \times i}}} F_{\mathfrak{a}}(x_0, \eta_{t_0}^{\times i}) \cdot (\Delta \eta_{t_0,s_1})^{\mathfrak{a}_1} \cdots (\Delta \eta_{t_0,s_i})^{\mathfrak{a}_i} \\ &\quad + \sum_{\substack{|\mathfrak{a}| = p+1 \\ \mathfrak{a} \in \mathbb{N}_0^{m \times i}}} \left[\int_0^1 F_{\mathfrak{a}}(x_0, \eta_{t_0} + \lambda \Delta \eta_{t_0,s_1}, \dots, \eta_{t_0} + \lambda \Delta \eta_{t_0,s_i}) \cdot (1-\lambda)^p d\lambda \right] \\ &\quad \times (p+1) \cdot (\Delta \eta_{t_0,s_1})^{\mathfrak{a}_1} \cdots (\Delta \eta_{t_0,s_i})^{\mathfrak{a}_i} \end{aligned}$$

holds for all $t_0, s_1, \dots, s_i \in [0, T]$.

Proof Fix $\omega \in \Omega$. Applying a Taylor approximation of order p to the function $u: [0, 1] \rightarrow \mathbb{R}^d$

$$u(\lambda) := L^i \chi(x_0(\omega), \eta_{t_0}(\omega) + \lambda \Delta \eta_{t_0,s_1}(\omega), \dots, \eta_{t_0}(\omega) + \lambda \Delta \eta_{t_0,s_i}(\omega))$$

results in

$$u(1) = u(0) + u'(0) + \dots + \frac{u^{(p)}(0)}{p!} + \int_0^1 u^{(p+1)}(\lambda) \frac{(1-\lambda)^p}{p!} d\lambda,$$

which is the assertion. \square

An estimate for the integrand in $I_{\alpha,t_0,t}$ is also required.

Lemma 8.4 *Let $t_0 \leq t$, $[t_0, t] \subset [0, T]$, and $\beta = (\beta_1, \dots, \beta_m)^\top \in (0, 1]^m$. Let $i = \iota(\alpha) \geq 1$. Then*

$$|(\Delta \eta_{t_0, s_1})^{\alpha_1} \cdots (\Delta \eta_{t_0, s_i})^{\alpha_i}| \leq (\|\eta\|_\beta)^{|\alpha|} \cdot (\Delta_{t_0, t})^{\|\alpha^\top \beta\|_1}$$

for all $s_1, \dots, s_i \in [t_0, t]$ and $\alpha \in \mathcal{M}_m$ with $\iota(\alpha) \geq 1$.

Proof Consider j in $\{1, \dots, \iota(\alpha)\}$. Then

$$\begin{aligned} |(\Delta \eta_{t_0, s_j})^{\alpha_j}| &= \left| (\Delta \eta_{t_0, s_1}^1)^{\alpha_j^1} \right| \cdots \left| (\Delta \eta_{t_0, s_j}^m)^{\alpha_j^m} \right| \\ &\leq \left(|\eta^1|_{\beta_1} (\Delta_{t_0, s_1})^{\beta_1} \right)^{\alpha_j^1} \cdots \left(|\eta^m|_{\beta_m} (\Delta_{t_0, s_j})^{\beta_m} \right)^{\alpha_j^m} \\ &\leq (\|\eta\|_\beta)^{\alpha_j^1 + \dots + \alpha_j^m} (\Delta_{t_0, t})^{\alpha_j^1 \beta_1 + \dots + \alpha_j^m \beta_m}. \end{aligned}$$

In addition with $i = \iota(\alpha)$,

$$|(\Delta \eta_{t_0, s_1})^{\alpha_1} \cdots (\Delta \eta_{t_0, s_i})^{\alpha_i}| = |(\Delta \eta_{t_0, s_1})^{\alpha_1}| \cdots |(\Delta \eta_{t_0, s_i})^{\alpha_i}|$$

which completes the proof. \square

The order of truncation error is estimated in the next lemma.

Lemma 8.5 *Let $\gamma > 0$ and let $\varepsilon = (\varepsilon_1, \dots, \varepsilon_m)^\top \in [0, 1]^m$ with $\|\varepsilon\|_1 < \gamma$. Then,*

$$\|x(t) - \Phi_\gamma(x_0, t_0, \Delta_{t_0, t})\|_2 \leq C_\gamma^\varepsilon \cdot (\Delta_{t_0, t})^{(\gamma+1-\|\varepsilon\|_1)}$$

for all $t_0, t \in [0, T]$ with $0 \leq t - t_0 \leq 1$.

Proof For $l = \lceil \gamma \rceil \in \mathbb{N}_0$, consider the integral equation expansion (8.4)

$$\begin{aligned} x(t) &= x_0 + \sum_{i=1}^l \int_{t_0}^t \cdots \int_{t_0}^{s_{i-1}} (L^i \chi)(x_0, \eta_{s_1}, \dots, \eta_{s_i}) ds_i \dots ds_1 \\ &\quad + \int_{t_0}^t \cdots \int_{t_0}^{s_l} (L^{l+1} \chi)(x(s_{l+1}), \eta_{s_1}, \dots, \eta_{s_{l+1}}) ds_{l+1} \dots ds_1. \end{aligned}$$

Using Lemma 8.3 to obtain the Taylor expansion of the integrands $(L^i \chi)(x_0, \eta_{s_1}, \dots, \eta_{s_i})$ in the last $m \cdot i$ -variables of order

$$k_i = \left\lceil \frac{\gamma + 1 - i}{\vartheta} \right\rceil - 1, \quad i = 1, \dots, l.$$

Thus

$$x(t) = x_0 + \sum_{i=1}^l \sum_{\substack{|\alpha| \leq k_i \\ \alpha \in \mathbb{N}_0^{m \times i}}} F_\alpha(x_0, \eta_{t_0}^{\times i}) \cdot I_{\alpha, t_0, t} + E_1 + E_2 \quad (8.5)$$

with

$$\begin{aligned} E_1 &= \int_{t_0}^t \dots \int_{t_0}^{s_l} (L^{l+1} \chi)(x(s_{l+1}), \eta_{s_1}, \dots, \eta_{s_{l+1}}) \, ds_{l+1} \dots ds_1, \\ E_2 &= \sum_{i=1}^l \sum_{\substack{|\alpha|=k_i+1 \\ \alpha \in \mathbb{N}_0^{m \times i}}} \int_{t_0}^t \dots \int_{t_0}^{s_{i-1}} (\Delta \eta_{t_0, s_1})^{\alpha_1} \dots (\Delta \eta_{t_0, s_i})^{\alpha_i} \cdot (k_i + 1) \\ &\quad \cdot \left(\int_0^1 F_\alpha(x_0, \eta_{t_0} + \lambda \Delta \eta_{t_0, s_1}, \dots, \eta_{t_0} + \lambda \Delta \eta_{t_0, s_i}) \cdot (1 - \lambda)^{k_i} \, d\lambda \right) ds_i \dots ds_1. \end{aligned}$$

Note that

$$\mathcal{M}_m^\gamma = \{\alpha \in \mathcal{M}_m : \iota(\alpha) + \|\alpha^\top \Theta\|_1 < \gamma + 1\} \subset \{\alpha \in \mathcal{M}_m : |\alpha| \leq k_{\iota(\alpha)}\}$$

for $|\alpha|\vartheta \leq \|\alpha^\top \Theta\|_1$. Hence, Eq.(8.5) can be written in the form

$$x(t) = x_0 + \Phi_\gamma(x_0, t_0, \Delta_{t_0, t}) + E_1 + E_2 + E_3,$$

with

$$\begin{aligned} \Phi_\gamma(x_0, t_0, \Delta_{t_0, t}) &= \sum_{i=1}^l \sum_{\substack{\alpha \in \mathcal{M}_m^\gamma \\ \iota(\alpha)=i}} F_\alpha(x_0, \eta_{t_0}^{\times i}) \cdot I_{\alpha, t_0, t} \\ E_3 &= \sum_{i=1}^l \sum_{\substack{|\alpha| \leq k_i \\ \alpha \in \mathbb{N}_0^{m \times i} \setminus \mathcal{M}_m^\gamma}} F_\alpha(x_0, \eta_{t_0}^{\times i}) \cdot I_{\alpha, t_0, t}. \end{aligned}$$

It then remains to estimate E_1 , E_2 and E_3 . In fact for every $\beta \in (0, 1]^m$

$$\begin{aligned}\|E_1\|_2 &\leq \left(\sup_{\substack{s_1, \dots, s_{l+1} \\ \in [0, T]}} \| (L^{l+1} \chi) (x(s_{l+1}), \eta_{s_1}, \dots, \eta_{s_{l+1}}) \|_2 \right) \frac{(\Delta_{t_0, t})^{(l+1)}}{(l+1)!}, \\ \|E_2\|_2 &\leq \sum_{i=1}^l \sum_{\substack{|\alpha|=k_i+1 \\ \alpha \in \mathbb{N}_0^{m \times i}}} \left(\frac{(\|\eta\|_\beta)^{|\alpha|}}{i!} (\Delta_{t_0, t})^{(i+\|\alpha^\top \beta\|_1)} r_\alpha \right) \\ \|E_3\|_2 &\leq \sum_{i=1}^l \sum_{\substack{|\alpha| \leq k_i \\ \alpha \in \mathbb{N}_0^{m \times i} \setminus \mathcal{M}_m^\gamma}} \left(\frac{(\|\eta\|_\beta)^{|\alpha|}}{i!} (\Delta_{t_0, t})^{(i+\|\alpha^\top \beta\|_1)} r_\alpha \right),\end{aligned}$$

where r_α is the random variable

$$r_\alpha = \sup_{\substack{\|w_1\|_2, \dots, \|w_i\|_2 \leq \|\beta\|_\infty \\ \|y\|_2 \leq \|x\|_\infty}} \|F_\alpha(y, w_1, \dots, w_i)\|_2.$$

Therefore for every $\beta \in (0, 1]^m$

$$\|x(t) - \Phi_\gamma(x_0, t_0, \Delta_{t_0, t})\|_2 \leq \sum_{\substack{i=1, \dots, l+1 \\ |\alpha| \leq \lceil \frac{\gamma}{\vartheta} \rceil \\ \alpha \in \mathbb{N}_0^{m \times i} \setminus \mathcal{M}_m^\gamma}} \left(\frac{(\|\eta\|_\beta)^{|\alpha|}}{i!} (\Delta_{t_0, t})^{(i+\|\alpha^\top \beta\|_1)} r_\alpha \right),$$

which implies that

$$\|x(t) - \Phi_\gamma(x_0, t_0, \Delta_{t_0, t})\|_2 \leq \sum_{\substack{\alpha \in \mathcal{M}_m \setminus \mathcal{M}_m^\gamma \\ \iota(\alpha) \leq \lceil \frac{\gamma}{\vartheta} \rceil + 1 \\ |\alpha| \leq \lceil \frac{\gamma}{\vartheta} \rceil}} \left(\frac{(\|\eta\|_\beta)^{|\alpha|}}{\iota(\alpha)!} (\Delta_{t_0, t})^{(\iota(\alpha)+\|\alpha^\top \beta\|_1)} r_\alpha \right).$$

Note that if $\beta = \Theta - \varepsilon / \lceil \frac{\gamma}{\vartheta} \rceil$ then

$$\begin{aligned}\iota(\alpha) + \|\alpha^\top \beta\|_1 &= \iota(\alpha) + \|\alpha^\top \Theta\|_1 - \left\| \alpha^\top \frac{\varepsilon}{\lceil \frac{\gamma}{\vartheta} \rceil} \right\|_1 \geq \gamma + 1 - \left\| \alpha^\top \frac{\varepsilon}{\lceil \frac{\gamma}{\vartheta} \rceil} \right\|_1 \\ &\geq \gamma + 1 - |\alpha| \frac{\|\varepsilon\|_1}{\lceil \frac{\gamma}{\vartheta} \rceil} \geq \gamma + 1 - \|\varepsilon\|_1\end{aligned}$$

for $i = 1, \dots, [\gamma] + 1$ and $\alpha \in \mathbb{N}_0^{m \times i} \setminus \mathcal{M}_m^\gamma$ with $|\alpha| \leq \lceil \frac{\gamma}{\vartheta} \rceil$. Therefore

$$\|x(t) - \Phi_\gamma(x_0, t_0, \Delta_{t_0, t})\|_2 \leq \sum_{\substack{\alpha \in \mathcal{M}_m \setminus \mathcal{M}_m^\gamma \\ \iota(\alpha) \leq [\gamma] + 1 \\ |\alpha| \leq \lceil \frac{\gamma}{\vartheta} \rceil}} \left(\frac{(\|\eta\|_\beta)^{|\alpha|}}{\iota(\alpha)!} r_\alpha \right) (\Delta_{t_0, t})^{(\gamma+1-\|\varepsilon\|_1)}$$

with $\beta = \Theta - \varepsilon / \lceil \frac{\gamma}{\vartheta} \rceil$.

This completes the proof of Lemma 8.5 and hence of Theorem 8.1. \square

8.7 Endnotes

Jentzen and Kloeden [87] (see also [76]) used Taylor expansions of the vector field of the RODE to obtain an implicit integral equation expansions for the solutions into which lower order approximations are inserted to provide an explicit approximation of the solution that were called RODE-Taylor expansion. This was illustrated in Chap. 1.

Here the simpler and more direct approach of Jentzen and Kloeden [78] is used to express such RODE-Taylor expansions. It is motivated by the Wagner-Platen multi-index notation that was used in [91] to formulate succinctly stochastic Taylor expansions and Taylor schemes for SODEs, see Chap. 6. The main difference is that matrix valued multi-indices are required to handle the possibly different Hölder exponents of the different components of the driving noise processes. As in the SODE case the coefficient functions are obtained by iterated application of differential operators to the vector field of the RODE, but instead of iterated integrals of the components of the noise processes single integrals of suitable powers of increments of them are used.

Part III

Numerical Schemes for Random Ordinary Differential Equations

Chapter 9

Numerical Methods for Ordinary and Stochastic Differential Equations

Taylor expansions are a very basic tool in numerical analysis. They allow one to derive one-step numerical schemes for ordinary differential equations (ODEs) of arbitrarily high order. In practice such Taylor schemes are rarely implemented, but are used instead as a theoretical comparison for determining the convergence orders of other schemes that have been derived by more heuristic methods. In view of the less robust nature of the Itô stochastic integral, stochastic Taylor expansions and the corresponding stochastic Taylor schemes are the essential starting point for the derivation of consistent higher order numerical schemes for stochastic differential equations (SODEs), see e.g. Kloeden and Platen [91] and Milstein [105].

Random ordinary differential equations (RODEs) are nonautonomous ODEs for each sample path, to which deterministic calculus can be applied pathwise. Typically, since the driving stochastic process in the RODE has at most Hölder continuous sample paths, the solution sample paths are continuously differentiable, but the sample paths of the derivative are at most Hölder continuous in time. Thus the resulting vector field after insertion of the driving stochastic process is at most Hölder continuous in time, no matter how smooth the vector field is in its original variables. The solutions of RODEs do not have sufficient smoothness to have Taylor expansions in the usual sense. Although classical numerical schemes for ODEs can be used pathwise for RODEs, they rarely attain their traditional order. Alternative kinds of Taylor expansions for RODEs that were introduced in Part II will be used in the following chapters to derive higher order numerical schemes for RODEs.

9.1 One-Step Numerical Schemes for ODEs

Consider an initial value problem (IVP)

$$\frac{dx}{dt} = f(t, x), \quad x(t_0) = x_0$$

with solution $x(t) = x(t, t_0, x_0)$ on an interval $[t_0, T]$ and consider a partition

$$t_0 < t_1 < \dots < t_n < t_{n-1} < \dots < t_N = T$$

with step size $h_n = t_{n+1} - t_n \equiv h > 0$.

The simplest numerical scheme for an approximation $x_n \approx x(t_n, t_0, x_0)$ is the Euler scheme

$$x_{n+1} = x_n + h_n f(t_n, x_n).$$

It is an example of a one-step numerical scheme, which has the general form

$$x_{n+1} = x_n + h_n \phi(h_n, t_n, x_n, x_{n+1}),$$

with an increment function ϕ .

Another example is the p -Taylor scheme based on the p -Taylor approximation

$$x_{n+1} = x_n + \sum_{j=1}^p \frac{1}{j!} h_n^j D^{j-1} f(t_n, x_n)$$

with the increment function

$$\phi(h, t, x) = \sum_{j=1}^p \frac{1}{j!} h^{j-1} D^{j-1} f(t, x),$$

where the total differential operator (scalar version) is defined as

$$D := \frac{\partial}{\partial t} + f \frac{\partial}{\partial x} \quad \text{with} \quad D^0 = Id.$$

The *Heun scheme*

$$x_{n+1} = x_n + \frac{1}{2} h_n \left[f(t_n, x_n) + f(t_{n+1}, x_n + hf(t_n, x_n)) \right]$$

with the increment function

$$\phi(h, t, x) = \frac{1}{2} [f(t, x) + f(t + h, x + hf(t, x))]$$

is a representative of the Runge–Kutta class of derivative-free schemes.

Numerical schemes for ODEs are often derived by heuristic arguments. They are called *consistent* if

$$\phi(0, t, x) \equiv f(t, x),$$

which essentially says that the line segment joining the numerical iterate to the starting point approximates the tangent line to the ODE solution through that point. It provides a quick test for the convergence of a numerical scheme, assuming that f and ϕ are continuous in all variables, although it says nothing about the rate of convergence.

Discretisation Error

For notational convenience a constant time step $h_n \equiv h > 0$ will now be used. The numerical solution at t_n , which obviously depends on h , will be denoted by $x_n^{(h)}$. The notation $|\cdot|$ represents the absolute value or the Euclidean norm.

The *global discretisation error* (GDE) of a numerical scheme is defined as

$$\mathcal{E}_n^G(h) := |x(t_n, t_0, x_0) - x_n^{(h)}|, \quad n = 0, 1, \dots, N_h := \frac{T - t_0}{h}.$$

A numerical scheme is said to *converge* if

$$\lim_{h \rightarrow 0} \max_{0 \leq n \leq N_h} \mathcal{E}_n^G(h) = 0$$

and, in particular, is said to have *global order p* if there exists $C_T \geq 0$ such that

$$\max_{0 \leq n \leq N_h} \mathcal{E}_n^G(h) \leq C_T h^p$$

for sufficiently small step size h .

On the other hand, the *local discretisation error* (LDE) is defined as

$$\mathcal{E}_{n+1}^L(h) := |x(t_{n+1}, t_n, x_n^{(h)}) - x_{n+1}^{(h)}|, \quad n = 0, 1, \dots, N_h - 1.$$

The LDE compares one step of the numerical scheme with the solution of the ODE starting at the same point. Obviously $\mathcal{E}_1^L = \mathcal{E}_1^G$, but in general $\mathcal{E}_{n+1}^L(h) \neq \mathcal{E}_{n+1}^G(h)$, because the solution $x(t, t_n, x_n^{(h)})$ with the initial value $x(t_n) = x_n^{(h)}$ is not the same as the sought solution $x(t, t_0, x_0)$. Nevertheless, the local discretisation error is convenient, because it can be easily estimated through a Taylor expansion and this can then be used to estimate the global discretisation error. In fact, for a p -Taylor scheme, the local discretisation error is just the truncation error of the corresponding p -Taylor approximation.

Theorem 9.1 Consider an explicit one-step numerical scheme

$$x_{n+1} = x_n + h_n \phi(h_n, t_n, x_n)$$

with local discretisation error of order $(p + 1)$, where $\phi(h, t, x)$ is continuous in h , t , x and uniformly Lipschitz in x (with Lipschitz constant κ). The scheme has global discretisation error of order p .

Proof For simplicity write h instead of h_n , and write $x(t)$ for the solution $x(t, t_n, x_n^{(h)})$. Then the GDE satisfies

$$\begin{aligned}
\mathcal{E}_{n+1}^G(h) &= |x(t_{n+1}) - x_{n+1}| \\
&\leq |x(t_{n+1}) - x(t_n) - h\phi(h, t_n, x(t_n))| \quad \text{local discretisation error} \\
&\quad + \left| x(t_n) + h\phi(h, t_n, x(t_n)) - \underbrace{x_n - h\phi(h, t_n, x_n)}_{x_{n+1}} \right| \\
&\leq \mathcal{E}_{n+1}^L(h) + |x(t_n) - x_n| + \underbrace{h |\phi(h, t_n, x(t_n)) - \phi(h, t_n, x_n)|}_{\text{Lipschitz: } \leq h\kappa |x(t_n) - x_n|} \\
&\leq \mathcal{E}_{n+1}^L(h) + (1 + \kappa h) \underbrace{|x(t_n) - x_n|}_{\mathcal{E}_n^G(h)} \\
&\leq C_T h^{p+1} + (1 + \kappa h) \mathcal{E}_n^G(h),
\end{aligned}$$

which gives a difference inequality with $\mathcal{E}_0^G(h) = 0$. Hence, by induction,

$$\mathcal{E}_n^G(h) \leq \frac{(1 + \kappa h)^n - 1}{(1 + \kappa h) - 1} C_T h^{p+1} \leq e^{\kappa(T-t_0)} \frac{C_T}{\kappa} h^p$$

i.e., the global error $\mathcal{E}_n^G(h) \sim \mathcal{O}(h^p)$. □

9.2 One-Step Numerical Schemes for Itô SODEs

Consider a d -dimensional Itô SODE

$$dX_t = f(X_t) dt + \sum_{j=1}^m \sigma_j(X_t) dW_t^j \tag{9.1}$$

with $f, \sigma_1, \dots, \sigma_m : \mathbb{R}^d \rightarrow \mathbb{R}^d$. Here $X_t = (X_t^1, \dots, X_t^d) \in \mathbb{R}^d$ and $W_t = (W_t^1, \dots, W_t^m)$ is an m -dimensional Wiener process, i.e., its components W_t^1, \dots, W_t^m are pairwise independent scalar Wiener processes. (Note that superscripts label components of vectors here).

Consider a partition $t_0 < t_1 < \dots < t_{N_T} = T$ of the interval $[t_0, T]$ with step sizes $h_n := t_{n+1} - t_n > 0$ and maximum step size $h := \max_n h_n$. Let $x_n^{(h)}$ be an approximation generated by some numerical schemes of X_{t_n} for a solution X_t of an SODE (9.1).¹ Such an approximation is said to have *strong approximation of order*

¹When the context is clear the superscript (h) can be dropped.

γ if there exists $K_{p,T} > 0$ such that

$$\left(\mathbb{E} \left[\sup_{n=0,\dots,N_T} \|x_n^{(h)} - X_{t_n}\|^p \right] \right)^{1/p} \leq K_{p,T} h^\gamma.$$

Usually $p = 1$ or 2 are considered.

9.3 Strong Taylor Schemes for Itô SODEs

Strong Taylor schemes for the SODE (9.1) are based on strong Taylor approximations (6.11) applied on successive discretisation subintervals $[t_n, t_{n+1}]$. They are based on the hierarchical set of multi-indices (see Chap. 6)

$$\Lambda_\gamma = \left\{ \alpha \in \mathcal{J}_m : l(\alpha) + n(\alpha) \leq 2\gamma \text{ or } l(\alpha) = n(\alpha) = \gamma + \frac{1}{2} \right\}$$

for $\gamma = \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$, where $n(\alpha)$ is the number of components of α that are equal to 0 and

$$\mathcal{J}_m = \{ \alpha = (j_1, \dots, j_l), j_1, \dots, j_l \in \{0, 1, 2, \dots, m\} : l \in \mathbb{N} \} \bigcup \{\emptyset\}$$

with \emptyset being the empty index of length $l(\emptyset) = 0$. In addition, they involve multiple stochastic integrals

$$I_{\alpha, t_n, t_{n+1}} = \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_l} \dots \int_{t_n}^{s_2} dW_{s_1}^{j_1} \dots dW_{s_{l-1}}^{j_{l-1}} dW_{s_l}^{j_l},$$

with $I_{\emptyset, n} = 1$, and iterated operators

$$L^\alpha := L^{j_1} \dots L^{j_l} \text{ for } \alpha = (j_1, \dots, j_l)$$

with L^\emptyset being the identity operator, i.e., $L^\emptyset f = f$ for all f , and

$$L^0 = \frac{\partial}{\partial t} + \sum_{k=1}^d f^k \frac{\partial}{\partial x^k} + \frac{1}{2} \sum_{k,l=1}^d \sum_{j=1}^m \sigma_j^k \sigma_j^l \frac{\partial^2}{\partial x^k \partial x^l},$$

$$L^j = \sum_{k=1}^d \sigma_j^k \frac{\partial}{\partial x^k}, \quad j = 1, \dots, m$$

for a smooth enough scalar-valued function of t and x .

Definition 9.1 The order γ strong Taylor scheme for the SODE (9.1) has the compact componentwise form

$$x_{n+1}^k = \sum_{\alpha \in \Lambda_\gamma} L^\alpha \chi^k(t_n, x_n) I_{\alpha,n}, \quad k = 1, \dots, d, \quad (9.2)$$

for $\gamma = \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$, where χ^k is the k th component of the identity mapping χ on \mathbb{R}^d .

By Theorem 10.6.4 in [91] the order γ strong Taylor scheme (9.2), which is based on the order γ strong Taylor approximation (6.11), converges strongly with order γ . Proofs in the literature of the above convergence orders, e.g., in the monographs Kloeden and Platen [91] and Milstein [105], assume that the coefficient functions F_α in the Itô-Taylor schemes are uniformly bounded on \mathbb{R}^d , i.e., the partial derivatives of appropriately high order of the SODE coefficient functions $f, \sigma_1, \dots, \sigma_m$ are uniformly bounded on \mathbb{R}^d .

These *standard assumptions* are not satisfied by many SODEs in important applications (see Hutzenthaler and Jentzen [67]). One way to overcome this problem is to use stopping time to restrict coefficient functions to a bounded set, another is to focus attention on SODEs with special dynamical properties (see, e.g., [64, 102, 107]). This yields the appropriate order estimates without requiring bounded derivatives of coefficients.

Remark 9.1 The situation for SODEs differs from that for ODEs in that there is not so much accumulation of errors in going from the local discretisation error to the global discretisation error (see Theorem 1.1 in Milstein [105] and also Jentzen and Kloeden [75]).

Example 9.1 For the scalar Itô SODE

$$dX_t = f(X_t) dt + \sigma(X_t) dW_t,$$

the Itô-Taylor scheme of strong order $\gamma = \frac{1}{2}$ with the hierarchical set $\Lambda_{\frac{1}{2}} = \{(0), (1)\}$ is the Euler–Maruyama scheme

$$x_{n+1} = x_n + f(x_n) h_n + \sigma(x_n) \Delta W_n,$$

where $\Delta W_n = I_{(1), t_n, t_{n+1}} = W_{t_{n+1}} - W_{t_n}$.

The Itô-Taylor scheme of strong order $\gamma = 1$ with the hierarchical set $\Lambda_1 = \{(0), (1), (1, 1)\}$ is the Milstein scheme

$$x_{n+1} = x_n + f(x_n) h_n + \sigma(x_n) \Delta W_n + \frac{1}{2} \sigma(x_n) \sigma'(x_n) [(\Delta W_n)^2 - h_n],$$

where the coefficient functions (recall that here $F_\alpha = L^\alpha \chi$) are

$$F_{(0)} = f, \quad F_{(1)} = \sigma, \quad F_{(1,1)} = \sigma\sigma',$$

and the iterated integrals are

$$\begin{aligned} I_{(0),t_n,t_{n+1}} &= \int_{t_n}^{t_{n+1}} dW_s = h_n, \\ I_{(1),t_n,t_{n+1}} &= \int_{t_n}^{t_{n+1}} dW_s = \Delta W_n, \\ I_{(1,1),t_n,t_{n+1}} &= \int_{t_n}^{t_{n+1}} \int_{t_n}^s dW_\tau dW_s = \frac{1}{2} [(\Delta W_n)^2 - h_n]. \end{aligned}$$

Pathwise Convergence

A numerical scheme is said to converge *pathwise* if

$$\sup_{n=0,\dots,N_T} \|X_{t_n}(\omega) - x_n^{(h)}(\omega)\| \rightarrow 0 \quad \text{as } h \rightarrow 0$$

for (at least) almost all $\omega \in \Omega$, where Ω is the sample space of the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Pathwise convergence of numerical schemes is interesting because numerical calculations of the approximating random variable $x_n^{(h)}$ are carried out path by path. In addition, the solutions of some SODEs are non-integrable, i.e., $\mathbb{E}[|X_t|] = \infty$ for some $t \geq 0$, so strong convergent approximation is not possible. It was seen in Theorem 6.2 that a strong Itô-Taylor approximation of order γ converges pathwise with order $\gamma - \varepsilon$ for all sufficiently small $\varepsilon > 0$. A similar result holds for strong RODE-Taylor schemes.

Theorem 9.2 *Under the standard assumptions a RODE-Taylor scheme based on an Itô-Taylor scheme of strong order $\gamma > 0$ converges pathwise with order $\gamma - \varepsilon$ for all sufficiently small $\varepsilon > 0$, i.e.,*

$$\sup_{i=0,\dots,N_T} \|X_{t_i}(\omega) - X_n^{(h)}(\omega)\| \leq C_{\varepsilon,T}^\gamma(\omega) \cdot h^{\gamma-\varepsilon}$$

for almost all $\omega \in \Omega$.

Pathwise convergence also holds when the vector field of the RODE $g \in \mathcal{C}^{2\gamma+1}$, i.e., the partial derivatives of g in the x -variables need not be uniformly bounded, although those in the y -variables are. This is stated as Theorem 10.1 in the next chapter. The proof there is based on that of Theorem 9.2 and a localisation argument. It does not depend on the specific structure of the strong Taylor schemes, just the fact that they converge pathwise under the standard assumptions.

9.4 Endnotes

See Kloeden and Platen [91] and Milstein [105] for detailed discussions on numerical schemes for Itô SODEs based on stochastic Taylor expansions. See also Milstein and Tretyakov [106] and the review Jentzen and Kloeden [77]. Hutzenthaler and Jentzen [67] discuss numerical schemes for SODEs which do not satisfy the standard assumptions.

Jentzen, Kloeden and Neuenkirch [74], Kloeden and Neuenkirch [90] and Talay [132] consider the pathwise convergence of strong Itô–Taylor schemes for SODEs.

Chapter 10

Itô–Taylor Schemes for RODEs with Itô Noise

RODEs driven by an Itô diffusion process, i.e., the solution of an Itô SODE, are the focus of this chapter. In particular, consider a RODE on \mathbb{R}^{d_1}

$$\frac{dx}{dt} = g(x, Y_t), \quad (10.1)$$

where Y_t is the solution of an Itô SODE in \mathbb{R}^{d_2}

$$dY_t = f(Y_t) dt + \sum_{j=1}^m \sigma_j(Y_t) dW_t^j \quad (10.2)$$

with m independent scalar Wiener processes W_t^1, \dots, W_t^m . These form an SODE in $\mathbb{R}^{d_1+d_2}$

$$d \begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} g(X_t, Y_t) \\ f(Y_t) \end{pmatrix} dt + \sum_{j=1}^m \begin{pmatrix} 0 \\ \sigma_j(Y_t) \end{pmatrix} dW_t^j. \quad (10.3)$$

The idea is to use numerical schemes for the SODE system (10.3) to obtain schemes for the RODE (10.1) that are pathwise convergent. By Theorem 9.2 an order γ strong Taylor scheme which has order γ strong convergence also has order $\gamma - \varepsilon$ pathwise convergence for arbitrarily small $\varepsilon > 0$. Hence it has order $\gamma - \varepsilon$ pathwise convergence when restricted to the RODE component (10.1).

The schemes to be introduced in this section involve the stochastic integrals

$$I_{\alpha,n}[u(\cdot)] := I_{\alpha}[u(\cdot)]_{t_n, t_{n+1}} = \int_{t_n}^{t_{n+1}} \dots \int_{t_n}^{s_2} u(s_1) dW_{s_1}^{j_1} \dots dW_{s_l}^{j_l},$$

for $\alpha = (j_1, \dots, j_l) \in \mathcal{J}_m$. In particular when $u \equiv 1$, write

$$I_{\alpha,n} := I_{\alpha}[1]_{t_n, t_{n+1}} = \int_{t_n}^{t_{n+1}} \cdots \int_{t_n}^{s_2} dW_{s_1}^{j_1} \cdots dW_{s_l}^{j_l}.$$

10.1 One-Step Schemes

10.1.1 Scalar Case

The scalar RODE driven by a scalar Itô diffusion Y_t ,

$$\frac{dx}{dt} = g(x, Y_t), \quad dY_t = f(Y_t)dt + \sigma(Y_t)dW_t, \quad (10.4)$$

where W_t is scalar Wiener process, can be written as a system of Itô SODEs

$$dX_t = g(X_t, Y_t)dt + 0dW_t, \quad (10.5)$$

$$dY_t = f(Y_t)dt + \sigma(Y_t)dW_t. \quad (10.6)$$

The system (10.5)–(10.6) can also be written in vector notation, as the Itô SODE (10.3) in \mathbb{R}^2

$$dU_t = \vec{f}(U_t)dt + \vec{\sigma}(U_t)dW_t, \quad (10.7)$$

where

$$U_t = \begin{pmatrix} X_t \\ Y_t \end{pmatrix}, \quad \vec{f}(U_t) = \begin{pmatrix} g(X_t, Y_t) \\ f(Y_t) \end{pmatrix}, \quad \vec{\sigma}(U_t) = \begin{pmatrix} 0 \\ \sigma(Y_t) \end{pmatrix}.$$

Note that here \vec{f} and $\vec{\sigma}$ are vector-valued functions different from f and σ .

In this case the differential operators L^0 and L^1 introduced in Chap. 6 reduce to

$$\begin{aligned} L^0 &= \frac{\partial}{\partial t} + g(x, y) \frac{\partial}{\partial x} + f(y) \frac{\partial}{\partial y} + \frac{1}{2} \sigma^2(y) \frac{\partial^2}{\partial y^2}, \\ L^1 &= \sigma(y) \frac{\partial}{\partial y}. \end{aligned}$$

Here $U^1 = X_t$ and $U^2 = Y_t$, so

$$\begin{aligned} L^0 \chi^1(U) &= g(X_t, Y_t), & L^0 \chi^2(U) &= f(Y_t), \\ L^1 \chi^1(U) &= 0, & L^1 \chi^2(U) &= \sigma(Y_t), \end{aligned}$$

where $\chi = (\chi^1, \chi^2) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is the identity function satisfying $\chi^i(U) = U^i$, for $i = 1, 2$. (Recall that vector components are denoted by superscripts).

Since the Y_t equation is an SODE in its own right, its order γ strong Taylor scheme is just a scalar version of the usual one, i.e.,

$$y_{n+1} = \sum_{\alpha \in \Lambda_\gamma} L^\alpha \chi^2(U)(t_n, y_n) I_{\alpha,n}. \quad (10.8)$$

On the other hand, the strong Taylor scheme for the x -component simplifies to

$$x_{n+1} = \sum_{\alpha \in \Lambda_\gamma^0} L^\alpha \chi^1(U)(t_n, x_n, y_n) I_{\alpha,n} \quad (10.9)$$

with the hierarchical set

$$\Lambda_\gamma^0 = \{\alpha \in \Lambda_\gamma \subset \mathcal{J}_1 : \alpha = \emptyset \text{ or } l(\alpha) \geq 1 \text{ with last component } j_l = 0\}. \quad (10.10)$$

Recall from Chap. 6 that the zero diffusion coefficient in the first component of the coupled system of SODEs (10.5)–(10.6) means that all terms corresponding to a multi-index α with final component $j_l = 1$ vanish in the RODE part of the strong Itô–Taylor expansion. Specifically, for any index α with $j_l = 1$,

$$L^\alpha \chi^1(U) = L^{\alpha-} L^{j_l} \chi^1(U) = L^{\alpha-} L^1 \chi^1(U) = L^{\alpha-} \emptyset \equiv 0,$$

where $\alpha-$ denotes the multi-index obtained by deleting the last component of α . On the other hand, terms such as

$$L^1 L^0 \chi^1(U) = L^1 g(X_t, Y_t) = \sigma(Y_t) \frac{\partial g}{\partial y}(X_t, Y_t)$$

do not vanish automatically.

10.1.2 Vector Case

Consider the general vector case described in the SODE (10.3) with the notation

$$V_t = (X_t^1, \dots, X_t^{d_1}, Y_t^1, \dots, Y_t^{d_2}) \in \mathbb{R}^d,$$

where $d = d_1 + d_2$. Then (10.3) can be written as the multi-dimensional Itô SODE

$$dV_t = \vec{f}(V_t) dt + \sum_{j=1}^m \vec{\sigma}_j(V_t) dW_t^j, \quad (10.11)$$

with

$$\vec{f}(V_t) = \begin{pmatrix} g(X_t, Y_t) \\ f(Y_t) \end{pmatrix}, \quad \vec{\sigma}_j(V_t) = \begin{pmatrix} 0 \\ \sigma_j(Y_t) \end{pmatrix}.$$

The order γ strong Taylor scheme for (10.11) reads

$$v_{n+1} = \sum_{\alpha \in \Lambda_\gamma} L^\alpha \chi(V)(t_n, v_n) I_{\alpha, n},$$

where the hierarchical set $\Lambda_\gamma \subset \mathcal{J}_m$ and the differential operators are given by

$$\begin{aligned} L^0 &= \frac{\partial}{\partial t} + \sum_{k=1}^{d_1} g^k \frac{\partial}{\partial X^k} + \sum_{k=1}^{d_2} f^k \frac{\partial}{\partial Y^k} + \sum_{k,l=1}^{d_2} \sum_{j=1}^m \frac{1}{2} \sigma_j^k \sigma_j^l \frac{\partial^2}{\partial Y^k \partial Y^l}, \\ L^j &= \sum_{k=1}^{d_2} \sigma_j^k \frac{\partial}{\partial Y^k} \quad j = 1, \dots, m. \end{aligned}$$

for smooth enough functions from $[0, T] \times \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ to \mathbb{R} . (Recall again that the components of vector-valued expressions are denoted by superscripts).

The x -component of the order γ strong Taylor scheme gives the order γ strong RODE-Taylor scheme in a componentwise form

$$x_{n+1}^k = \sum_{\alpha \in \Lambda_\gamma^0} L^\alpha \chi^k(V)(t_n, x_n, y_n) I_{\alpha, n}, \quad k = 1, 2, \dots, d_1,$$

where $\chi^k(V) = V^k$, and $\Lambda_\gamma^0 \subset \Lambda_\gamma$ is defined as in (10.10).

Pathwise convergence

Theorem 9.2 gives the order $\gamma - \varepsilon$ pathwise convergence for an order γ RODE-Taylor scheme under the standard assumptions, i.e., that all the derivatives appearing in the scheme are uniformly bounded. The standard assumptions there are, in fact, too strong just for the pathwise convergence of the Taylor scheme (10.9) and for many applications. They ensure the strong convergence of the full Taylor scheme (10.8)–(10.9), but the x -component scheme (10.9) can still converge pathwise even when the full scheme does not converge in the strong sense, provided the noise is approximated to the required order. This is possible directly or through the y -component scheme (10.8) when the noise is a simple process such as a Wiener process or an Ornstein–Uhlenbeck process.

In fact, pathwise convergence also holds when the vector field of the RODE $g \in \mathcal{C}^{2\gamma+1}$, i.e., the partial derivatives of g in the x -variables need not be uniformly bounded, although those in the y -variables are. The proof, which is given at the end of the chapter, is based on that of Theorem 9.2 and a localisation argument. It does not depend on the specific structure of the strong Taylor schemes, just the fact that they converge pathwise under the standard assumptions, which follows by the

Burkholder–Davis–Gundy inequality and a Borel–Cantelli argument when all of the error moments converge with the same order γ .

Theorem 10.1 Suppose that $g \in \mathcal{C}^{2\gamma+1}$ with the partial derivatives of g in the y -variables uniformly bounded, but not necessarily those in the x -variables, and that the coefficients of the Itô SODE for the driving process satisfy standard assumptions. Then the order γ RODE-Taylor scheme converges pathwise with order $\gamma - \varepsilon$ for all $\varepsilon > 0$, i.e.,

$$\sup_{n=0, \dots, N_T} \|X_{t_n}(\omega) - x_n^{(h)}(\omega)\| \leq C_{\varepsilon, T}^{(\gamma)}(\omega) \cdot h^{\gamma-\varepsilon}$$

for almost all $\omega \in \Omega$.

For example the Euler–Maruyama scheme has pathwise order $\frac{1}{2} - \varepsilon$ and the Milstein scheme has pathwise order $1 - \varepsilon$ when the vector field of the RODE is just appropriately often continuously differentiable (without the partial derivatives necessarily being uniformly bounded).

10.1.3 Examples

The above schemes will be illustrated for the scalar system (10.4). For notational compactness partial derivatives will be denoted by subscripts. Function g and its partial derivatives are evaluated at (x_n, y_n) , and functions f and σ are evaluated at y_n .

Case I: $\gamma = \frac{1}{2}$

The hierarchical sets are $\Lambda_{\frac{1}{2}} = \{\emptyset, (0), (1)\}$, $\Lambda_{\frac{1}{2}}^0 = \{\emptyset, (0)\}$ and the corresponding RODE-Taylor scheme is the Euler–Maruyama scheme which reduces to

$$x_{n+1} = x_n + gh_n.$$

This is also the RODE-Taylor scheme obtainable from the Milstein scheme with $\gamma = 1$ since $\Lambda_1 = \{\emptyset, (0), (1), (1, 1)\}$ and $\Lambda_1^0 = \{\emptyset, (0)\}$.

Case II: $\gamma = \frac{3}{2}$

The hierarchical sets are

$$\begin{aligned}\Lambda_{\frac{3}{2}} &= \{\emptyset, (0), (1), (1, 1), (0, 1), (1, 0), (0, 0), (1, 1, 1)\}, \\ \Lambda_{\frac{3}{2}}^0 &= \{\emptyset, (0), (1, 0), (0, 0)\},\end{aligned}$$

and the corresponding order $\frac{3}{2}$ RODE-Taylor scheme is given by

$$x_{n+1} = x_n + gh_n + \sigma g_y I_{(1,0),n} + \frac{1}{2} \left(gg_x + fg_y + \frac{1}{2} (\sigma)^2 g_{yy} \right) h_n^2. \quad (10.12)$$

This scheme includes the multiple stochastic integral

$$I_{(1,0),n} = \int_{t_n}^{t_{n+1}} \int_{t_n}^t dW_s dt,$$

which is correlated to the simple integral

$$\Delta W_n = I_{(1),n} = \int_{t_n}^{t_{n+1}} dW_s.$$

They can be generated by using two independent $\mathcal{N}(0, 1)$ -distributed random variables \mathcal{N}_1 and \mathcal{N}_2 (see Chap. 14) via

$$I_{(1),n} = \sqrt{h_n} \mathcal{N}_1, \quad I_{(1,0),n} = \frac{1}{2} h_n^{3/2} \left(\mathcal{N}_1 + \frac{1}{\sqrt{3}} \mathcal{N}_2 \right).$$

Case III: $\gamma = 2$

The hierarchical sets are

$$\begin{aligned} A_2 &= A_{\frac{3}{2}} \bigcup \{(1, 1, 0), (1, 0, 1), (0, 1, 1), (1, 1, 1, 1)\} \\ A_2^0 &= \{\emptyset, (0), (1, 0), (0, 0), (1, 1, 0)\}, \end{aligned}$$

and the corresponding RODE-Taylor scheme of order 2 reads

$$\begin{aligned} x_{n+1} &= x_n + gh_n + \sigma g_y I_{(1,0),n} + \frac{1}{2} \left(gg_x + \sigma g_y + \frac{1}{2} (\sigma)^2 g_{yy} \right) h_n^2 \\ &\quad + (\sigma \sigma_y g_y + (\sigma)^2 g_{yy}) I_{(1,1,0),n}. \end{aligned} \tag{10.13}$$

Note that the schemes (10.12) and (10.13) include coefficients of the SODE of the driving noise as well as an additional multiple stochastic integral.

10.1.4 Derivative-Free Explicit Strong Schemes

The order γ strong Taylor schemes involve derivatives of the coefficients functions of the SODE (10.11) that may be difficult to determine in higher dimensional examples. The derivative-free explicit strong schemes in Kloeden and Platen [91] are Runge-Kutta like schemes that replace such derivatives by appropriate finite difference quotients to ensure the same order γ strong convergence.

For a single Wiener process, i.e., $m = 1$, the k th component of the explicit order $\gamma = 1.0$ strong scheme has the form

$$\begin{aligned} v_{n+1}^k &= v_n^k + \vec{f}^k(v_n)h_n + \vec{\sigma}^k(v_n)\Delta W_n \\ &\quad + \frac{1}{2\sqrt{h_n}}\{\vec{\sigma}^k(\tilde{v}_n) - \vec{\sigma}^k(v_n)\}\{(\Delta W_n)^2 - h_n\}, \end{aligned}$$

with the support function

$$\tilde{v}_n = v_n + \vec{f}(v_n)h_n + \vec{\sigma}(v_n)\sqrt{h_n}.$$

The corresponding RODE scheme is given componentwise by

$$x_{n+1}^k = x_n^k + g^k(x_n, y_n)h_n, \quad k = 0, \dots, d_1$$

which is the same as the Euler–Maruyama and Milstein schemes.

The componentwise order $\gamma = 1.5$ explicit strong scheme is

$$\begin{aligned} v_{n+1}^k &= v_n^k + \vec{f}^k(v_n)h_n + \vec{\sigma}^k(v_n)\Delta W_n \\ &\quad + \frac{1}{2\sqrt{h_n}} \left[(\vec{f}^k(v_+) - \vec{f}^k(v_-)) I_{(1,0),n} + (\vec{\sigma}^k(v_+) - \vec{\sigma}^k(v_-)) I_{(1,1),n} \right] \\ &\quad + \frac{1}{2h_n} \left[(\vec{f}^k(v_+) - 2\vec{f}^k(v_n) + \vec{f}^k(v_-)) I_{(0,0),n} \right. \\ &\quad \left. + (\vec{\sigma}^k(v_+) - 2\vec{\sigma}^k(v_n) + \vec{\sigma}^k(v_-)) I_{(0,1),n} \right] \\ &\quad + \frac{1}{2h_n} [\vec{\sigma}^k(\tilde{v}_+) - \vec{\sigma}^k(\tilde{v}_-) - \vec{\sigma}^k(v_+) + \vec{\sigma}^k(v_-)] I_{(1,1,1),n}, \end{aligned}$$

where

$$\begin{aligned} v_{\pm} &= v_n + \vec{f}(v_n)h_n \pm \vec{\sigma}(v_n)\sqrt{h_n}, \\ \tilde{v}_{\pm} &= v_{\pm} \pm \vec{\sigma}(v_{\pm})\sqrt{h_n}. \end{aligned}$$

Since $\vec{\sigma}^1 \equiv 0$ for $k = 1$ the x -component gives the RODE scheme

$$\begin{aligned} x_{n+1} &= x_n + g(x_n, y_n)h_n + \frac{1}{2\sqrt{h_n}} [g(\tilde{x}, \tilde{y}_+) - g(\tilde{x}, \tilde{y}_-)] I_{(1,0),n} \\ &\quad + \frac{1}{4} [g(\tilde{x}, \tilde{y}_+) - 2g(x_n, y_n) + g(\tilde{x}, \tilde{y}_-)] h_n \end{aligned}$$

where

$$\tilde{x} = x_n + g(x_n, y_n)h_n, \quad \tilde{y}_{\pm} = y_n + f(y_n)h_n \pm \sigma(y_n)\sqrt{h_n}.$$

10.2 Implicit Strong Schemes

There are many differential equations that arise often in practice for which explicit schemes may perform poorly, while implicit schemes offer better numerical stability properties.

The order $\gamma = 1.0$ implicit strong Taylor scheme given in Kloeden and Platen [91] is a drift-implicit version of the Milstein scheme. For the vector SODE (10.11) with single Wiener process, the componentwise order 1.0 implicit strong Taylor scheme reads

$$\begin{aligned} v_{n+1}^k &= v_n^k + \left\{ \theta \vec{f}^k(v_{n+1}) + (1 - \theta) \vec{f}^k(v_n) \right\} h_n \\ &\quad + \vec{\sigma}^k(v_n) \Delta W_n + L^1 \vec{\sigma}^k(v_n) I_{(1,1),n}, \quad k = 1, \dots, d, \end{aligned}$$

where the parameter $\theta \in [0, 1]$ characterises the degree of implicitness. (When $\theta = 0$, it reduces to usual explicit Milstein scheme).

For the scalar RODE (10.4) its x -component gives the order $\gamma = 1.0$ implicit strong RODE-Taylor scheme

$$x_{n+1} = x_n + [\theta g(x_{n+1}, y_{n+1}) + (1 - \theta)g(x_n, y_n)] h_n,$$

which is often called the *theta-scheme* in the context of ODEs.

The family of $\gamma = 1.5$ implicit strong Taylor scheme for the vector SODE (10.11) with single Wiener process given in [91] is, componentwise,

$$\begin{aligned} v_{n+1}^k &= v_n^k + \left[\theta_1 \vec{f}^k(v_{n+1}) + (1 - \theta_1) \vec{f}^k(v_n) \right] h_n \\ &\quad + \left(\frac{1}{2} - \theta_1 \right) \left[\theta_2 L^0 \vec{f}^k(v_{n+1}) + (1 - \theta_2) L^0 \vec{f}^k(v_n) \right] h_n^2 \\ &\quad + \vec{\sigma}^k(v_n) \Delta W_n + L^0 \vec{\sigma}^k(v_n) I_{(0,1),n} + L^1 \vec{f}^k(v_n) (I_{(1,0),n} - \theta_1 \Delta W_n h_n) \\ &\quad + L^1 \vec{\sigma}^k(v_n) I_{(1,1),n} + L^1 L^1 \vec{\sigma}^k(v_n) I_{(1,1,1),n}, \end{aligned}$$

where the parameters $\theta_1, \theta_2 \in [0, 1]$ indicate the extent of implicitness. The corresponding implicit RODE-Taylor scheme then reads

$$\begin{aligned} x_{n+1} &= x_n + [\theta_1 g(x_{n+1}, y_{n+1}) + (1 - \theta_1)g(x_n, y_n)] h_n \\ &\quad + \left(\frac{1}{2} - \theta_1 \right) [\theta_2 L^0 g(x_{n+1}, y_{n+1}) + (1 - \theta_2)L^0 g(x_n, y_n)] h_n^2 \\ &\quad + L^1 g(x_n, y_n) (I_{(1,0),n} - \theta_1 \Delta W_n h_n). \end{aligned} \tag{10.14}$$

Derivative-free implicit scheme

When $\theta_1 = \frac{1}{2}$, the implicit RODE-Taylor scheme (10.14) reduces to

$$x_{n+1} = x_n + \frac{1}{2} [g(x_{n+1}, y_{n+1}) + g(x_n, y_n)] h_n + L^1 g(x_n, y_n) \left(I_{(1,0),n} - \frac{1}{2} \Delta W_n h_n \right).$$

Replacing the derivatives by finite difference quotients gives the derivative-free implicit scheme with $\gamma = 1.5$:

$$\begin{aligned} x_{n+1} &= x_n + \frac{1}{2} [g(x_{n+1}, y_{n+1}) + g(x_n, y_n)] h_n \\ &\quad + \frac{1}{2\sqrt{h_n}} [g(\tilde{x}, \tilde{y}_+) - g(\tilde{x}, \tilde{y}_-)] \left(I_{(1,0),n} - \frac{1}{2} \Delta W_n h_n \right) \end{aligned}$$

where

$$\tilde{x} = x_n + g(x_n, y_n)h_n, \quad \tilde{y}_{\pm} = y_n + f(y_n)h_n \pm \sigma(y_n)\sqrt{h_n}.$$

10.3 Multi-step Schemes

For simplicity first consider the SODE (10.7), based on the coupled RODE-SODE in \mathbb{R}^2

$$dU_t = \vec{f}(U_t)dt + \vec{\sigma}(U_t)dW_t. \quad (10.15)$$

An Euler–Maruyama type linear s -step method for (10.15) is given by

$$\sum_{j=0}^s a_j u_{n-j} = h \sum_{j=0}^s b_j \vec{f}_{n-j} + \Delta W_n \sum_{j=1}^s c_j \vec{\sigma}_{n-j},$$

where u_{n-j} is an approximation of U at t_{n-j} , and \vec{f}_{n-j} and $\vec{\sigma}_{n-j}$ are functions \vec{f} and $\vec{\sigma}$ evaluated at u_{n-j} , respectively. Since $\vec{\sigma}^{-1} \equiv 0$ its x -component reduces to

$$\sum_{j=0}^s a_j x_{n-j} = h \sum_{j=0}^s b_j g(x_{n-j}, y_{n-j}). \quad (10.16)$$

Typical examples of (10.16) are the Adams–Bashforth and Adams–Moulton methods which coincide with their counterparts in the deterministic case. In the scalar case, the x -component gives the *RODE-Adams–Bashforth-2 scheme*

$$x_{n+1} = x_n + \left\{ \frac{3}{2} g(x_n, y_n) - \frac{1}{2} g(x_{n-1}, y_{n-1}) \right\} h_n$$

and the *RODE–Adams–Moulton-2 scheme*

$$x_{n+1} = x_n + \left\{ \frac{5}{12}g(x_{n+1}, y_{n+1}) + \frac{8}{12}g(x_n, y_n) - \frac{1}{12}g(x_{n-1}, y_{n-1}) \right\} h_n.$$

These both have order $\gamma = 1.0$ convergence.

Higher order multi-step schemes can be derived using strong Itô–Taylor expansions and require multiple stochastic integral terms to achieve a higher order of strong convergence. For the special case of the SODE (10.11) based on a RODE–SODE pair with the diffusion coefficient functions $\vec{\sigma}_j(V)$, the differential operators reduce to $L^j \chi^k(V) \equiv 0$ for $k \leq d_1$, i.e., the components of the RODE. Hence only the multi-indices in the subset

$$\Lambda_\gamma^0 = \{\alpha \in \Lambda_\gamma : \text{the last component } j_l = 0\} \cup \{\emptyset\}$$

of Λ_γ appear in these RODE components of the stochastic Taylor expansion. The corresponding remainder set is

$$\mathcal{R}(\Lambda_\gamma^0) = \{\alpha \in \mathcal{J}_m \setminus \Lambda_\gamma^0 : -\alpha \in \Lambda_\gamma^0\}.$$

Note that $(j) \in \mathcal{R}(\Lambda_\gamma^0)$ for $j = 1, \dots, m$, since $(j) \notin \Lambda_\gamma^0$ but $-(j) = \emptyset \in \Lambda_\gamma^0$.

For example, when $m = 1$ and $\gamma = \frac{3}{2}$, the sets Λ_γ , $\mathcal{R}(\Lambda_\gamma)$ and the reduced sets Λ_γ^0 , $\mathcal{R}(\Lambda_\gamma^0)$ are given by

$$\begin{aligned} \Lambda_{3/2} &= \{\emptyset, (1), (0), (1, 1), (1, 0), (0, 1), (0, 0), (1, 1, 1)\} \\ \mathcal{R}(\Lambda_{3/2}) &= \left\{ (0, 1, 1), (1, 1, 0), (0, 1, 0), (1, 0, 1), (0, 0, 1), \right. \\ &\quad \left. (1, 0, 0), (0, 0, 0), (1, 1, 1, 1), (0, 1, 1, 1) \right\} \\ \Lambda_{3/2}^0 &= \{\emptyset, (0), (1, 0), (0, 0)\} \\ \mathcal{R}(\Lambda_{3/2}^0) &= \{(1), (1, 1, 0), (0, 1, 0), (1, 0, 0), (0, 0, 0)\}. \end{aligned}$$

The Taylor expansions for the Itô diffusion components of the SODE (10.11) and other functions of the solutions still require all of the multi-indices in the original hierarchical set Λ_γ . It is decoupled from the RODE components of the Taylor scheme.

Definition 10.1 An order γ s -step stochastic linear multi-step method (SLMM) for the multi-dimensional Itô SODE (10.11) has the general form

$$\begin{aligned} \sum_{j=0}^s C_{\emptyset, j} x_{n-j} &= h \sum_{j=0}^s C_{(0), j} g(x_{n-j}, \hat{y}_{n-j}) \\ &+ \sum_{j=1}^s \sum_{\alpha \in \Lambda_\gamma^0 \setminus \{\emptyset, (0)\}} L^\alpha \chi^1(V)(x_{n-j}, \hat{y}_{n-j}) (C_{\alpha, j} I_{\alpha, n-j} + C_{\alpha, j}^* I_{\alpha-, n-j} h), \end{aligned}$$

where the consistency conditions are given as

$$\begin{cases} \sum_{j=0}^s C_{\emptyset,j} = 0, & \sum_{j=0}^s (s-j)C_{\emptyset,j} = \sum_{j=0}^s C_{(0),j}, \\ C_{\alpha,i} = \sum_{j=0}^{i-1} C_{\emptyset,j} & \text{for } i = 1, \dots, s, \\ C_{\alpha,j}^* = \sum_{j=0}^{i-1} ((s-1-j)C_{\emptyset,j} - C_{(0),j}) & \text{for } i = 1, \dots, s. \end{cases} \quad (10.17)$$

for $\alpha \in \Lambda_\gamma^0 \setminus \{\emptyset, (0)\}$.

Here \hat{y}_{n-j} is an approximation of Y_t at t_{n-j} by a scheme of high enough order or Y_{t_n} itself when Y_t can be generated exactly. The scheme reduces to an explicit scheme when $C_{(0),0} = 0$. By the order conditions this happens when, e.g., $C_{\alpha,1}^* = 0$ for all $\alpha \in \Lambda_\gamma^0 \setminus \{\emptyset, (0)\}$.

Derivation of higher order multi-step schemes

Higher order multi-step schemes for RODEs driven by an Itô diffusion process are derived using strong stochastic Taylor expansions for the SODE formed by coupled RODE-SODE pair. They involve partial derivatives of the coefficient functions, but, due to the special structure of the coupled pair (i.e., no diffusion coefficient in the RODE part) they do not require the intensity of the noise to be restricted as for general SODEs (see [19]). Their derivation and that of corresponding order conditions will be illustrated here via the scalar RODE driven by a scalar SODE (10.4), in the form of the Itô SODEs (10.5) and (10.6).

In order to develop higher order schemes for (10.4), the local error of the linear s -step scheme (10.16) will first be estimated. When $s = 2$, the local error \mathcal{E}_n^L of (10.16) is given by

$$\mathcal{E}_n^L := \left| \sum_{j=0}^2 a_j X_{t_{n-j}} - h \sum_{j=0}^2 b_j g(X_{t_{n-j}}, Y_{t_{n-j}}) \right|. \quad (10.18)$$

Estimations of the local error require additional stochastic integrals $\hat{I}_{\alpha,t_{n-k}}^{t_n}$ defined by combinations of stochastic integrals I_α between t_{n-k} to t_n . For example, for $\alpha = (1, 1, 0)$ and $k = 3$, the stochastic integral $\hat{I}_{(1,1,0),t_{n-3}}^{t_n}$ is given by

$$\begin{aligned} \hat{I}_{(1,1,0),t_{n-3}}^{t_n}[\cdot] &= I_{(1,1,0),n-3}[\cdot] + I_{(1,1),n-3}[\cdot](I_{(0),n-2} + I_{(0),n-1}) \\ &\quad + I_{(1),n-3}[\cdot](I_{(1,0),n-2} + I_{(1),n-2}I_{(0),n-1} + I_{(1,0),n-1}) \\ &\quad + I_{(1,1,0),n-2}[\cdot] + I_{(1,1),n-2}[\cdot]I_{(0),n-1} \\ &\quad + I_{(1),n-2}[\cdot]I_{(1,0),n-1} + I_{(1,1,0),n-1}[\cdot]. \end{aligned}$$

Order 1.0 consistency conditions

The hierarchical and remainder sets for $\gamma = 1$ are $\Lambda_1^0 = \{\emptyset, (0)\}$ and $\mathcal{R}(\Lambda_1^0) = \{(1), (0, 0)(1, 0)\}$, respectively. Then the corresponding Itô–Taylor expansions are

$$\begin{aligned} X_{t_n} &= X_{t_{n-1}} + g(X_{t_{n-1}}, Y_{t_{n-1}}) I_{(0), n-1} + \sum_{\alpha \in \mathcal{R}(\Lambda_1^0)} I_{\alpha, n-1} [L^\alpha \chi^1(V)], \\ X_{t_{n-1}} &= X_{t_{n-2}} + g(X_{t_{n-2}}, Y_{t_{n-2}}) I_{(0), n-2} + \sum_{\alpha \in \mathcal{R}(\Lambda_1^0)} I_{\alpha, n-2} [L^\alpha \chi^1(V)]. \end{aligned}$$

Since the g terms in the local error are already multiplied by h they only need to be expanded to a lower order, i.e., for the hierarchical and remainder sets $\Lambda_0 = \{\emptyset\}$ and $\mathcal{R}(\Lambda_0) = \{(0), (1)\}$. Thus

$$\begin{aligned} g(X_{t_n}, Y_{t_n}) &= g(X_{t_{n-1}}, Y_{t_{n-1}}) + \sum_{\alpha \in \mathcal{R}(\Lambda_0)} I_{\alpha, n-1} [L^\alpha g], \\ g(X_{t_{n-1}}, Y_{t_{n-1}}) &= g(X_{t_{n-2}}, Y_{t_{n-2}}) + \sum_{\alpha \in \mathcal{R}(\Lambda_0)} I_{\alpha, n-2} [L^\alpha g]. \end{aligned}$$

Note that since $L^1 \chi^1(V) = 0$, the term for $\alpha = (1)$ vanishes. Thus the above expressions are equivalent to

$$\begin{aligned} g(X_{t_n}, Y_{t_n}) &= g(X_{t_{n-1}}, Y_{t_{n-1}}) + \sum_{\alpha \in \mathcal{R}(\Lambda_1^0)} I_{(j_1(\alpha)), n-1} [L^\alpha \chi^1(V)], \\ g(X_{t_{n-1}}, Y_{t_{n-1}}) &= g(X_{t_{n-2}}, Y_{t_{n-2}}) + \sum_{\alpha \in \mathcal{R}(\Lambda_1^0)} I_{(j_1(\alpha)), n-2} [L^\alpha \chi^1(V)], \end{aligned}$$

where $j_1(\alpha)$ is the first component of α .

Substituting X_{t_n} , $X_{t_{n-1}}$, $g(X_{t_n}, Y_{t_n})$ and $g(X_{t_{n-1}}, Y_{t_{n-1}})$ in (10.18) results in

$$\begin{aligned} \mathcal{E}_n^L &= |(a_0 + a_1 + a_2) X_{t_{n-2}} \\ &\quad + h (2a_0 + a_1 - (b_0 + b_1 + b_2)) g(X_{t_{n-2}}, Y_{t_{n-2}}) + R_1|. \end{aligned}$$

The remainder term R_1 is given by

$$\begin{aligned} R_1 &= \sum_{\alpha \in \mathcal{R}(\Lambda_1^0)} \left\{ a_0 \hat{I}_{\alpha, t_{n-2}}^{t_n} [L^\alpha \chi^1(V)] + a_1 I_{\alpha, n-2} [L^\alpha \chi^1(V)] \right. \\ &\quad \left. - h \left(b_0 \hat{I}_{(j_1(\alpha)), t_{n-2}}^{t_n} [L^\alpha \chi^1(V)] + b_1 I_{(j_1(\alpha)), n-2} [L^\alpha \chi^1(V)] \right) \right\}. \end{aligned}$$

When

$$a_0 + a_1 + a_2 = 0, \quad 2a_0 + a_1 - (b_0 + b_1 + b_2) = 0, \quad (10.19)$$

the local error $\mathcal{E}_n^L \leq C(h^{3/2})$ and the SLMMs (10.16) with the consistency conditions (10.19) satisfy order 1 convergence. Here the consistency conditions (10.19) are the same as the deterministic ones.

Order 1.5 consistency conditions

In order to achieve the same order as the order 1.5 Itô–Taylor scheme, one has to deal with the remainder term R_1 . In this case $\mathcal{R}(A_1^0) \setminus \{(1)\} = \{(1, 0), (0, 0)\}$ and $L^1 g$ and $L^0 g$ terms appear in R_1 .

Consider the SLMM in the following form:

$$\begin{aligned} \sum_{j=0}^2 a_j x_{n-j} &= \sum_{j=0}^2 b_j g(x_{n-j}, Y_{t_{n-j}}) h + \sum_{j=1}^2 L^1 g(x_{n-j}, Y_{t_{n-j}}) (c_j I_{(1,0),n-j} + \tilde{c}_j I_{(1),n-j} h) \\ &\quad + \sum_{j=1}^2 L^0 g(x_{n-j}, Y_{t_{n-j}}) (d_j I_{(0,0),n-j} + \tilde{d}_j I_{(0),n-j} h). \end{aligned} \quad (10.20)$$

Then the local error \mathcal{E}_n^L of (10.20) is given by

$$\begin{aligned} \mathcal{E}_n^L &= \left| \sum_{j=0}^2 a_j X_{t_{n-j}} - \sum_{j=0}^2 b_j g(X_{t_{n-j}}, Y_{t_{n-j}}) h \right. \\ &\quad - \sum_{j=1}^2 L^1 g(X_{t_{n-j}}, Y_{t_{n-j}}) (c_j I_{(1,0),n-j} + \tilde{c}_j I_{(1),n-j} h) \\ &\quad \left. - \sum_{j=1}^2 L^0 g(X_{t_{n-j}}, Y_{t_{n-j}}) (d_j I_{(0,0),n-j} + \tilde{d}_j I_{(0),n-j} h) \right|. \end{aligned} \quad (10.21)$$

The hierarchical set for $\gamma = 1.5$ is now given by

$$A_{3/2}^0 = \{\emptyset, (0), (1, 0), (0, 0)\}$$

and the corresponding Itô–Taylor expansions are

$$\begin{aligned} X_{t_n} &= \sum_{\alpha \in A_{3/2}^0} L^\alpha \chi^1(V)(X_{t_{n-1}}, Y_{t_{n-1}}) I_{\alpha,n-1} + \sum_{\alpha \in \mathcal{R}(A_{3/2}^0) \setminus \{(1)\}} I_{\alpha,n-1}[L^\alpha \chi^1(V)], \\ X_{t_{n-1}} &= \sum_{\alpha \in A_{3/2}^0} L^\alpha \chi^1(V)(X_{t_{n-2}}, Y_{t_{n-2}}) I_{\alpha,n-2} + \sum_{\alpha \in \mathcal{R}(A_{3/2}^0) \setminus \{(1)\}} I_{\alpha,n-2}[L^\alpha \chi^1(V)]. \end{aligned}$$

For $A_{1/2} = \{\emptyset, (0), (1)\}$,

$$\mathcal{R}(A_{1/2}) = \{(0, 0), (1, 0), (0, 1), (1, 1)\},$$

and thus

$$g(X_{t_n}, Y_{t_n}) = g(X_{t_{n-1}}, Y_{t_{n-1}}) + L^1 g(X_{t_{n-1}}, Y_{t_{n-1}}) I_{(1),n-1}$$

$$\begin{aligned}
& + L^0 g(X_{t_{n-1}}, Y_{t_{n-1}}) I_{(0), n-1} + \sum_{\alpha \in \mathcal{R}(\Lambda_{1/2})} I_{\alpha, n-1} [L^\alpha g] \\
g(X_{t_{n-1}}, Y_{t_{n-1}}) & = g(X_{t_{n-2}}, Y_{t_{n-2}}) + L^1 g(X_{t_{n-2}}, Y_{t_{n-2}}) I_{(1), n-2} \\
& + L^0 g(X_{t_{n-2}}, Y_{t_{n-2}}) I_{(0), n-2} + \sum_{\alpha \in \mathcal{R}(\Lambda_{1/2})} I_{\alpha, n-2} [L^\alpha g],
\end{aligned}$$

and

$$\begin{aligned}
L^1 g(X_{t_{n-1}}, Y_{t_{n-1}}) & = L^1 g(X_{t_{n-2}}, Y_{t_{n-2}}) + I_{(1), n-2} [L^1 L^1 g] + I_{(0), n-2} [L^0 L^1 g] \\
L^0 g(X_{t_{n-1}}, Y_{t_{n-1}}) & = L^0 g(X_{t_{n-2}}, Y_{t_{n-2}}) + I_{(1), n-2} [L^1 L^0 g] + I_{(0), n-2} [L^0 L^0 g].
\end{aligned}$$

Substituting X_{t_n} , $X_{t_{n-1}}$, $g(X_{t_n}, Y_{t_n})$, $g(X_{t_{n-1}}, Y_{t_{n-1}})$, $L^1 g(X_{t_{n-1}}, Y_{t_{n-1}})$ and $L^0 g(X_{t_{n-1}}, Y_{t_{n-1}})$ into (10.21) yields the local error

$$\begin{aligned}
\mathcal{E}_n^L = & \left| \sum_{j=0}^2 a_j X_{t_{n-2}} + (2a_0 + a_1 - b_0 - b_1 - b_2) h g(X_{t_{n-2}}, Y_{t_{n-2}}) \right. \\
& + \left(a_0 \hat{I}_{(1,0), t_{n-2}}^{t_n} + a_1 I_{(1,0), n-2} - b_0 \hat{I}_{(1), t_{n-2}}^{t_n} h - b_1 I_{(1), n-2} h - c_1 I_{(1,0), n-1} \right. \\
& - c_2 I_{(1,0), n-2} - \tilde{c}_1 I_{(1), n-1} h - \tilde{c}_2 I_{(1), n-2} h \Big) L^1 g(X_{t_{n-2}}, Y_{t_{n-2}}) \\
& + \left(a_0 \hat{I}_{(0,0), t_{n-2}}^{t_n} + a_1 I_{(0,0), n-2} - b_0 \hat{I}_{(0), t_{n-2}}^{t_n} h - b_1 I_{(0), n-2} h - d_1 I_{(0,0), n-1} \right. \\
& \left. \left. - d_2 I_{(0,0), n-2} - \tilde{d}_1 I_{(0), n-1} h - \tilde{d}_2 I_{(0), n-2} h \right) L^0 g(X_{t_{n-2}}, Y_{t_{n-2}}) + R_2 \right|.
\end{aligned}$$

Note that

$$\mathcal{R}(\Lambda_{1/2}) = \{\alpha : (\alpha, 0) \in \mathcal{R}(\Lambda_{3/2}^0) \setminus \{(1)\}\},$$

and $L^{(\alpha,0)} \chi^1(V) = L^\alpha L^0 \chi^1(V) = L^\alpha g$ while $L^1 \chi^1(V) = 0$. Then the remainder term R_2 can be written as

$$\begin{aligned}
R_2 = & \sum_{\alpha \in \mathcal{R}(\Lambda_{3/2}^0) \setminus \{(1)\}} \left\{ a_0 \hat{I}_{\alpha, t_{n-2}}^{t_n} [L^\alpha \chi^1(V)] + a_1 I_{\alpha, n-2} [L^\alpha \chi^1(V)] \right. \\
& - b_0 \hat{I}_{(j_1(\alpha), j_2(\alpha)), t_{n-2}}^{t_n} [L^\alpha \chi^1(V)] h - b_1 I_{(j_1(\alpha), j_2(\alpha)), n-2} [L^\alpha \chi^1(V)] h \\
& - c_1 I_{(j_1(\alpha)), n-2} [L^\alpha \chi^1(V)] I_{(j_2(\alpha), j_3(\alpha)), n-1} - \tilde{c}_1 I_{(j_1(\alpha)), n-2} [L^\alpha \chi^1(V)] I_{(j_2(\alpha)), n-1} h \\
& \left. - d_1 I_{(j_1(\alpha)), n-2} [L^\alpha \chi^1(V)] I_{(j_2(\alpha), j_3(\alpha)), n-1} - \tilde{d}_1 I_{(j_1(\alpha)), n-2} [L^\alpha \chi^1(V)] I_{(j_2(\alpha)), n-1} h \right\},
\end{aligned}$$

where $j_i(\alpha)$ is the i th component of α .

In addition, the terms with c_1 and \tilde{c}_1 appear when $\alpha = (1, 1, 0)$ or $(0, 1, 0)$, and the terms with d_1 and \tilde{d}_1 appear when $\alpha = (1, 0, 0)$ or $(0, 0, 0)$. The coefficients of $L^1 g(X_{t_{n-2}}, Y_{t_{n-2}})$ in \mathcal{E}_n^L can be reduced to

$$(a_0 + a_1 - c_2)I_{(1,0),n-2} + (a_0 - c_1)I_{(1,0),n-1} \\ + (a_0 - b_0 - b_1 - \tilde{c}_2)I_{(1),n-2}h - (b_0 + \tilde{c}_1)I_{(1),n-1}h.$$

Similarly the coefficients of $L^0 g(X_{t_{n-2}}, Y_{t_{n-2}})$ can be reduced to

$$(a_0 + a_1 - d_2)I_{(0,0),n-2} + (a_0 - d_1)I_{(0,0),n-1} \\ + (a_0 - b_0 - b_1 - \tilde{d}_2)I_{(0),n-2}h - (b_0 + \tilde{d}_1)I_{(0),n-1}h.$$

Hence if the following consistency conditions

$$\begin{cases} c_1 = d_1 = a_0, & c_2 = d_2 = a_0 + a_1, \\ \tilde{c}_1 = \tilde{d}_1 = -b_0, & \tilde{c}_2 = \tilde{d}_2 = a_0 - b_0 - b_1, \end{cases} \quad (10.22)$$

are satisfied, the coefficients of $L^1 g(X_{t_{n-2}}, Y_{t_{n-2}})$ and $L^0 g(X_{t_{n-2}}, Y_{t_{n-2}})$ terms will both vanish and the SLMMs (10.20) satisfy 1.5-order convergence.

Order 2.0 consistency conditions

Following a similar process as above,

$$\Lambda_2^0 = \{\emptyset, (0), (1, 0), (0, 0), (1, 1, 0)\}$$

and a 2-step SLMM of order 2.0 reads

$$\sum_{j=0}^2 a_j x_{n-j} = h \sum_{j=0}^2 b_j g(x_{n-j}, Y_{t_{n-j}}) + \sum_{j=1}^2 L^1 g(x_{n-j}, Y_{t_{n-j}}) \left(c_j I_{(1,0),n-j} + \tilde{c}_j I_{(1),n-j}h \right) \\ + \sum_{j=1}^2 L^0 g(x_{n-j}, Y_{t_{n-j}}) \left(d_j I_{(0,0),n-j} + \tilde{d}_j I_{(0),n-j}h \right) \\ + \sum_{j=1}^2 L^1 L^1 g(x_{n-j}, Y_{t_{n-j}}) \left(e_j I_{(1,1,0),n-j} + \tilde{e}_j I_{(1,1),n-j}h \right).$$

under the consistency conditions

$$\begin{cases} e_1 = a_0, & e_2 = a_0 + a_1, \\ \tilde{e}_1 = -b_0, & \tilde{e}_2 = a_0 - b_0 - b_1. \end{cases} \quad (10.23)$$

The higher order SLMMs in Definition 10.1 with more steps can be generated in the same manner.

Example 10.1 The order 1.5 implicit 2-step scheme for the scalar RODE (10.4) driven by a scalar Itô stochastic differential equation is

$$\begin{aligned}
x_n &= \frac{1}{2}x_{n-1} + \frac{1}{2}x_{n-2} + h \left(\frac{1}{2}g_n + \frac{1}{2}g_{n-1} + \frac{1}{2}g_{n-2} \right) \\
&\quad + \left(L^1 g_{n-1} I_{(1,0),n-1} + \frac{1}{2} L^1 g_{n-2} I_{(1,0),n-2} \right) - \frac{1}{2} L^1 g_{n-1} h I_{(1),n-1} \\
&\quad + \left(\frac{1}{4} L^0 g_{n-1} I_{(0,0),n-1} + \frac{1}{4} L^0 g_{n-2} I_{(0,0),n-2} \right),
\end{aligned} \tag{10.24}$$

where the coefficients g_{n-j} are the values of function g at (x_{n-j}, \hat{y}_{n-j}) for $j = 0, 1$ and 2 , and \hat{y}_{n-j} is an approximation of Y_t at t_{n-j} by a scheme of high enough order or $Y_{t_{n-j}}$ itself if it can be generated exactly.

The coefficient constants

$$\begin{aligned}
a_0 &= 1, \quad a_1 = -\frac{1}{2}, \quad a_2 = -\frac{1}{2}, \quad b_0 = \frac{1}{2}, \quad b_1 = \frac{1}{2}, \quad b_2 = \frac{1}{2}, \\
c_1 &= 1, \quad c_2 = \frac{1}{2}, \quad d_1 = -\frac{1}{2}, \quad d_2 = 0, \quad e_1 = \frac{1}{4}, \quad e_2 = \frac{1}{4},
\end{aligned}$$

satisfy the consistency conditions (10.19) and (10.22) for convergence order 1.5, but not the conditions (10.23) for order 2.0.

Note that the multi-step scheme (10.24) contains the coefficient functions f and σ of the Itô SDE for the driving process through the operators L^0 and L^1 applied to the RODE vector field g , specifically

$$\begin{aligned}
L^0 g(x, y) &= g(x, y) \frac{\partial g}{\partial x}(x, y) + f(y) \frac{\partial g}{\partial y}(x, y) + \frac{1}{2} \sigma^2(y) \frac{\partial^2 g}{\partial y^2}(x, y), \\
L^1 g(x, y) &= \sigma(y) \frac{\partial g}{\partial y}(x, y).
\end{aligned}$$

10.4 RODEs with Affine Noise

When a RODE has affine noise, many higher order terms in the above Taylor schemes will vanish and the schemes can be reduced to ones with simpler structures. This will be illustrated below in terms of the scalar RODE (10.4) with affine noise, i.e., the vector field g is of the form

$$g(x, y) = p(x) + q(x)y.$$

Since the partial derivatives of g are

$$g_x(x, y) = p_x(x) + q_x(x)y, \quad g_y(x, y) = q(x),$$

the second derivative term vanishes in L^0 and gives

$$L^0 = \frac{\partial}{\partial t} + g(x, y) \frac{\partial}{\partial x} + f(y) \frac{\partial}{\partial y}$$

for function g and its partial derivatives. In particular,

$$\begin{aligned} L^0 g(x, y) &= g(x, y)g_x(x, y) + f(y)g_y(x, y) \\ &= g(x, y)(p_x(x) + q_x(x)y) + f(y)q(x), \end{aligned}$$

and

$$L^0 g_y(x, y) = g(x, y)g_{xy}(x, y) + f(y)g_{yy}(x, y) = g(x, y)q_x(x),$$

where the subscripts denote partial derivatives.

As a result, for example, the order $\frac{3}{2}$ RODE-Taylor scheme (10.12) simplifies to

$$\begin{aligned} x_{n+1} &= x_n + g(x_n, y_n)h_n + \sigma(y_n)q(x_n) I_{(1,0),n} \\ &\quad + \frac{1}{2} (g(x_n, y_n)g_x(x_n, y_n) + f(y_n)q(x_n)) h_n^2, \end{aligned}$$

since $g_y = q$ and $g_{yy} = 0$.

10.5 Proof of Theorem 10.1

The proof is based on that of Theorem 6.2, due to Jentzen, Kloeden and Neuenkirch [80], combined with a localisation argument.

Proof Let N be some sufficiently large number and define the stopping times

$$\begin{aligned} \tau^{(N)}(\omega) &= \inf \{t \geq 0 : |X_t(\omega)| > N\} \wedge 2T, \\ \tau_n^{(N)}(\omega) &= \inf \{t \geq 0 : |x_n(t, \omega)| > N\} \wedge 2T. \end{aligned}$$

The interval $[0, 2T]$ is used here to handle stopping times and other technical issues, but the result will be restricted later to the smaller interval $[0, T]$. Then $\tau^{(N)} \rightarrow 2T$ almost surely as $N \rightarrow \infty$ since the sample paths of X_t are continuous functions.

Choose a function $\pi_N \in \mathcal{C}_b^\infty(\mathbb{R}^{d_1}; [0, 1])$ such that

$$\pi_N(x) = \begin{cases} 1, & |x| \leq N, \\ 0, & |x| > N + 1 \end{cases}$$

and define the truncated function $g_{\pi_N} \in \mathcal{C}_b^{2\gamma+1}$ by

$$g_{\pi_N}(x, y) = g(x, y) \cdot \pi_N(x).$$

Consider the truncated degenerate SODE

$$dZ_t = g_{\pi_N}(Z_t, Y_t) dt + 0 dW_t, \quad (10.25)$$

with the initial value

$$Z_0(\omega) = \begin{cases} X_0(\omega), & |X_0(\omega)| \leq N, \\ \frac{N}{\|X_0(\omega)\|} X_0(\omega), & |X_0| > N. \end{cases}$$

Coupling (10.25) with the driving SODE (10.2) for Y_t forms the modified system of SODEs in $\mathbb{R}^{d_1+d_2}$

$$d \begin{pmatrix} Z_t \\ Y_t \end{pmatrix} = \begin{pmatrix} g_{\pi_N}(Z_t, Y_t) \\ f(Y_t) \end{pmatrix} dt + \begin{pmatrix} 0 \\ \sigma(Y_t) \end{pmatrix} dW_t. \quad (10.26)$$

Denote by z_n the Itô–Taylor approximation on $[0, 2T]$ of order γ applied to the truncated equation (10.25), and denote by $\mathbf{1}_A$ the indicator function, i.e., $\mathbf{1}_A(x) = 1$ for $x \in A$ and $\mathbf{1}_A(x) = 0$ for $x \notin A$.

First, it is clear that

$$\begin{aligned} \tau_n^{(N)}(\omega) &= \inf \{t \geq 0 : x_n(t, \omega) > N\} \wedge 2T \\ &= \inf \{t \geq 0 : z_n(t, \omega) > N\} \wedge 2T, \end{aligned}$$

thus

$$z_n(t \wedge \tau_n^{(N)}) \mathbf{1}_{\{\tau_n^{(N)} > 0\}} = x_n(t \wedge \tau_n^{(N)}) \mathbf{1}_{\{\tau_n^{(N)} > 0\}}, \quad t \in [0, 2T], \quad \text{a.s.} \quad (10.27)$$

Second, as shown in the appendix of [80],

$$X_{t \wedge \tau_n^{(N)}} \mathbf{1}_{\{\tau_n^{(N)} > 0\}} = Z_{t \wedge \tau_n^{(N)}} \mathbf{1}_{\{\tau_n^{(N)} > 0\}}, \quad t \geq 0. \quad (10.28)$$

In addition, since the coefficients of the modified system of SODEs (10.26) satisfy the standard assumptions, according to Proposition 2 in [80], i.e., Theorem 6.2, the interpolated order γ strong Taylor scheme $(z_n(t), y_n(t))$ applied to (10.26) converges strongly with order γ and pathwise with order $\gamma - \varepsilon$ to its solution (Z_t, Y_t) on the time interval. Thus for every $\varepsilon > 0$, there exists a finite non-negative random variable $\zeta_{\varepsilon, 2T}^{(N)}$ such that

$$\sup_{t \in [0, 2T]} |Z_t(\omega) - z_n(t, \omega)| \leq \zeta_{\varepsilon, 2T}^{(N)}(\omega) \cdot n^{-\gamma+\varepsilon}, \quad \text{a.s.} \quad (10.29)$$

for all $n \in \mathbb{N}$.

Using (10.27), (10.28) and (10.29) to obtain

$$\sup_{t \in [0, \tau_n^{(N)} \wedge \tau_n^{(N)}]} |X_t(\omega) - x_n(t, \omega)| \leq \zeta_{\varepsilon, 2T}^{(N)}(\omega) \cdot n^{-\gamma+\varepsilon} \quad (10.30)$$

for all $n \in \mathbb{N}$, i.e., $x_n(t)$ converges pathwise to X_t with order $\gamma - \varepsilon$ on $[0, \tau_n^{(N)} \wedge \tau^{(N)}]$.

Now notice that (10.29) implies

$$\lim_{n \rightarrow \infty} \sup_{t \in [0, T]} |Z_t(\omega) - z_n(t, \omega)| = 0, \quad \forall \omega \in \Omega,$$

which yields

$$\lim_{N \rightarrow \infty} \tau^{(N)} \geq \inf\{t \geq 0 : |Z_t(\omega)| > N\} \wedge 2T.$$

Also notice that (10.27) implies

$$\tau^{(N)}(\omega) = \inf\{t \geq 0 : |X_t(\omega)| > N\} \wedge 2T \leq \inf\{t \geq 0 : |Z_t(\omega)| > N\} \wedge 2T.$$

Therefore

$$\lim_{n \rightarrow \infty} \tau_n^{(N)} \geq \tau^{(N)}. \quad (10.31)$$

Inequalities (10.30) and (10.31) together give that for any $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\limsup_{n \rightarrow \infty} \sup_{t \in [0, \tau^{(N)}(\omega) - \delta]} n^{\gamma - \varepsilon} |x_n(t, \omega) - X_t(\omega)| \leq \zeta_{\varepsilon, 2T}^{(N)}(\omega)$$

for almost all $\omega \in \Omega$. This implies that $x_n(t, \omega)$ converges pathwise to $X_t(\omega)$ with order $\gamma - \varepsilon$ on every compact interval contained in $[0, \tau^{(N)}]$.

Pick $\delta = T/2$ and define

$$\Omega_N := \{\omega \in \Omega : \tau^{(N)}(\omega) \geq 3T/2\}.$$

Then since $\tau^{(N)} \rightarrow 2T$ almost surely as $N \rightarrow \infty$,

$$\mathbb{P}\left(\bigcup_{N \in \mathbb{N}} \Omega_N\right) = 1.$$

It then follows that

$$\sup_{t \in [0, T]} |X_t(\omega) - x_n(t, \omega)| \leq \zeta_{\varepsilon, T} n^{-\gamma + \varepsilon},$$

which completes the proof. \square

10.6 Endnotes

This chapter is based on Asai and Kloeden [9, 10]. For multi-step schemes for SODEs see [19]. Shardlow and Taylor [120] proved the pathwise convergence of the

Euler–Maruyama and Milstein schemes directly without using strong convergence or standard assumptions.

Chapter 11

Numerical Schemes for RODEs with Affine Noise

Consider a d -dimensional RODE with m -dimensional affine noise of the form

$$\frac{dx}{dt} = f_0(t, x) + \sum_{j=1}^m f_j(t, x) \eta_t^j.$$

This can be written in a compact integral equation form

$$x(t) = x(t_0) + \sum_{j=0}^m \int_{t_0}^t f_j(s, x(s)) \eta_s^j ds \quad (11.1)$$

with a conventional fictitious “noise” component $\eta_t^0 \equiv 1$.

Recall from Chap. 7 that the affine-RODE-Taylor expansion of $F(t, x(t))$ for a hierarchical set \mathcal{H} , where $x(t)$ is a solution of the affine RODE (11.1), is truncated to give the affine-RODE-Taylor approximation for the hierarchical set \mathcal{H} :

$$\Phi_{\mathcal{H}}(t, x(t); F) = \sum_{\alpha \in \mathcal{H}} F_\alpha(t_0, x(t_0)) I_\alpha [1]_{t_0, t}. \quad (11.2)$$

The notations used here are taken from from Chap. 7. In particular, for each multi-index $\alpha = (j_1, \dots, j_l) \in \mathcal{J}_m$ and function $F : [t_0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$, the coefficient function F_α is defined recursively by

$$F_\alpha = \begin{cases} F & : l = 0 \\ L^{j_1} F_{-\alpha} & : l \geq 1 \end{cases},$$

where the partial differential operators L^0, L^1, \dots, L^m are defined by

$$L^0 = \frac{\partial}{\partial t} + \sum_{k=1}^d f_0^k \frac{\partial}{\partial x^k}, \quad L^j = \sum_{k=1}^d f_j^k \frac{\partial}{\partial x^k}, \quad j = 1, \dots, m.$$

Let $\{t_0, t_1, \dots, t_n, \dots, T\}$ be a partition of the time interval $[t_0, T]$ with step sizes $h_n = t_{n+1} - t_n$ and maximal step size $h := \max_n h_n$. Define the multiple integrals

$$I_{\alpha,n}[u(\cdot)] := I_{\alpha}[u(\cdot)]_{t_n, t_{n+1}} = \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_1} \cdots \int_{t_n}^{s_l} u(s_1) \eta_{s_1}^{j_1} \cdots \eta_{s_l}^{j_l} ds_1 \cdots ds_l,$$

for $\alpha = (j_1, \dots, j_l) \in \mathcal{H}$. In particular when $u \equiv 1$ write

$$I_{\alpha,n} := \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_1} \cdots \int_{t_n}^{s_l} \eta_{s_1}^{j_1} \cdots \eta_{s_l}^{j_l} ds_1 \cdots ds_l.$$

In addition, the following new notation will be used:

$$\chi_{\alpha}(\cdot) := L^{\alpha} \chi(\cdot), \quad \chi_{\alpha}^k(\cdot) = L^{\alpha} \chi^k(\cdot) \quad (\text{componentwise})$$

where $L^{\alpha} = L^{j_1} L^{j_2} \cdots L^{j_l}$ and $\chi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is the identity function on \mathbb{R}^d .

11.1 Affine-RODE Taylor Schemes for Bounded Noise

The approximation (11.2) will be applied to the function $F(t, x) \equiv x^k$, the k th component of the identity function $F(x) \equiv \chi(x) = x$ on \mathbb{R}^d for each $k = 1, \dots, d$. The order of these approximations depends on the nature of the driving noise processes in the RODE (11.1). The cases considered here include noise with continuous sample paths as well as piecewise continuous paths such as a compound Poisson process.

For RODEs with affine bounded noise, the multiple integrals $I_{\alpha}[1]_{t_0, t_0+h}$ are then of order $h^{l(\alpha)}$ (see Theorem 11.1 below), so the appropriate hierarchical set for affine-RODE-Taylor schemes is

$$\Lambda_{\gamma} = \{\alpha \in \mathcal{J}_m : l(\alpha) \leq \gamma\}.$$

The general multi-dimensional case with $d, m = 1, 2, \dots$ the *affine-RODE-Taylor scheme of order γ* for $\gamma = 1, 2, 3, \dots$ is defined componentwise by

$$x_{n+1}^k = x_n^k + \sum_{\alpha \in \Lambda_{\gamma} \setminus \{\emptyset\}} \chi_{\alpha}^k(t_n, x_n) I_{\alpha,n}, \quad k = 1, \dots, d. \quad (11.3)$$

With multi-indices $\alpha = (0), (0, 0), (0, 0, 0), \dots$, the multiple integrals are merely deterministic and thus need no special techniques to evaluate. In fact, for $l(\alpha) = l \geq 1$,

$$I_{(0,\dots,0),n} = \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_l} \cdots \int_{t_n}^{s_2} 1 \, ds_1 \cdots ds_l = \frac{1}{l!} h_n^l.$$

Theorem 11.1 Assume that noise sample paths are continuous or essentially bounded on bounded time intervals and that all the appearing derivatives of f_0, f_1, \dots, f_m exist and are continuous. Then the affine-RODE-Taylor scheme (11.3) has pathwise order of convergence γ .

Proof For each $k = 1, \dots, d$, the pathwise local discretisation error of the order γ affine-RODE-Taylor scheme (11.3) for the k th component is given by

$$\mathcal{E}_{n,k}^L := \left| \sum_{\alpha \in \mathcal{R}(A_\gamma)} I_{\alpha,n} [\chi_\alpha^k(\cdot, x(\cdot, \omega))] \right|.$$

Under the standard assumptions, the RODE (11.1) has a unique solution on a finite time interval $[t_0, T]$. Since the sample paths of its solution are continuous, there is a finite $r(\omega, T)$ such that $|x(t, \omega)| < r(\omega, T)$ for all $t \in [t_0, T]$.

Since the sample paths of the stochastic processes η_t^j are assumed to be almost everywhere bounded

$$C(\alpha, \omega, T) := \text{ess sup}_{t \in [t_0, T]} |\eta_t^j(\omega)| < \infty,$$

and thus

$$|I_{\alpha,n}| \leq C(\alpha, \omega, T)^{l(\alpha)-n(\alpha)} h_n^{l(\alpha)},$$

where $n(\alpha)$ is the number of zero elements in α , $l(\alpha)$ is the length of α and $h_n = t_{n+1} - t_n$.

Define

$$C_{\chi_\alpha}^k(\omega, T) := \sup_{\substack{t \in [t_0, T] \\ |x(t, \omega)| < r(\omega, T)}} |\chi_\alpha^k(t, x(t, \omega))| < \infty.$$

Then $\mathcal{E}_{n,k}^L$ can be estimated as

$$\begin{aligned} \mathcal{E}_{n,k}^L &\leq \sum_{|\alpha|=\gamma+1} |I_{\alpha,n}[\chi_\alpha^k(\cdot, x(\cdot))]| \\ &\leq \sum_{|\alpha|=\gamma+1} C_{\chi_\alpha}^k(\omega, T) |I_{\alpha,n}| \\ &\leq \sum_{|\alpha|=\gamma+1} C_{\chi_\alpha}^k(\omega, T) C(\alpha, \omega, T)^{\gamma+1-n(\alpha)} h_n^{\gamma+1} \end{aligned}$$

since all $\alpha \in \mathcal{R}(\Lambda_\gamma)$ have length $l(\alpha) = \gamma + 1$. Obviously $\mathcal{E}_{n,k}^L \sim \mathcal{O}(h_n^{\gamma+1})$, and thus the total pathwise local discretisation error $\mathcal{E}_n^L \sim \mathcal{O}(h_n^{\gamma+1})$. It then follows from Theorem 9.1 that the affine-RODE-Taylor scheme (11.3) has global order of convergence γ . \square

The Affine-RODE Euler Scheme

The Euler approximation is the simplest nontrivial Taylor scheme. It corresponds to the hierarchical set Λ_1 and has the convergence order $\gamma = 1$. It is given componentwise by

$$x_{n+1}^k = x_n^k + f_0^k(t_n, x_n) h_n + \sum_{j=1}^m f_j^k(t_n, x_n) I_{(j),n} \quad (11.4)$$

for $k = 1, \dots, d$, where $h_n = t_{n+1} - t_n = \int_{t_n}^{t_{n+1}} ds$ and

$$I_{(j),n} = \int_{t_n}^{t_{n+1}} \eta_s^j ds, \quad j = 1, \dots, m.$$

The Affine-RODE-Taylor Scheme of Order 2

The *affine-RODE-Taylor scheme of order 2* is given componentwise by

$$\begin{aligned} x_{n+1}^k &= x_n^k + f_0^k(t_n, x_n) h_n + \sum_{j=1}^m f_j^k(t_n, x_n) I_{(j),n} \\ &\quad + \frac{1}{2} L^0 f_0^k(t_n, x_n) h_n^2 + \sum_{\substack{j_1, j_2=0 \\ j_1+j_2 \neq 0}}^m L^{j_1} f_{j_2}^k(t_n, x_n) I_{(j_1, j_2),n} \end{aligned} \quad (11.5)$$

for $k = 1, \dots, d$.

The Affine-RODE-Taylor Scheme of Order 3

The *affine-RODE-Taylor scheme of order 3* is given componentwise by

$$\begin{aligned} x_{n+1}^k &= x_n^k + f_0^k(t_n, x_n) h_n + \sum_{j=1}^m f_j^k(t_n, x_n) I_{(j),n} \\ &\quad + \frac{1}{2} L^0 f_0^k(t_n, x_n) h_n^2 + \sum_{\substack{j_1, j_2=0 \\ j_1+j_2 \neq 0}}^m L^{j_1} f_{j_2}^k(t_n, x_n) I_{(j_1, j_2),n} \\ &\quad + \frac{1}{6} L^0 L^0 f_0^k(t_n, x_n) h_n^3 + \sum_{\substack{j_1, j_2, j_3=0 \\ j_1+j_2+j_3 \neq 0}}^m L^{j_1} L^{j_2} f_{j_3}^k(t_n, x_n) I_{(j_1, j_2, j_3),n} \end{aligned} \quad (11.6)$$

for $k = 1, \dots, d$.

11.2 Affine-RODEs with Special Structure

The Taylor schemes (11.3) can be simplified considerably when the drift and noise coefficients f_0, f_1, \dots, f_m of the affine-RODE (11.1) satisfy special structural properties such as additive or commutative noise.

11.2.1 Additive Noise

A RODE with affine noise is said to have *additive noise* if the noise coefficients f_1, \dots, f_m are all constant or depend only on t . Then all of the spatial derivatives of these noise coefficients vanish and, hence, so do the corresponding higher order terms in their affine-RODE-Taylor schemes.

The affine-RODE-Taylor schemes thus contain only the multi-indices in the hierarchical set

$$\Lambda_\gamma^0 = \{\alpha \in \mathcal{J}_m : \mathbf{l}(\alpha) \leq \gamma, j_l = 0 \text{ if } \mathbf{l}(\alpha) = l \geq 1\}.$$

For example, the second order affine-RODE-Taylor scheme (11.5) then reduces to

$$x_{n+1}^k = x_n^k + f_0^k(t_n, x_n) h_n + \sum_{j=1}^m f_j^k(t_n, x_n) I_{(j),n} + \sum_{j=0}^m L^j f_0^k(t_n, x_n) I_{(j,0),n}$$

for $k = 1, \dots, d$.

11.2.2 Commutative Noise

Another major simplification occurs under *commutative noise*, i.e., when the drift and noise coefficients satisfy

$$L^i f_j^k(t, x) \equiv L^j f_i^k(t, x) \quad \text{for all } i, j = 0, 1, \dots, m.$$

Then, by the generalized integration-by-parts identities,

$$I_{(i,j),n} + I_{(j,i),n} = I_{(i),n} I_{(j),n}, \quad i, j = 0, 1, \dots, m,$$

the sum of terms

$$L^i f_j^k(t_n, x_n) I_{(i,j),n} + L^j f_i^k(t_n, x_n) I_{(j,i),n}$$

simplifies to

$$L^i f_j^k(t_n, x_n) I_{(i),n} I_{(j),n},$$

which involves more easily computed multiple integrals of lower multiplicity. For example, for a RODE with affine commutative noise the k th component of the affine-RODE Taylor scheme (11.5) simplifies to

$$\begin{aligned} x_{n+1}^k &= x_n^k + f_0^k(t_n, x_n) h_n + \sum_{j=1}^m f_j^k(t_n, x_n) I_{(j),n} \\ &\quad + \frac{1}{2} L^0 f_0^k(t_n, x_n) h_n^2 + \sum_{0 \leq i < j \leq m} L^i f_j^k(t_n, x_n) I_{(i),n} I_{(j),n} \end{aligned}$$

for $k = 1, \dots, d$.

11.3 Affine-RODE Derivative-Free Schemes

A disadvantage of Taylor schemes is that the derivatives of various orders of the drift and noise coefficients must be first derived and then evaluated at each step. In the past this made the implementation of such schemes a complicated undertaking, but this is no longer such a difficulty these days with symbolic manipulators. Nevertheless it is useful to have approximations and schemes that avoid the use of derivatives of the drift and noise coefficients in much the same way that Runge–Kutta schemes do in the more traditional setting since these often have other computational advantages.

11.3.1 Finite Difference Approximation of Derivatives

Derivative-free schemes can be derived by replacing the derivatives in RODE-Taylor schemes by finite difference quotients of an appropriate order. Since the Euler or affine-RODE-Taylor scheme of order 1 contains no derivatives of f_0, f_1, \dots, f_m , the second order Taylor scheme (11.5) in the scalar “autonomous”¹ case with a single noise integral, i.e., with $d = m = 1$, will be considered. The underlying RODE reads

$$\frac{dx}{dt} = f_0(x) + f_1(x) \eta_t$$

¹Strictly speaking a RODE is always nonautonomous since the noise depends explicitly on time. The word autonomous is used here to indicate that the drift and noise coefficients f_0, f_1, \dots, f_m do not depend on the time variable.

and the affine-RODE-Taylor scheme of order 2 is given by

$$\begin{aligned}x_{n+1} = & x_n + f_0(x_n) h_n + f_1(x_n) I_{(1),n} + \frac{1}{2} L^0 f_0(x_n) h_n^2 + L^0 f_1(x_n) I_{(0,1),n} \\& + L^1 f_0(x_n) I_{(1,0),n} + L^1 f_1(x_n) I_{(1,1),n},\end{aligned}$$

or equivalently

$$\begin{aligned}x_{n+1} = & x_n + f_0(x_n) h_n + f_1(x_n) I_{(1),n} + \frac{1}{2} f_0(x_n) f_0'(x_n) h_n^2 \\& + f_0(x_n) f_1'(x_n) I_{(0,1),n} + f_1(x_n) f_0'(x_n) I_{(1,0),n} + f_1(x_n) f_1'(x_n) I_{(1,1),n},\end{aligned}\quad (11.7)$$

where $'$ denotes the differentiation with respect to x .

By the classical Taylor expansions for the coefficient functions

$$f_j(x) f_i'(x) = \frac{1}{h} (f_i(x + f_j(x) h) - f_i(x)) + \mathcal{O}(h), \quad i, j = 0, 1,$$

the derivative terms in the affine-RODE-Taylor scheme (11.7) are given by

$$\begin{aligned}& f_j(x_n) f_i'(x_n) I_{(i,j),n} \\= & \left(\frac{1}{h_n} (f_i(x_n + f_j(x_n) h_n) - f_i(x_n)) + \mathcal{O}(h_n) \right) I_{(i,j),n} \\= & \frac{1}{h_n} (f_i(x_n + f_j(x_n) h_n) - f_i(x_n)) I_{(i,j),n} + \mathcal{O}(h_n^3)\end{aligned}\quad (11.8)$$

for $i, j = 0, 1$, since $\mathcal{O}(h_n) I_{(i,j),n} = \mathcal{O}(h_n^3)$.

The remainder in (11.8) is of the same order as the local discretisation error, so the term on the left can be replaced by that on the right without reducing the global order of the resulting scheme. This gives the *second order derivative-free scheme* for the scalar “autonomous” case with a single noise ($d = m = 1$):

$$\begin{aligned}x_{n+1} = & x_n + \frac{1}{2} f_0(x_n) h_n + f_1(x_n) I_{(1),n} + \frac{1}{2} f_0(x_n + f_0(x_n) h_n) h_n \\& + \frac{1}{h_n} \sum_{\substack{i,j=0 \\ i+j \neq 0}}^1 (f_i(x_n + f_j(x_n) h_n) - f_i(x_n)) I_{(i,j),n}.\end{aligned}$$

The vector version of the second order derivative-free scheme for “autonomous” RODE with affine noise has the k th component given by²

²When the context is clear the notation x is used for both scalar and vector.

$$\begin{aligned} x_{n+1}^k &= x_n^k + \frac{1}{2} f_0^k(x_n) h_n + \sum_{j=1}^m f_j^k(x_n) I_{(j),n} + \frac{1}{2} f_0^k(x_n + f_0(x_n) h_n) h_n \\ &\quad + \frac{1}{h_n} \sum_{\substack{i,j=0 \\ i+j \neq 0}}^m (f_i^k(x_n + f_j(x_n) h_n) - f_i^k(x_n)) I_{(i,j),n} \end{aligned}$$

for $k = 1, \dots, d$. In the usual ODE case, i.e., with $f_j(x) \equiv 0$ for $j = 1, \dots, m$, this is just the classical second order Runge–Kutta scheme for ODEs that is known as the Heun scheme. This principle can be extended to obtain higher order derivative-free schemes. See [91] for analogous higher order derivative-free schemes for Itô stochastic differential equations.

11.3.2 Runge–Kutta Schemes for Affine-RODE

Heuristic adaptations of the traditional Runge–Kutta schemes for ODEs to RODEs with affine noise will usually not attain their asserted order due to lack of sufficient smoothness in the time variable. Consequently, new classes need to be derived for RODEs. The affine noise structure of affine-RODEs facilitates this task. This subsection follows Kloeden and Rössler [93], to which the reader is referred for details and proofs.

Consider an “autonomous” affine-RODE of the form

$$\frac{dx}{dt} = f_0(x) + \sum_{j=1}^m f_j(x) \eta_t^j \quad (11.9)$$

with $f_j : \mathbb{R}^d \rightarrow \mathbb{R}^d$ at least three times continuously differentiable. Explicit and implicit Runge–Kutta schemes of up to order 3 for the affine-RODE (11.9) have the form

$$\begin{aligned} x_{n+1} &= x_n + \sum_{i=1}^s b_i^{(0)} f_0(\xi_i) h_n + \sum_{i=1}^s \sum_{j_1=1}^m b_i^{(j_1)} f_{j_1}(x_n) I_{(j_1),n} \\ &\quad + \frac{1}{h_n} \sum_{i=1}^s \sum_{\substack{j_1, j_2=0 \\ j_1+j_2 \neq 0}}^m b_i^{(j_1, j_2)} f_{j_2}(\xi_i^{(j_1)}) I_{(j_1, j_2),n} \\ &\quad + \frac{1}{h_n^2} \sum_{i=1}^s \sum_{\substack{j_1, j_2, j_3=0 \\ j_1+j_2+j_3 \neq 0}}^m b_i^{(j_1, j_2, j_3)} f_{j_3}(\xi_i^{(j_1, j_2)}) I_{(j_1, j_2, j_3),n} \end{aligned} \quad (11.10)$$

for $n = 0, 1, \dots, N - 1$ with stage functions

Table 11.1 Tableau of coefficients for the Runge–Kutta method (11.10)

A	$A^{(j_1)(j_1)}$	$A^{(j_1, j_2)(j_2)}$	$A^{(j_1, j_2)(j_1)}$
$\mathbf{b}^{(0)\top}$	$\mathbf{b}^{(j_1)\top}$	$\mathbf{b}^{(j_1, j_2)\top}$	$\mathbf{b}^{(j_1, j_2, j_3)\top}$

$$\begin{aligned}\xi_i &= x_n + \sum_{j=1}^s A_{ij} f_0(\xi_j) h_n, \\ \xi_i^{(j_1)} &= x_n + \sum_{j=1}^s A_{ij}^{(j_1)(j_1)} f_{j_1}(x_n) h_n, \\ \xi_i^{(j_1, j_2)} &= x_n + \sum_{j=1}^s A_{ij}^{(j_1, j_2)(j_1)} f_{j_1}(\xi_j^{(j_2)}) h_n + \sum_{j=1}^s A_{ij}^{(j_1, j_2)(j_2)} f_{j_2}(\xi_j^{(j_1)}) h_n,\end{aligned}$$

where the integer s denotes the number of stages.

For $\alpha \in \mathcal{J}_m$, write $\mathbf{b}^\alpha = (b_1^\alpha, \dots, b_s^\alpha)^\top$. Each Runge–Kutta scheme is characterised by its coefficients and weights, which can be represented by a Butcher tableau as presented in Table 11.1 for an s -stage method. Coefficients which do not appear in the tableau are equal to zero.

Theorem 11.2 Denote $\mathbf{e} := (1, 1, \dots, 1)^\top$. The Runge–Kutta method (11.10) has the order of convergence 1 if the coefficients of (11.10) satisfy the conditions

$$\begin{array}{ll} (1) & \mathbf{b}^{(0)\top} \mathbf{e} = 1 \\ (3) & \mathbf{b}^{(j_1, j_2)\top} \mathbf{e} = 0 \end{array}$$

$$\begin{array}{ll} (2) & \mathbf{b}^{(j_1)\top} \mathbf{e} = 1 \\ (4) & \mathbf{b}^{(j_1, j_2, j_3)\top} \mathbf{e} = 0; \end{array}$$

has the order of convergence 2 if in addition the conditions

$$(5) \quad \mathbf{b}^{(0)\top} A \mathbf{e} = \frac{1}{2}$$

$$(6) \quad \mathbf{b}^{(j_1, j_2)\top} A^{(j_1)(j_1)} \mathbf{e} = 1$$

$$(7) \quad \mathbf{b}^{(j_1, j_2, j_3)\top} A^{(j_1, j_2)(j_2)} \mathbf{e} = 0$$

$$(8) \quad \mathbf{b}^{(j_1, j_2, j_3)\top} A^{(j_1, j_2)(j_1)} \mathbf{e} = 0$$

are satisfied; and has order of convergence 3 if in addition the conditions

$$(9) \quad \mathbf{b}^{(0)\top} (A(A\mathbf{e})) = \frac{1}{6}$$

$$(10) \quad \mathbf{b}^{(0)\top} ((A\mathbf{e})(A\mathbf{e})) = \frac{1}{3}$$

$$(11) \quad \mathbf{b}^{(j_1, j_2, j_3)\top} (A^{(j_1, j_2)(j_2)} (A^{(j_1)(j_1)} \mathbf{e})) = 1$$

$$(12) \quad \mathbf{b}^{(j_1, j_2, j_3)\top} (A^{(j_1, j_2)(j_1)} (A^{(j_2)(j_2)} \mathbf{e})) = 0$$

$$(13) \quad \mathbf{b}^{(j_1, j_2)\top} ((A^{(j_1)(j_1)} \mathbf{e})(A^{(j_1)(j_1)} \mathbf{e})) = 0$$

$$(14) \quad \mathbf{b}^{(j_1, j_2, j_3)\top} ((A^{(j_1, j_2)(j_2)} \mathbf{e})(A^{(j_1, j_2)(j_2)} \mathbf{e})) = 0$$

$$(15) \quad \mathbf{b}^{(j_1, j_2, j_3)\top} ((A^{(j_1, j_2)(j_2)} \mathbf{e})(A^{(j_1, j_2)(j_1)} \mathbf{e})) = 1$$

$$(16) \quad \mathbf{b}^{(j_1, j_2, j_3)\top} ((A^{(j_1, j_2)(j_1)} \mathbf{e})(A^{(j_1, j_2)(j_1)} \mathbf{e})) = 0$$

are satisfied.

Denote by γ_{ODE} and γ_{RODE} the order of convergence of the Runge–Kutta scheme applied to the ODE with vector field f_0 only and to the affine-RODE (11.9), respectively. Since $\gamma_{\text{ODE}} \geq \gamma_{\text{RODE}}$, the scheme converges at least with order $\gamma = \gamma_{\text{RODE}}$.

Example 11.1 The Euler scheme (11.4) is an example of an $s = 1$ stage explicit Runge–Kutta scheme of order 1. Its coefficients $b_1^{(0)} = b_1^{(j_1)} = 1$ and $b_1^{(j_1, j_2)} = b_1^{(j_1, j_2, j_3)} = 0$ satisfy conditions (1)–(4) of Theorem 11.2.

Example 11.2 The coefficients for a Runge–Kutta method of order 2 are given in Table 11.2. They satisfy conditions (1)–(8) of Theorem 11.2.

Example 11.3 Some typical coefficients defining explicit order 3 Runge–Kutta schemes with $s = 4$ stages are presented in Tables 11.3 and 11.4.

Since there are more degrees of freedom in choosing these coefficients, it is possible to derive a Runge–Kutta scheme converging with order 4 if it is applied to an ODE (see Table 11.4). Therefore, if the weights $\mathbf{b}_i^{(0)}$ and the coefficients A_{ij} are chosen such that conditions (1), (5), (9) and (10) of Theorem 11.2 and, additionally, the conditions

$$\begin{aligned} \mathbf{b}^{(0)\top} (A((A\mathbf{e})(A\mathbf{e}))) &= \frac{1}{12}, & \mathbf{b}^{(0)\top} ((A\mathbf{e})(A(A\mathbf{e}))) &= \frac{1}{8}, \\ \mathbf{b}^{(0)\top} (A(A(A\mathbf{e}))) &= \frac{1}{24}, & \mathbf{b}^{(0)\top} ((A\mathbf{e})(A\mathbf{e})(A\mathbf{e})) &= \frac{1}{4}, \end{aligned}$$

are satisfied, then the Runge–Kutta scheme is of order 4 when $f_j \equiv 0$ for $j = 1, \dots, k$ in (11.9), see [61]. Since the weights $\mathbf{b}_i^{(0)}$ and the coefficients A_{ij} are not linked to the other weights and coefficients by the order conditions, one can substitute the coefficients for $\mathbf{b}^{(0)}$ and A in the Runge–Kutta schemes presented by the coefficients of any well known Runge–Kutta method having at least order γ .

Table 11.2 The explicit Runge–Kutta scheme with $\gamma_{\text{ODE}} = 2$ and $\gamma_{\text{RODE}} = 2$

	1	1	0	0
(2, 2)	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$

Table 11.3 The explicit Runge–Kutta scheme with $\gamma_{\text{ODE}} = 3$ and $\gamma_{\text{RODE}} = 3$

$\frac{1}{3}$	2	-1	0
0 $\frac{2}{3}$	0 1	0 -1	1 1
0 0 0	0 0 1	0 0 0	1 1 0
(3, 3)	$\frac{1}{4}$ 0 $\frac{3}{4}$ 0	1 0 0 0	$-\frac{3}{2}$ $-\frac{1}{2}$ 2 0
			$-\frac{1}{2}$ $\frac{1}{2}$ $-\frac{1}{2}$ $\frac{1}{2}$

Table 11.4 The explicit Runge–Kutta scheme with $\gamma_{\text{ODE}} = 4$ and $\gamma_{\text{RODE}} = 3$

	$\frac{1}{3}$	$-\frac{1}{3}$	1	2	0	-1	0	0	1	1	1	0
				0 1		0 -1						
				0 0 1		0 0 0						
(4, 3)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	1 0 0 0	$-\frac{3}{2}$	$-\frac{1}{2}$	2 0	$-\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$

11.4 Linear Multi-step Methods for Affine RODEs

Linear multi-step methods (LMMs) were derived for RODEs driven by an Itô diffusion, i.e., solution of an Itô SDE, in Chap. 10 using Itô stochastic Taylor expansions. They can also be applied to RODEs with a special affine structure. Since the driving noise does not need to be an Itô diffusion, a similar approach, but using affine-RODE-Taylor approximations, will be introduced here. The resulting multi-step methods are then, in particular, directly applicable to the affine-RODE (11.1).

Euler-type s -step LMMs for a RODE (11.1) with m -dimensional affine noise have the general componentwise form

$$\sum_{i=0}^s a_i x_{n-i}^k = \sum_{i=0}^s b_i f_0^k(t_{n-i}, x_{n-i}) h_n + \sum_{i=1}^s c_i \sum_{j=1}^m f_j^k(t_{n-i}, x_{n-i}) I_{(j), n-i}, \quad (11.11)$$

for $k = 1, \dots, d$. The coefficients in (11.11) need to satisfy certain consistency conditions for the LMMs (11.11) to have order 1. It will be shown next how these can be derived, for simplicity in the case $s = 2$.

The pathwise local discretisation error \mathcal{E}_n^L of the LMMs (11.11) is of the same order as the pathwise local discretisation error $\mathcal{E}_{n,k}^L$ for the k th component, which is given by

$$\begin{aligned} \mathcal{E}_{n,k}^L := & \left| \sum_{i=0}^2 a_i x^k(t_{n-i}) - \sum_{i=0}^2 b_i f_0^k(t_{n-i}, x(t_{n-i})) h_n \right. \\ & \left. - \sum_{i=1}^2 c_i \sum_{j=1}^m f_j^k(t_{n-i}, x(t_{n-i})) I_{(j), n-i} \right|. \end{aligned} \quad (11.12)$$

The affine-RODE-Taylor expansions

$$\begin{aligned} x^k(t_n) &= x^k(t_{n-1}) + f_0^k(t_{n-1}, x(t_{n-1})) I_{(0), n-1} \\ &\quad + \sum_{j=1}^m f_j^k(t_{n-1}, x(t_{n-1})) I_{(j), n-1} + \sum_{j_1, j_2=0}^m I_{(j_1, j_2), n-1} [L^{j_1} f_{j_2}^k(\cdot, x(\cdot))] \\ x^k(t_{n-1}) &= x^k(t_{n-2}) + f_0^k(t_{n-2}, x(t_{n-2})) I_{(0), n-2} \end{aligned}$$

$$+ \sum_{j=1}^m f_j^k(t_{n-2}, x(t_{n-2})) I_{(j),n-2} + \sum_{j_1, j_2=0}^m I_{(j_1, j_2), n-2} [L^{j_1} f_{j_2}^k(\cdot, x(\cdot))]$$

and classical chain rule expansions

$$\begin{aligned} f_0^k(t_n, x(t_n)) &= f_0^k(t_{n-1}, x(t_{n-1})) + \sum_{j=0}^m I_{(j),n-1} [L^j f_0^k(\cdot, x(\cdot))] \\ f_0^k(t_{n-1}, x(t_{n-1})) &= f_0^k(t_{n-2}, x(t_{n-2})) + \sum_{j=0}^m I_{(j),n-2} [L^j f_0^k(\cdot, x(\cdot))] \\ f_j^k(t_{n-1}, x(t_{n-1})) &= f_j^k(t_{n-2}, x(t_{n-2})) + \sum_{j_1=0}^m I_{(j_1),n-2} [L^{j_1} f_j^k(\cdot, x(\cdot))] \end{aligned}$$

are substituted into the local discretisation error (11.12) to yield

$$\begin{aligned} \mathcal{E}_{n,k}^L &= \left| (a_0 + a_1 + a_2)x^k(t_{n-2}) + (2a_0 + a_1 - b_0 - b_1 - b_2)f_0^k(t_{n-2}, x(t_{n-2}))h_n \right. \\ &\quad + \left. \sum_{j=1}^m \left(a_0 I_{(j),n-1} + (a_0 + a_1)I_{(j),n-2} - c_1 I_{(j),n-1} - c_2 I_{(j),n-2} \right) \right. \\ &\quad \left. \cdot f_j^k(t_{n-2}, x(t_{n-2})) + R_1 \right|, \end{aligned}$$

where R_1 is a remainder term given by

$$\begin{aligned} R_1 &= a_0 \sum_{j_1, j_2=0}^m \left(I_{(j_1, j_2), n-2} [L^{j_1} f_{j_2}^k(\cdot, x(\cdot))] + I_{(j_1), n-2} [L^{j_1} f_{j_2}^k(\cdot, x(\cdot))] I_{(j_2), n-1} \right. \\ &\quad \left. + I_{(j_1, j_2), n-1} [L^{j_1} f_{j_2}^k(\cdot, x(\cdot))] \right) + a_1 \sum_{j_1, j_2=0}^m I_{(j_1, j_2), n-2} [L^{j_1} f_{j_2}^k(\cdot, x(\cdot))] \\ &\quad - b_0 \sum_{j=0}^m \left(I_{(j), n-2} [L^j f_0^k(\cdot, x(\cdot))] + I_{(j), n-1} [L^j f_0^k(\cdot, x(\cdot))] \right) h_n \\ &\quad - b_1 \sum_{j=0}^m I_{(j), n-2} [L^j f_0^k(\cdot, x(\cdot))] h_n - c_1 \sum_{j_1=0}^m \sum_{j_2=1}^m I_{(j_1), n-2} [L^{j_1} f_{j_2}^k(\cdot, x(\cdot))] I_{(j_2), n-1}. \end{aligned}$$

Clearly R_1 is of order h_n^2 . When the coefficients of the LMMs (11.11) satisfy the following consistency conditions

$$\begin{cases} a_0 + a_1 + a_2 = 0, & b_0 + b_1 + b_2 = 2a_0 + a_1 \\ c_1 = a_0, & c_2 = a_0 + a_1, \end{cases}$$

the LMMs (11.11) have order 1 convergence.

In order to construct higher order schemes, the integrands in the remainder term R_1 are expanded similarly. Its terms with constant coefficients are now included in the numerical scheme and a new remainder term R_2 with threefold noise integrals is obtained. The corresponding LMMs, which now contain terms with the derivatives $L^0 f_0^k$, $L^j f_0^k$, $L^0 f_j^k$ and $L^{j_1} f_{j_2}^k$, have the following componentwise form:

$$\begin{aligned} \sum_{i=0}^2 a_i x_{n-i}^k &= \sum_{i=0}^2 b_i f_0^k(t_{n-i}, x_{n-i}) h_n + \sum_{i=1}^2 c_i \sum_{j=1}^m f_j^k(t_{n-i}, x_{n-i}) I_{(j), n-i} \\ &+ \sum_{i=1}^2 d_i \sum_{j=0}^m L^j f_0^k(t_{n-i}, x_{n-i}) I_{(j), n-i} h_n \\ &+ \sum_{i=1}^2 e_i \sum_{j_1, j_2=0}^m L^{j_1} f_{j_2}^k(t_{n-i}, x_{n-i}) I_{(j_1, j_2), n-i}, \end{aligned}$$

for $k = 1, \dots, d$. They are of order 2 if in addition the following consistency conditions are satisfied:

$$\begin{cases} d_1 = -b_0, & d_2 = a_0 - b_0 - b_1 \\ e_1 = a_0, & e_2 = a_0 + a_1. \end{cases}$$

Higher order LMMs as well as LMMs with more steps for the affine-RODE (11.1) can be generated in the same manner. In general, an s -step LMM for an affine-RODE (11.1) with order γ has the general form

$$\begin{aligned} \sum_{i=0}^s C_{\emptyset, i} x_{n-i}^k &= \sum_{i=0}^s C_{(0), i} f_0^k(t_{n-i}, x_{n-i}) h_n + \sum_{i=1}^s \sum_{j=1}^m C_{(j), i} f_j^k(t_{n-i}, x_{n-i}) I_{(j), n-i} \\ &+ \sum_{i=1}^s \sum_{\alpha \in \Lambda_\gamma^0 \setminus \{(0)\}} C_{\alpha, i}^* L^{\alpha-} f_0^k(t_{n-i}, x_{n-i}) I_{\alpha-, n-i} h_n \\ &+ \sum_{i=1}^s \sum_{\alpha \in \Lambda_\gamma \setminus \{\emptyset, (0), (j)\}} C_{\alpha, i} \chi_\alpha^k(t_{n-i}, x_{n-i}) I_{\alpha, n-i}, \end{aligned}$$

where Λ_γ^0 is given by

$$\Lambda_\gamma^0 = \{\alpha \in \Lambda_\gamma : \text{the last component } j_l = 0\}$$

and the coefficients $C_{\alpha, i}$ and $C_{\alpha, i}^*$ satisfy consistency conditions up to γ -order.

11.5 Endnotes

This chapter is based mostly on Asai and Kloeden [11]. See also the endnotes for Chap. 7. The material on Runge–Kutta schemes for affine-RODEs is from Kloeden and Rössler [93], which considers corresponding schemes for affine control systems.

The derivation of Runge–Kutta schemes for affine-RODEs is based on a different kind of Taylor-like expansions, now called B-series, that was introduced by John Butcher [22] to investigate Runge–Kutta schemes for ODEs and then extended to SODEs by Burrage and Burrage [21] and Rössler [117]. Debrabant and Kværnø [37] show that they are essentially equivalent to the approach using stochastic Taylor expansions in Kloeden and Platen [91]. Roughly speaking, it puts the emphasis on individual partial derivatives multiplied by terms involving sums of stochastic integrals instead of on individual stochastic integrals multiplied by terms involving sums of partial derivatives.

Chapter 12

RODE-Taylor Schemes: General Case

Random ordinary differential equations are ordinary differential equations that contain a stochastic process in their vector field functions. They are thus nonautonomous ODEs for each sample path and can be analysed pathwise with deterministic calculus. However, the resulting vector field after the insertion of the driving stochastic process is at most Hölder continuous in time, no matter how smooth the vector field is in its original variables. Thus the solutions to RODEs do not have sufficient smoothness to have Taylor expansions in the conventional sense. Classical numerical schemes for ODEs can be used pathwise for these RODEs, but they will rarely attain their traditional order.

In this chapter the RODE-Taylor approximations for the solutions of RODEs that were introduced in Chap. 8 will be used to derive systematically higher order numerical schemes for these RODEs.

12.1 RODE-Taylor Schemes

Recall that the integral equation expansions of solutions to the RODE

$$\frac{dx}{dt} = g(x, \eta_t(\omega)), \quad x(t_0) = x_0 \quad (12.1)$$

can be simplified further to (8.4), under Assumption 8.1.1 that the driving stochastic process $(\eta_t)_{t \geq 0}$ is Hölder continuous. Recall that driving process is inserted into the second component of the vector field $g(x, \mathbf{w})$, which is smooth in both variables.

Since $L^i \chi : \mathbb{R}^d \times \mathbb{R}^{m \times i} \rightarrow \mathbb{R}^d$ is in $\mathcal{V}_i = \mathcal{C}^\infty(\mathbb{R}^d \times \mathbb{R}^{m \times i}, \mathbb{R}^d)$, we can approximate it by a Taylor expansion in its last $m \times i$ variables. These will be used in the integral equation expansion (8.4) to construct a temporally discretised approximation of the solution of the RODE (12.1). The resulting numerical schemes will be called *RODE-Taylor schemes*.

The same notation introduced in Sect. 8.1 will be adopted here and recalled briefly for the reader's convenience.

For any matrix multi-index $\mathbf{a} \in \mathcal{M}_m = \bigcup_{i=1}^{\infty} \mathbb{N}_0^{m \times i}$ with $\iota(\mathbf{a}) = i$, let

$$F_{\mathbf{a}} = \frac{1}{\mathbf{a}!} (\partial^{\mathbf{a}} L^{\iota(\mathbf{a})} \chi) \in \mathcal{V}_{\iota(\mathbf{a})}.$$

For any vector $v = (v_1, \dots, v_d) \in \mathbb{R}^d$ with $d \in \mathbb{N}$, denote by $v^{\times n}$ ($n \in \mathbb{N}$) the $d \times n$ matrix

$$v^{\times n} := \underbrace{(v, \dots, v)}_{n \text{ times}} = \begin{pmatrix} v_1 & \dots & v_1 \\ \vdots & & \vdots \\ v_d & \dots & v_d \end{pmatrix}.$$

The following specific subsets of matrix multi-indices

$$\mathcal{M}_m^\gamma := \left\{ \mathbf{a} \in \mathcal{M}_m \mid \iota(\mathbf{a}) + \|\mathbf{a}^\top \Theta\|_1 < \gamma + 1 \right\}$$

will be used, where $\gamma > 0$ and $\Theta = (\theta_1, \dots, \theta_m) \in (0, 1]^m$.

For $\gamma > 0$ define the function $\Phi_\gamma : \mathbb{R}^d \times [0, T] \times [0, 1] \times \Omega \rightarrow \mathbb{R}^d$ by

$$\Phi_\gamma(x, t, h, \omega) := \sum_{\mathbf{a} \in \mathcal{M}_m^\gamma} F_{\mathbf{a}}(x, \eta_t(\omega)^{\times \iota(\mathbf{a})}) \cdot I_{\mathbf{a}, t, t+h}(\omega), \quad (12.2)$$

where for $i = \iota(\mathbf{a}) \geq 1$,

$$I_{\mathbf{a}, t, t+h} = \int_t^{t+h} \int_t^{s_1} \cdots \int_t^{s_{i-1}} ((\Delta \eta_{t, s_1})^{\mathbf{a}_1} \cdots (\Delta \eta_{t, s_i})^{\mathbf{a}_i}) \, ds_i \cdots ds_2 ds_1. \quad (12.3)$$

Consider an equidistant partition $(t_n)_{n=0,1,\dots,N_h}$ of $[0, T]$ with constant stepsize $h \in (0, 1]$:

$$N_h := \left\lceil \frac{T}{h} \right\rceil, \quad t_n := nh, \quad \text{for } n = 0, 1, \dots, N_h - 1 \text{ and } t_{N_h} := T,$$

where $\lceil k \rceil$ for a real number $k \geq 0$ is the smallest integer $\tilde{k} \in \mathbb{N}_0$ such that $\tilde{k} \geq k$.

The pathwise γ -RODE-Taylor is defined by

$$x_{n+1}^{(h)}(\omega) = \Phi_\gamma(x_n^{(h)}(\omega), t_n, \Delta_{t_n, t_{n+1}}, \omega), \quad n = 0, 1, \dots, N_h - 1, \quad (12.4)$$

with the initial value $x_0^{(h)}(\omega) = x_0(\omega)$ for each $\omega \in \Omega$ and $\Delta_{t_n, t_{n+1}} = t_{n+1} - t_n$. It generates a discrete time stochastic process $(X_n^{(h)})_{n=0,1,\dots,N_h}$.

To facilitate error estimates, the scheme (12.4) will extend to a continuous time stochastic process $(\bar{X}_t^{(h)})_{t \in [0, T]}$ defined by $\bar{X}_{t_n}^{(h)} = x_n^{(h)}$ for $n = 0, 1, \dots, N_h$ and

$$\bar{X}_t^{(h)}(\omega) = \Phi_\gamma(\bar{X}_{t_n}^{(h)}, t_n, \Delta_{t_n, t}, \omega) \quad \text{for } t \in (t_n, t_{n+1}), n = 0, 1, \dots, N_h - 1.$$

Clearly the process $(\bar{X}_t^{(h)})_{t \in [0, T]}$ has continuous sample paths and it is also non-anticipative if the stochastic process η_t is non-anticipative.

Theorem 12.1 (Global discretisation error of the RODE-Taylor schemes) *Let $\gamma > 0$ and let $\varepsilon = (\varepsilon_1, \dots, \varepsilon_m) \in [0, 1]^m$ with $\|\varepsilon\|_1 = \sum_{i=1}^m \varepsilon_i < \gamma$. Then,*

$$\sup_{0 \leq t \leq T} \|x(t, \omega) - \bar{X}_t^{(h)}(\omega)\|_2 \leq K_\gamma^\varepsilon(\omega) \cdot h^{(\gamma - \|\varepsilon\|_1)}$$

for all $0 < h \leq h_0(\omega)$ and all $\omega \in \Omega$ with the nonnegative random variables

$$K_\gamma^\varepsilon = (C_\gamma^\varepsilon T) e^{\kappa_\gamma T}, \quad h_0 = (K_\gamma^\varepsilon)^{1/(\|\varepsilon\|_1 - \gamma)} \wedge 1,$$

where $\beta = \Theta - \lceil \frac{\varepsilon}{\vartheta} \rceil$ and

$$C_\gamma^\varepsilon := \sum_{\substack{\alpha \in \mathcal{M}_m \setminus \mathcal{M}_m^\gamma \\ \iota(\alpha) \leq \lceil \gamma + 1 \rceil \\ |\alpha| \leq \lceil \frac{\varepsilon}{\vartheta} \rceil}} \frac{(\|\eta\|_\beta)^{|\alpha|}}{\iota(\alpha)!} \sup_{\substack{\|w_1\|_2, \dots, \|w_{\iota(\alpha)}\|_2 \leq \|\eta\|_\infty \\ \|y\|_2 \leq |x|_\infty}} \|F_\alpha(y, w_1, \dots, w_{\iota(\alpha)})\|_2,$$

$$\kappa_\gamma := \sum_{\alpha \in \mathcal{M}_m^\gamma \setminus \{\emptyset\}} (2\|\eta\|_\infty)^{|\alpha|} \left(\sup_{0 \leq t \leq T} \sup_{\|y - x(t, \omega)\| \leq 1} \|\partial_x F_\alpha(y, \eta_t^{\times \iota(\alpha)})\|_2 \right).$$

Proof The global discretisation order is obtained from that of the local discretisation error given by the truncation error of the corresponding Taylor approximation in Theorem 8.1 and the classical theorem for ODEs. It needs the Lipschitz estimate for the function Φ_γ presented in Lemma 12.1 below. \square

Lemma 12.1 *Let $[t_0, t_0 + h] \subset [0, T]$ with $0 \leq h \leq 1$ and let $\gamma > 0$. Then, for each $\omega \in \Omega$ there exists $\kappa_\gamma > 0$ such that*

$$\|(\Phi_\gamma(x_1, t_0, h, \omega) - x_1) - (\Phi_\gamma(x_2, t_0, h, \omega) - x_2)\|_2 \leq h \kappa_\gamma(\omega) \|x_1 - x_2\|_2$$

for every $x_1, x_2 \in \mathbb{R}^d$ with $\|x(t_0, \omega) - x_1\|_2 \leq 1$ and $\|x(t_0, \omega) - x_2\|_2 \leq 1$.

Proof First note that

$$\Phi_\gamma(x_j, t_0, h, \omega) - x_j = \sum_{\alpha \in \mathcal{M}_m^\gamma \setminus \{\emptyset\}} F_\alpha(x_j, \eta_{t_0}^{\times \iota(\alpha)}) \cdot I_{\alpha, t_0, t_0+h}$$

for $j = 1$ and 2 . Moreover, for $\alpha \in \mathcal{M}_m$ with $\iota(\alpha) \geq 1$

$$|I_{\alpha, t_0, t_0+h}| \leq \frac{1}{\iota(\alpha)!} h^{\iota(\alpha)} (2\|\eta\|_\infty)^{|\alpha|} \leq h (2\|\eta\|_\infty)^{|\alpha|},$$

and by the fundamental theorem of calculus,

$$F_\alpha(x_2, \eta_{t_0}^{\times \iota(\alpha)}) - F_\alpha(x_1, \eta_{t_0}^{\times \iota(\alpha)}) = \int_0^1 (\partial_x F_\alpha)(x_1 + s(x_2 - x_1), \eta_{t_0}^{\times \iota(\alpha)}) \cdot (x_2 - x_1) ds.$$

Combining the above results gives

$$\begin{aligned} & \|(\Phi_\gamma(x_1, t_0, h, \omega) - x_1) - (\Phi_\gamma(x_2, t_0, h, \omega) - x_2)\|_2 \\ & \leq \sum_{\alpha \in \mathcal{M}_m^\gamma \setminus \{\emptyset\}} \|F_\alpha(x_1, \eta_{t_0}^{\times \iota(\alpha)}(\omega)) - F_\alpha(x_2, \eta_{t_0}^{\times \iota(\alpha)}(\omega))\|_2 |I_{\alpha, t_0, t_0+h}(\omega)| \\ & \leq h \cdot \kappa_\gamma(\omega) \cdot \|x_1 - x_2\|_2, \end{aligned}$$

which is the assertion of the Lemma. \square

12.1.1 The Essential RODE-Taylor Schemes

Although the γ -RODE-Taylor scheme was defined for every $\gamma > 0$, only some specific $\gamma > 0$ are important. For the vector $\Theta = (\theta_1, \dots, \theta_m)$ in Assumption 8.1.1, define the set

$$\begin{aligned} \Lambda_\Theta &:= \mathbb{N}_0 + \theta_1 \mathbb{N}_0 + \dots + \theta_m \mathbb{N}_0 \\ &= \left\{ k + l_1 \theta_1 + \dots + l_m \theta_m \mid k, l_1, \dots, l_m \in \mathbb{N}_0 \right\}. \end{aligned}$$

Then Λ_Θ coincides with a sequence $0 = \gamma_0 < \gamma_1 < \gamma_2 < \dots$. For any $\gamma > 0$, let

$$[\gamma] = [\gamma]_\Theta := \min_{\substack{\gamma_i \geq \gamma \\ \gamma_i \in \Lambda_\Theta}} \gamma_i.$$

Clearly, $\gamma \leq [\gamma]$ and $[\gamma] \in \Lambda_\Theta$.

Recall from Lemma 8.2 that for $\gamma, \tilde{\gamma} > 0$

$$\mathcal{M}_m^\gamma = \mathcal{M}_m^{\tilde{\gamma}} \quad \text{if and only if} \quad [\gamma] = [\tilde{\gamma}].$$

In particular, $\mathcal{M}_m^\gamma = \mathcal{M}_m^{[\gamma]}$ and $\Phi_\gamma = \Phi_{[\gamma]}$. This implies that for $\gamma \in (\gamma_n, \gamma_{n+1}]$ with some $n \in \mathbb{N}_0$, the γ -RODE-Taylor scheme is the same as the γ_{n+1} -RODE-Taylor scheme. Hence only the γ_n -RODE-Taylor schemes for these γ_n need to be considered.

12.2 Examples of the RODE-Taylor Schemes

The three examples presented in Sect. 8.5 to illustrate RODE-Taylor expansions will be revisited here to illustrate the RODE-Taylor schemes. In particular, given $0 \leq t_n < t_{n+1} \leq T$ and $x_n \in \mathbb{R}^d$, the γ -RODE-Taylor scheme (12.4) with constant stepsize $h \in (0, 1]$ will be determined for different values of γ , Θ , d and m .

In the following examples all the coefficient functions are evaluated at (η_{t_n}, x_n) . For simplicity denote by $\underbrace{g_x \dots x}_{k \text{ times}} \underbrace{w \dots w}_{l \text{ times}}$ the partial derivative $\partial_x^k \partial_w^l g$, and denote

$$\begin{aligned} I_{\alpha,n} &:= I_{\alpha,t_n, t_{n+1}} \text{ where } I_{\alpha,t_n, t_{n+1}} \text{ is defined by (12.3). In particular,} \\ I_{(i),n} &= \int_{t_n}^{t_{n+1}} (\Delta \eta_{t_n, t})^i dt, \quad \text{for } i = 1, \dots, 6 \\ I_{(1,0),n} &= \int_{t_n}^{t_{n+1}} \Delta \eta_{t_n, t} \Delta_{t_n, t} dt, \quad I_{(0,1),n} = \int_{t_n}^{t_{n+1}} \int_{t_n}^s \Delta \eta_{t_n, t} dt ds, \\ I_{(1,1),n} &= \int_{t_n}^{t_{n+1}} \int_{t_n}^s \Delta \eta_{t_n, s} \Delta \eta_{t_n, t} dt ds, \quad I_{(2,0),n} = \int_{t_n}^{t_{n+1}} (\Delta \eta_{t_n, t})^2 \Delta_{t_n, t} dt, \\ I_{(0,2),n} &= \int_{t_n}^{t_{n+1}} \int_{t_n}^s (\Delta \eta_{t_n, t})^2 dt ds, \quad I_{(2,1),n} = \int_{t_n}^{t_{n+1}} \int_{t_n}^s (\Delta \eta_{t_n, s})^2 \Delta \eta_{t_n, t} dt ds, \\ I_{(1,2),n} &= \int_{t_n}^{t_{n+1}} \int_{t_n}^s \Delta \eta_{t_n, s} (\Delta \eta_{t_n, t})^2 dt ds, \quad I_{(3,0),n} = \int_{t_n}^{t_{n+1}} (\Delta \eta_{t_n, t})^3 \Delta_{t_n, t} dt, \\ I_{(0,3),n} &= \int_{t_n}^{t_{n+1}} \int_{t_n}^s (\Delta \eta_{t_n, t})^3 dt ds, \quad I_{(1,0,0),n} = \frac{1}{2} \int_{t_n}^{t_{n+1}} \Delta \eta_{t_n, t} (\Delta_{t_n, t})^2 dt, \\ I_{(0,1,0),n} &= \int_{t_n}^{t_{n+1}} \int_{t_n}^s \Delta \eta_{t_n, t} \Delta_{t_n, t} dt ds, \quad I_{(0,0,1),n} = \int_{t_n}^{t_{n+1}} \int_{t_n}^s \int_{t_n}^t \Delta \eta_{t_n, \tau} d\tau dt ds. \end{aligned}$$

Case I: $\gamma = 0.6, 1.6, 2.1$, $\Theta = (0.3)$, $d = m = 1$

In this scalar case $\gamma = 0.6$, $\gamma = 1.6$ and $\gamma = 2.1$ are considered. The corresponding multi-index sets are, respectively,

$$\mathcal{M}_1^{0.6} = \{\emptyset, (0), (1)\},$$

$$\mathcal{M}_1^{1.6} = \{\emptyset, (0), (1), (2), (3), (4), (5), (0, 0), (1, 0), (0, 1)\}$$

and

$$\mathcal{M}_1^{2.1} = \left\{ \begin{array}{l} \emptyset, (0), (1), (2), (3), (4), (5), (6), (0, 0), (1, 0), (0, 1), \\ (1, 1), (2, 0), (0, 2), (3, 0), (0, 3), (2, 1), (1, 2), (0, 0, 0) \end{array} \right\}.$$

Therefore the 0.6-RODE-Taylor scheme reads

$$x_{n+1}^{(h)} = x_n + hg + g_w \int_{t_n}^{t_{n+1}} \Delta \eta_{t_n, t} dt,$$

the 1.6-RODE-Taylor scheme reads

$$x_{n+1}^{(h)} = x_n + hg + \sum_{i=1}^5 \frac{1}{i!} \partial_w^i g I_{(i),n} + g_x g \frac{h^2}{2} + g_{xw} g I_{(1,0),n} + g_x g_w I_{(0,1),n}$$

and the 2.1-RODE-Taylor scheme reads

$$\begin{aligned} x_{n+1}^{(h)} = & x_n + hg + \sum_{i=1}^5 \frac{1}{i!} \partial_w^i g I_{(i),n} + g_x g \frac{h^2}{2} + g_{xw} g I_{(1,0),n} + g_x g_w I_{(0,1),n} \\ & + \frac{1}{6!} \partial_w^6 g I_{(6),n} + g_{wx} g_w I_{(1,1),n} + \frac{1}{2} g_{wwx} g I_{(2,0),n} \\ & + \frac{1}{2} g_x g_{ww} I_{(0,2),n} + \frac{1}{2} g_{wwx} g_w I_{(2,1),n} + \frac{1}{2} g_{wx} g_{ww} I_{(1,2),n} \\ & + \frac{1}{6} (\partial_w^3 g_x) g I_{(3,0),n} + \frac{1}{6} g_x g_{www} I_{(0,3),n} + \frac{1}{6} g_x^2 g h^3 + \frac{1}{6} g_{xx} g^2 h^3. \end{aligned}$$

Case II: $\gamma = 3, \Theta = \left(\frac{1}{2}\right), d = m = 1$

Here the RODE and the driving stochastic process η_t are 1-dimensional and the noise process could be a Wiener process.

Choose $\gamma = \gamma_6 = 3.0 \in \mathbb{N}_0 + \frac{1}{2} \mathbb{N}_0$ with the multi-index set

$$\mathcal{M}_1^3 = \left\{ \emptyset, (0), (1), (2), (3), (4), (5), (0, 0), (1, 0), (0, 1), (1, 1), (2, 0), \right. \\ \left. (0, 2), (2, 1), (1, 2), (3, 0), (0, 3), (0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1) \right\}.$$

The 3.0-RODE-Taylor scheme reads

$$\begin{aligned} x_{n+1}^{(h)} = & x_n + hg + \sum_{i=1}^5 \frac{1}{i!} \partial_w^i g I_{(i),n} + \frac{1}{2} g_x g h^2 + g_{wx} g I_{(1,0),n} \\ & + g_x g_w I_{(0,1),n} + g_{wx} g_w I_{(1,1),n} + \frac{1}{2} g_{wwx} g I_{(2,0),n} + \frac{1}{2} g_x g_{ww} I_{(0,2),n} \\ & + \frac{1}{2} g_{wwx} g_w I_{(2,1),n} + \frac{1}{2} g_{wx} g_{ww} I_{(1,2),n} + \frac{1}{6} (\partial_w^3 g_x) g I_{(3,0),n} + \frac{1}{6} g_x^2 g h^3 \\ & + \frac{1}{6} g_x g_{www} I_{(0,3),n} + \frac{1}{6} g_{xx} g^2 h^3 + (g_{wx} g_x g + g_{wx} g^2) I_{(1,0,0),n} \\ & + (g_x g_{wx} g + g_{xx} g_w g) I_{(0,1,0),n} + (g_x^2 g_w + g_{xx} g g_w) I_{(0,0,1),n}. \end{aligned}$$

Case III: $\gamma = 1.5, \Theta = \left(\frac{1}{2}, \frac{3}{4}\right), d = m = 2$

Both the RODE and the driving stochastic process $\eta_t = (\eta_t^1, \eta_t^2)$ are 2-dimensional. In view of the choice of Θ , the first component of the noise process could be a Wiener process and the second could be a fractional Brownian motion with Hurst Index $H = \frac{3}{4}$.

Choose $\gamma = \gamma_5 = 1.5$ with the multi-index set

$$\mathcal{M}_2^{1.5} = \left\{ \emptyset, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \right\}.$$

The 1.5-RODE-Taylor scheme reads

$$\begin{aligned} x_{n+1}^{(h)} = & x_n + hg + g_{w_1} \int_{t_n}^{t_{n+1}} \Delta \eta_{t_n, t}^1 dt + g_{w_2} \int_{t_n}^{t_{n+1}} \Delta \eta_{t_n, t}^2 dt \\ & + g_{w_1 w_2} \int_{t_n}^{t_{n+1}} \Delta \eta_{t_n, t}^1 \Delta \eta_{t_n, t}^2 dt + \frac{1}{2} g_{w_1 w_1} \int_{t_n}^{t_{n+1}} (\Delta \eta_{t_n, t}^1)^2 dt \\ & + \frac{1}{2} ((g_{x_1}) g^1 + (g_{x_2}) g^2) h^2, \end{aligned}$$

where g^k denotes the k th component of function g .

12.3 RODEs with Affine Noise

RODE-Taylor schemes simplify considerable when the vector field has an affine structure in the noise, i.e.,

$$g(x, w) = f(x) + \sigma(x)w.$$

Then all terms involving two or more partial differentiations in the noise variable w vanish. Given a multi-index set 鳌 (*fēng*), the effective multi-index set 穿 (*yún*) is obtained from 鳌 by deleting all multi-indices in which the first component is 2 or higher for a scalar noise, and deleting all multi-indices in which the sum of the noise components is 2 or more for higher dimensional noises.

Example 12.1 The 1.5-RODE-Taylor scheme corresponding to the multi-index set 鳌 = {(0, 0), (1, 0), (2, 0), (0, 1)} reduces to

$$x_{n+1} = x_n + hg + g_w \int_{t_n}^{t_{n+1}} \Delta \eta_{t_n, t} dt + \frac{1}{2} g_x g h^2$$

with the effective index set 穿 = {(0, 0), (1, 0), (0, 1)}.

Example 12.2 For the multi-index set 鳌 = {(0, 0), (1, 0), (2, 0), (3, 0), (0, 1), (1, 1)} the corresponding 2.0-RODE-Taylor scheme reads

$$x_{n+1} = x_n + hg + g_w \int_{t_n}^{t_{n+1}} \Delta \eta_{t_n, t} dt + \frac{1}{2} g_x g h^2$$

$$+ g_x g_w \int_{t_n}^{t_{n+1}} \int_{t_n}^t \Delta \eta_{t_n, s} \, ds \, dt + g_{wx} g \int_{t_n}^{t_{n+1}} \Delta \eta_{t_n, t} \Delta_{t_n, t} \, dt$$

with the effective index set $\bar{\mathbb{A}} = \{(0, 0), (1, 0), (0, 1), (1, 1)\}$.

Example 12.3 The 1.5-RODE-Taylor scheme with a 2-dimensional Wiener process corresponding to the multi-index set $\bar{\mathbb{A}} = \{(0, 0, 0), (1, 0, 0), (0, 1, 0), (1, 1, 0), (2, 0, 0), (0, 2, 0), (0, 0, 1)\}$ reduces to

$$x_{n+1} = x_n + hg + g_{w1} \int_{t_n}^{t_{n+1}} \Delta \eta_{t_n, t}^1 \, dt + g_{w2} \int_{t_n}^{t_{n+1}} \Delta \eta_{t_n, t}^2 \, dt + \frac{1}{2} g_x g h^2$$

with effective index set $\bar{\mathbb{A}} = \{(0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1)\}$.

12.4 Other Numerical Schemes for RODEs

Some other numerical schemes for RODEs from the literature that were mentioned in Sect. 1.1 are compared here briefly with the above RODE-Taylor schemes. They are also one-step schemes, so only the functions $\Phi(x, t, h)$ on the right-hand side will be provided below.

12.4.1 The Local Linearisation Scheme for RODEs

A local linearisation method (LL) for RODEs, which was proposed by Carbonell, Jimenez, Biscay and de la Cruz [30], was already mentioned in Chap. 1. In the context of this chapter, the LL scheme has the form

$$x_{n+1} = \Phi_{LL}(x_n, t_n, h_n), \quad (12.5)$$

where the function Φ_{LL} is defined in a general form by

$$\Phi_{LL}(x, t, h) = x + \int_t^{t+h} e^{g_x \cdot (t+h-s)} \, ds \cdot g + \sum_{j=1}^m \int_t^{t+h} e^{g_x \cdot (t+h-s)} \Delta \eta_{t,s}^j \, ds \cdot g_{w_j},$$

or, after simplifying the first integral,

$$\Phi_{LL}(x, t, h) = x + \int_0^h e^{g_x \cdot s} \, ds \cdot g + \sum_{j=1}^m \int_t^{t+h} e^{g_x \cdot (t+h-s)} \Delta \eta_{t,s}^j \, ds \cdot g_{w_j}.$$

In the LL scheme, the functions g , g_x and g_{w_j} are evaluated at (x_n, η_{t_n}) , where x_n is the numerical simulation of $x(t_n)$. Note that $g_x(x_n, \eta_{t_n})$ is a $d \times d$ -matrix, so $e^{g_x(x_n, \eta_{t_n}) \cdot (t_n + h_n - s)}$ is a matrix exponential function, while the $g_{w_j}(x_n, \eta_{t_n})$ are vectors.

In [30] it was assumed that all of the components of the driving noise were of the same type, i.e., $\Theta = (\theta, \dots, \theta) \in (0, 1]^m$, and that the vector field g was globally Lipschitz continuous in the state variable and uniformly continuous in the noise variables. Under these assumptions, it was shown in [30] that the LL scheme converges with order 2θ . It is thus comparable with a 2θ -RODE-Taylor scheme, which has the following form:

$$x_{n+1}^{(h)} = \begin{cases} x_n + hg + \sum_{j=1}^m g_{w_j} \int_{t_n}^{t_n+h} \Delta \eta_{t_n,s}^j \, ds, & \theta \leq \frac{1}{2} \\ x_n + hg + \sum_{j=1}^m g_{w_j} \int_{t_n}^{t_n+h} \Delta \eta_{t_n,s}^j \, ds + \frac{1}{2} g_x g \, h^2, & \theta > \frac{1}{2}. \end{cases}$$

In the second case where $\theta > 1/2$,

$$\begin{aligned} x_{n+1}^{(h)} &= x_n + hg + \sum_{j=1}^m g_{w_j} \int_{t_n}^{t_n+h} \Delta \eta_{t_n,s}^j \, ds + \frac{1}{2} g_x g \, h^2 \\ &= x_n + \left(h + \frac{1}{2} h^2 g_x \right) g + \sum_{j=1}^m g_{w_j} \int_{t_n}^{t_n+h} \Delta \eta_{t_n,s}^j \, ds \\ &= x_n + (g_x)^{-1} \left(hg_x + \frac{1}{2} (hg_x)^2 \right) g + \sum_{j=1}^m g_{w_j} \int_{t_n}^{t_n+h} \Delta \eta_{t_n,s}^j \, ds \\ &\approx x_n + (g_x)^{-1} (e^{hg_x} - I) g + \sum_{j=1}^m g_{w_j} \int_{t_n}^{t_n+h} \Delta \eta_{t_n,s}^j \, ds \\ &= x_n + \int_0^h e^{g_x \cdot s} \, ds \cdot g + \sum_{j=1}^m \int_{t_n}^{t_n+h} \Delta \eta_{t_n,s}^j \, ds \cdot g_{w_j} \\ &\approx \Phi_{LL}(y_n, t_n, h), \end{aligned}$$

provided g_x is invertible. The coefficients and integrals of the RODE-Taylor scheme are easier to compute when the noise process is scalar or low dimensional and the state space dimensional is high, since the LL scheme (12.5) requires one to calculate a large exponential matrix $e^{g_x \cdot (t_n + h - s)}$ for $s \in [t_n, t_n + h]$ in approximations of the integrals

$$\int_{t_n}^{t_n+h} e^{g_x \cdot (t_n + h - s)} \Delta \eta_{t_n,s}^j \, ds, \quad j = 1, \dots, m.$$

They still, however, require the same number of evaluations of the driving noise process η_t (see Chap. 14).

12.4.2 The Averaged Euler Scheme

As already mentioned in Chap. 1, Grüne and Kloeden [54] introduced the averaged Euler scheme for RODEs with a vector field g of the affine structure ($m = d + 1$)

$$g(x, w_1, \dots, w_m, w_{m+1}) = \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix} + w_{m+1} \cdot \sigma(x)$$

for a smooth function $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^d$. The averaged Euler scheme for such RODEs reads

$$x_{n+1} = \Phi_{AE}(x_n, t_n, h_n),$$

where Φ_{AE} is given by

$$\Phi_{AE}(x, t, h) = x + \int_t^{t+h} \begin{pmatrix} \eta_s^1 \\ \vdots \\ \eta_s^m \end{pmatrix} ds + \sigma(x) \int_t^{t+h} \eta_s^{m+1} ds,$$

and has order 1. It is easy to see that the averaged Euler scheme is exactly the same as the 1-RODE-Taylor scheme for such RODEs.

12.4.3 Heuristic RODE-Taylor Schemes

In an earlier paper [87] Jentzen and Kloeden introduced numerical schemes for RODEs that converge with arbitrary high order. The schemes there are very similar to the RODE-Taylor schemes in this chapter, but use a more heuristic derivation, which was sketched in Chap. 1.

For the case $d = m = 1$ and $\Theta = (\frac{1}{2})$ the schemes presented in [87] are the RODE-Taylor schemes of order $\gamma = 0.5, 1.0, 1.5, 2.0$ and 2.5 . However, for higher order they contain not only the terms of the RODE-Taylor schemes given in this chapter, but often more terms than needed for the required order. The reason for this is the recursive technique used in [87].

A tedious calculation in [87] gives the following order 3.0-RODE-Taylor-scheme

$$x_{n+1} = \Phi_{HE}(x_n, t_n, h_n),$$

with

$$\Phi_{HE}(x, t, h) = \Phi_{3.0}(x, t, h) + \frac{1}{2} g_{xx} g_w^2 \int_t^{t+h} \left(\int_t^s \Delta \eta_{t,\tau} d\tau \right)^2 ds,$$

where $\Phi_{3.0}(x, t, h)$ is the 3.0-RODE-Taylor scheme (12.2).

12.5 Endnotes

Jentzen and Kloeden [87] (see also [76]) uses Taylor expansions of the vector field of the RODE to obtain an implicit integral equation expansion for the solutions into which lower order approximations are inserted to provide an explicit approximation of the solution that will be called RODE-Taylor expansion. This was sketched in Chap. 1.

This chapter is based on the more rigorous and direct approach of Jentzen and Kloeden [78] that was presented and used to derive RODE-Taylor approximations in Chap. 8. It is motivated by the Wagner–Platen multi-index notation that was used in [91] to formulate succinctly stochastic Taylor expansions and Taylor schemes for SODEs, see also Chap. 6. The main difference is that matrix-valued multi-indices are now required to handle the possibly different Hölder exponents of the different components of the driving noise processes. As in the SODE case the coefficient functions are obtained by iterated application of differential operators to the vector field of the RODE, but instead of iterated integrals of the components of the noise processes single integrals of suitable powers of increments of them are used.

Becker [14] provides MATLAB programs for some of the schemes in this chapter.

The Chinese characters 風 and 云 used for various multi-index sets are pronounced *fēng* and *yún*, respectively. Their actual meanings “wind” and “cloud” have no significance here.

Chapter 13

Numerical Stability

It is well known from the theory of classical Runge–Kutta schemes for ODEs that implicit schemes are necessary for the stable integration of stiff ODEs. Since RODEs are pathwise ODEs, many of the numerical stability concepts for ODEs can in principle be applied to RODEs too. However RODEs are intrinsically nonautonomous due to the time-varying nature of the driving noise process. Moreover, the nonlinearities and global effects are much more prevalent in nonautonomous and random systems. These restrict the useful means of simpler numerical stability ideas for linear ODEs in the context of RODEs.

For the above reasons, this chapter will focus on the nonlinear numerical concept of B-stability. This means that the non-expansive structure of trajectories of ODEs is preserved, i.e., any two solutions x_n and \tilde{x}_n of the scheme satisfy

$$\|x_{n+1} - \tilde{x}_{n+1}\| \leq \|x_n - \tilde{x}_n\|, \quad n = 0, 1, \dots, N-1,$$

when the ODE satisfies a dissipative one-sided Lipschitz condition. Recall that no explicit or linear implicit Runge–Kutta scheme is ever B-stable, so one may expect the same resulting from RODEs.

In this chapter it will be assumed that the vector field in the RODE

$$\frac{dx}{dt} = g(x, \eta_t) \tag{13.1}$$

satisfies the following uniform one-sided Lipschitz condition.

Assumption 13.0.1 (*One-sided Lipschitz condition*) There exists a constant $\kappa \in \mathbb{R}$ such that

$$\langle g(x, w) - g(\tilde{x}, w), x - \tilde{x} \rangle \leq \kappa \|x - \tilde{x}\|^2 \tag{13.2}$$

for all $x, \tilde{x} \in \mathbb{R}^d$ and $w \in \mathbb{R}^m$. When $\kappa < 0$, this condition is called a dissipative one-sided Lipschitz condition.

Below, the partial derivatives of g with respect to the state variable and noise variable will be denoted by g_x, g_w (g_{xx}, g_{ww} , etc.), respectively. The operator norm on matrices induced by the above vector norm $\|\cdot\|$ will also be denoted by $\|\cdot\|$ when the context is clear. For simplicity, an equidistant partition of the time interval $[0, T]$ under consideration with the constant step size $h = h_N = T/N$ and discretisation times $t_n = nh$ for $n = 0, 1, \dots, N$ will be used.

13.1 B-Stability of the Implicit Averaged Schemes

Two implicit averaged numerical schemes, which use the averaged noise integral

$$I_n(\omega) = I_n^{(h)}(\omega) := \frac{1}{h} \int_{nh}^{(n+1)h} \eta_s(\omega) \, ds$$

within the vector field of the RODE (13.1), will be considered here. They are the *implicit averaged Euler scheme* (IAES) given by

$$x_{n+1} = x_n + g(x_{n+1}, I_n) h$$

and the *implicit averaged midpoint scheme* (IAMS) given by

$$x_{n+1} = x_n + g\left(\frac{1}{2}(x_n + x_{n+1}), I_n\right) h$$

for $n = 0, 1, \dots, N - 1$.

The convergence and unique solvability of these schemes were established in [12].

The Implicit Averaged Euler Scheme

Consider two solutions x_n and \tilde{x}_n of the IAES and denote

$$\delta := g(x_{n+1}, I_n) h - g(\tilde{x}_{n+1}, I_n) h$$

Then,

$$(x_{n+1} - \tilde{x}_{n+1}) = (x_n - \tilde{x}_n) + \delta$$

and hence

$$\begin{aligned} \|x_{n+1} - \tilde{x}_{n+1}\|^2 &= \|(x_n - \tilde{x}_n) + \delta\|^2 \\ &= \|x_n - \tilde{x}_n\|^2 + 2 \langle x_n - \tilde{x}_n, \delta \rangle + \|\delta\|^2 \\ &= \|x_n - \tilde{x}_n\|^2 + 2 \langle x_{n+1} - \tilde{x}_{n+1} - \delta, \delta \rangle + \|\delta\|^2 \\ &= \|x_n - \tilde{x}_n\|^2 + 2 \langle x_{n+1} - \tilde{x}_{n+1}, \delta \rangle - \|\delta\|^2 \end{aligned}$$

$$\leq \|x_n - \tilde{x}_n\|^2 + 2 \langle x_{n+1} - \tilde{x}_{n+1}, \delta \rangle. \quad (13.3)$$

By the one-sided Lipschitz condition (13.2) and the assumption that $\kappa \leq 0$,

$$\langle x_{n+1} - \tilde{x}_{n+1}, \delta \rangle \leq \kappa h \|x_{n+1} - \tilde{x}_{n+1}\|^2 \leq 0. \quad (13.4)$$

Hence, (13.3) and (13.4) together give

$$\|x_{n+1} - \tilde{x}_{n+1}\|^2 \leq \|x_n - \tilde{x}_n\|^2,$$

which establishes the B-stability of the implicit averaged Euler scheme.

The Implicit Averaged Midpoint Scheme

Similarly, for the IAMS

$$(x_{n+1} - \tilde{x}_{n+1}) = (x_n - \tilde{x}_n) + \delta, \quad (13.5)$$

but now with

$$\delta := g \left(\frac{1}{2} (x_n + x_{n+1}), I_n \right) h - g \left(\frac{1}{2} (\tilde{x}_n + \tilde{x}_{n+1}), I_n \right) h.$$

Note that (13.5) is equivalent to

$$\frac{1}{2} (x_n + x_{n+1}) - \frac{1}{2} (\tilde{x}_n + \tilde{x}_{n+1}) = (x_n - \tilde{x}_n) + \frac{1}{2} \delta.$$

Hence,

$$\begin{aligned} \|x_{n+1} - \tilde{x}_{n+1}\|^2 &= \|(x_n - \tilde{x}_n) + \delta\|^2 \\ &= \|x_n - \tilde{x}_n\|^2 + 2 \langle x_n - \tilde{x}_n, \delta \rangle + \|\delta\|^2 \\ &= \|x_n - \tilde{x}_n\|^2 + \|\delta\|^2 + 2 \left\langle \frac{1}{2} (x_n + x_{n+1}) - \frac{1}{2} (\tilde{x}_n + \tilde{x}_{n+1}) - \frac{1}{2} \delta, \delta \right\rangle \\ &= \|x_n - \tilde{x}_n\|^2 + 2 \left\langle \frac{1}{2} (x_{n+1} + x_n) - \frac{1}{2} (\tilde{x}_{n+1} + \tilde{x}_n), \delta \right\rangle. \end{aligned}$$

Finally, the one-sided Lipschitz condition (13.2) on g and the assumption that $\kappa \leq 0$ yield

$$\|x_{n+1} - \tilde{x}_{n+1}\|^2 \leq \|x_n - \tilde{x}_n\|^2,$$

which implies that the implicit averaged midpoint scheme is B-stable.

13.2 B-Stability of the Implicit Multi-step Schemes

Stochastic linear multi-step schemes (SLMMs) were developed in Chap. 10 for RODEs driven by an Itô diffusion process, i.e., the solution of an Itô stochastic differential equation. The essential ideas to prove the B-stability of implicit linear multi-step schemes will be illustrated here by the fully scalar case of the order 1.5 implicit SLMM.

Consider a scalar RODE (13.1) driven by a scalar Itô stochastic differential equation

$$dY_t = f(Y_t) dt + \sigma(Y_t) dW_t, \quad (13.6)$$

that together form a coupled Itô SODE in \mathbb{R}^2

$$d \begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} g(X_t, Y_t) \\ f(Y_t) \end{pmatrix} dt + \begin{pmatrix} 0 \\ \sigma(Y_t) \end{pmatrix} dW_t.$$

Assume that the coefficients of the RODE (13.1) and the SODE (13.6) satisfy the standard assumptions, i.e., all relevant derivatives are continuous and uniformly bounded.

Recall from Chap. 10 that the order 1.5 implicit SLMM is given by

$$\begin{aligned} x_{n+1} &= \frac{1}{2}x_n + \frac{1}{2}x_{n-1} + h \left(\frac{1}{2}g_{n+1} + \frac{1}{2}g_n + \frac{1}{2}g_{n-1} \right) \\ &\quad + \left(L^1 g_n I_{(1,0),n} + \frac{1}{2} L^1 g_{n-1} I_{(1,0),n-1} \right) - \frac{1}{2} L^1 g_n h I_{(1),n} \\ &\quad + \left(\frac{1}{4} L^0 g_n I_{(0,0),n} + \frac{1}{4} L^0 g_{n-1} I_{(0,0),n-1} \right), \end{aligned} \quad (13.7)$$

where the coefficients are evaluated at (x_{n+1}, y_{n+1}) , (x_n, y_n) , and (x_{n-1}, y_{n-1}) , respectively. In this case the differential operators L^0 and L^1 reduce to

$$L^0 = \frac{\partial}{\partial t} + g(x, y) \frac{\partial}{\partial x} + f(y) \frac{\partial}{\partial y} + \frac{1}{2} \sigma^2(y) \frac{\partial^2}{\partial y^2}, \quad L^1 = \sigma(y) \frac{\partial}{\partial y}.$$

The difference between any two solutions x_n and \tilde{x}_n of the scheme (13.7) is calculated to be

$$\begin{aligned} x_{n+1} - \tilde{x}_{n+1} &= \frac{1}{2}(x_n - \tilde{x}_n) + \frac{1}{2}(x_{n-1} - \tilde{x}_{n-1}) \\ &\quad + \frac{1}{2}h \left\{ (g(x_{n+1}, y_{n+1}) - g(\tilde{x}_{n+1}, y_{n+1})) + (g(x_n, y_n) - g(\tilde{x}_n, y_n)) \right. \\ &\quad \left. + (g(x_{n-1}, y_{n-1}) - g(\tilde{x}_{n-1}, y_{n-1})) \right\} \end{aligned}$$

$$\begin{aligned}
& + \left(I_{(1,0),n} - \frac{1}{2}hI_{(1),n} \right) (L^1 g(x_n, y_n) - L^1 g(\tilde{x}_n, y_n)) \\
& + \frac{1}{2} I_{(1,0),n-1} (L^1 g(x_{n-1}, y_{n-1}) - L^1 g(\tilde{x}_{n-1}, y_{n-1})) \\
& + \frac{1}{4} I_{(0,0),n} (L^0 g(x_n, y_n) - L^0 g(\tilde{x}_n, y_n)) \\
& + \frac{1}{4} I_{(0,0),n-1} (L^0 g(x_{n-1}, y_{n-1}) - L^0 g(\tilde{x}_{n-1}, y_{n-1})). \tag{13.8}
\end{aligned}$$

Taking inner product of both sides of (13.8) with $x_{n+1} - \tilde{x}_{n+1}$ yields

$$\begin{aligned}
& \|x_{n+1} - \tilde{x}_{n+1}\|^2 = \frac{1}{2}h \langle g(x_{n+1}, y_{n+1}) - g(\tilde{x}_{n+1}, y_{n+1}), x_{n+1} - \tilde{x}_{n+1} \rangle \\
& + \frac{1}{2} \langle x_n - \tilde{x}_n, x_{n+1} - \tilde{x}_{n+1} \rangle + \frac{1}{2}h \langle g(x_n, y_n) - g(\tilde{x}_n, y_n), x_{n+1} - \tilde{x}_{n+1} \rangle \\
& + \left(I_{(1,0),n} - \frac{1}{2}hI_{(1),n} \right) \langle L^1 g(x_n, y_n) - L^1 g(\tilde{x}_n, y_n), x_{n+1} - \tilde{x}_{n+1} \rangle \\
& + \frac{1}{4} I_{(0,0),n} \langle L^0 g(x_n, y_n) - L^0 g(\tilde{x}_n, y_n), x_{n+1} - \tilde{x}_{n+1} \rangle \\
& + \frac{1}{2} \langle x_{n-1} - \tilde{x}_{n-1}, x_{n+1} - \tilde{x}_{n+1} \rangle \\
& + \frac{1}{2}h \langle g(x_{n-1}, y_{n-1}) - g(\tilde{x}_{n-1}, y_{n-1}), x_{n+1} - \tilde{x}_{n+1} \rangle \\
& + \frac{1}{2} I_{(1,0),n-1} \langle L^1 g(x_{n-1}, y_{n-1}) - L^1 g(\tilde{x}_{n-1}, y_{n-1}), x_{n+1} - \tilde{x}_{n+1} \rangle \\
& + \frac{1}{4} I_{(0,0),n-1} \langle L^0 g(x_{n-1}, y_{n-1}) - L^0 g(\tilde{x}_{n-1}, y_{n-1}), x_{n+1} - \tilde{x}_{n+1} \rangle.
\end{aligned}$$

Then the one-sided Lipschitz condition (13.2) and the standard Lipschitz assumption $\|L^\alpha x - L^\alpha y\| \leq K\|x - y\|$ together with the Schwarz inequality give

$$\begin{aligned}
& \left(1 - \frac{1}{2}hL \right) \|x_{n+1} - \tilde{x}_{n+1}\|^2 \\
& \leq \frac{1}{2} \|x_n - \tilde{x}_n\| \cdot \|x_{n+1} - \tilde{x}_{n+1}\| + \frac{1}{2}hK \|x_n - \tilde{x}_n\| \cdot \|x_{n+1} - \tilde{x}_{n+1}\| \\
& \quad + K \left(|I_{(1,0),n}| + \frac{1}{2}h |I_{(1),n}| \right) \|x_n - \tilde{x}_n\| \cdot \|x_{n+1} - \tilde{x}_{n+1}\| \\
& \quad + \frac{K}{4} |I_{(0,0),n}| \cdot \|x_n - \tilde{x}_n\| \cdot \|x_{n+1} - \tilde{x}_{n+1}\|
\end{aligned} \tag{13.9}$$

$$\begin{aligned}
& + \frac{1}{2} \|x_{n-1} - \tilde{x}_{n-1}\| \cdot \|x_{n+1} - \tilde{x}_{n+1}\| + \frac{1}{2} h K \|x_{n-1} - \tilde{x}_{n-1}\| \cdot \|x_{n+1} - \tilde{x}_{n+1}\| \\
& + \frac{K}{2} |I_{(1,0),n-1}| \cdot \|x_{n-1} - \tilde{x}_{n-1}\| \cdot \|x_{n+1} - \tilde{x}_{n+1}\| \\
& + \frac{K}{4} |I_{(0,0),n-1}| \cdot \|x_{n-1} - \tilde{x}_{n-1}\| \cdot \|x_{n+1} - \tilde{x}_{n+1}\|.
\end{aligned}$$

Applying triangle inequalities to (13.9) results in

$$\|x_{n+1} - \tilde{x}_{n+1}\| \leq \frac{1}{(1 - \frac{1}{2}hL)} \left(K_n \|x_n - \tilde{x}_n\| + K_{n-1} \|x_{n-1} - \tilde{x}_{n-1}\| \right), \quad (13.10)$$

where

$$\begin{aligned}
K_n &= \frac{1}{2}(1 + Kh) + \frac{1}{2}K \left(2 |I_{(1,0),n}| + h |I_{(1),n}| + \frac{1}{2} I_{(0,0),n} \right) \\
&\sim \frac{1}{2}(1 + Kh) + \mathcal{O}(h^{\frac{3}{2}}), \\
K_{n-1} &= \frac{1}{2}(1 + Kh) + K \left(|I_{(1,0),n-1}| + \frac{1}{4} I_{(0,0),n-1} \right) \\
&\sim \frac{1}{2}(1 + Kh) + \mathcal{O}(h^{\frac{3}{2}}).
\end{aligned}$$

Subtracting $\|x_{n-1} - \tilde{x}_{n-1}\|$ from both sides of the inequality (13.10) then yields

$$\begin{aligned}
& \|x_{n+1} - \tilde{x}_{n+1}\| - \|x_{n-1} - \tilde{x}_{n-1}\| \\
& \leq \frac{1}{(1 - \frac{1}{2}hL)} \left(K_n \|x_n - \tilde{x}_n\| - (1 - \frac{1}{2}hL - K_{n-1}) \|x_{n-1} - \tilde{x}_{n-1}\| \right) \\
& \leq \frac{K_n}{(1 - \frac{1}{2}hL)} (\|x_n - \tilde{x}_n\| - \|x_{n-1} - \tilde{x}_{n-1}\|)
\end{aligned}$$

when

$$K_n \leq 1 - \frac{1}{2}hL - K_{n-1}.$$

It follows immediately that

$$\|x_{n+1} - \tilde{x}_{n+1}\| \leq \|x_n - \tilde{x}_n\|$$

provided $L \leq -2K + \mathcal{O}(h^{\frac{1}{2}})$. The SLMM (13.7) is thus B-stable for all step sizes $h > 0$, provided the dissipativity (indicated by L) is strong enough to overcome the effect of the nonlinearities (indicated by K).

13.3 Endnotes

This chapter is based on the work in Asai, Herrmann and Kloeden [12] and Asai and Kloeden [10]. B-stability of various numerical schemes under a dissipative one-sided Lipschitz condition is discussed in Stuart and Humphries [128]. Numerical stability of ODE is treated extensively, e.g., in Griffiths and Higham [53], Hairer and Wanner [59, 60] and Hairer, Nørsett and Wanner [61].

Chapter 14

Stochastic Integrals: Simulation and Approximation

The numerical schemes derived in the previous chapters require the simulation of noise process η_t and its integrals

$$I_{(j_1, \dots, j_l), t_n, t_{n+1}} = \int_{t_n}^{t_{n+1}} \cdots \int_{t_n}^{s_{l-1}} \eta_{s_l}^{j_l} \cdots \eta_{s_1}^{j_1} ds_l \cdots ds_1$$

on each discretised subinterval $[t_n, t_{n+1}]$. The components η_t^j of the driving noise process are assumed to be at least Lebesgue integrable in time, in particular with essentially bounded sample paths.

In general, for processes with continuous or piecewise continuous sample paths, the integrals can be calculated using Riemann sums on much finer partition of the discretisation subinterval so that the error is dominated by local discretisation error of the scheme itself. For example, the averaged numerical schemes discussed in Chap. 13 involve the averaged noise integral

$$I_n(\omega) := \frac{1}{h_n} \int_{t_n}^{t_{n+1}} \eta_s(\omega) ds \quad (14.1)$$

on each discretisation subinterval $[t_n, t_{n+1}]$ with step size h_n . On the other hand, some integrals can be simulated directly if the distributions of η_t^j are known. For example, if the noise process η_t is a Wiener process or an Ornstein–Uhlenbeck (OU) process, then sample paths of the integrals can be simulated directly.

In this chapter, four types of noise processes, namely a Wiener process, an OU process, a compound Poisson process and a fractional Brownian motion (fBm), are introduced and their integrals are evaluated.

14.1 Calculating a Finer Approximation of the Same Sample Path

To investigate the convergence of a computed quantity, a finer partition of the time interval $[t_0, T]$ is used. If the step size is halved, then the original partition

$$t_0 < t_1 < \cdots < t_n < t_{n+1} < \cdots < t_N = T$$

with $h_n = t_{n+1} - t_n$ is replaced by a new partition

$$t'_0 = t_0 < t'_1 < t'_2 < \cdots < t'_{2N} = T$$

with

$$t'_{2n} = t_n, \quad t'_{2n+1} = t_n + \frac{1}{2}(t_{n+1} - t_n), \quad t'_{2n+2} = t_{n+1}.$$

The Lévy construction of a Wiener process allows an appropriate value of $W_{t'_{2n+1}}(\omega)$ to be calculated when $W_{t_n}(\omega)$ and $W_{t_{n+1}}(\omega)$ are known, i.e., for the same sample path. It is given by

$$W_{t'_{2n+1}}(\omega) = \frac{1}{2} (W_{t_n}(\omega) + W_{t_{n+1}}(\omega)) + \frac{\sqrt{h_n}}{2} Z_n(\omega),$$

where $Z_n \sim \mathcal{N}(0, 1)$ and different Z_n 's are independent. This relationship holds for the random variables as well as for the individual sample paths.

As a hint of the proof note that the new increments

$$\begin{aligned} \Delta' W_{2n} &= W_{t'_{2n+1}} - W_{t'_{2n}} = W_{t'_{2n+1}} - W_{t_n} = \frac{\sqrt{h_n}}{2} Z_n + \frac{1}{2} \Delta W_n \\ \Delta' W_{2n+1} &= W_{t'_{2n+2}} - W_{t'_{2n+1}} = W_{t_{n+1}} - W_{t'_{2n+1}} = -\frac{\sqrt{h_n}}{2} Z_n + \frac{1}{2} \Delta W_n \end{aligned}$$

are independent and are $\mathcal{N}(0, \frac{1}{2}h_n)$ -distributed with

$$\Delta' W_{2n} + \Delta' W_{2n+1} = \Delta W_n.$$

In particular, since these new increments are independent, their variances can be summed to give

$$\begin{aligned} \text{Var}[\Delta' W_{2n} + \Delta' W_{2n+1}] &= \text{Var}[\Delta' W_{2n}] + \text{Var}[\Delta' W_{2n+1}] \\ &= \frac{1}{2}h_n + \frac{1}{2}h_n = h_n = \text{Var}[\Delta W_n]. \end{aligned}$$

A similar construction can be derived in much the same way when the new time instant is not the midpoint of two others.

14.2 Integral of a Wiener Process

The integral $I(W) = \int_0^h W_t dt$ is an $\mathcal{N}(0, \frac{1}{3}h^3)$ -distributed random variable, so the integral $I = I(W)/h$ in the averaged numerical schemes (14.1) for RODEs can be simulated directly as

$$I = \frac{1}{\sqrt{3}} h^{1/2} Z,$$

where Z is an $\mathcal{N}(0, 1)$ -distributed random variable.

If both $I(W)$ and ΔW are needed in one numerical scheme, they can be simulated together by using the fact that they are correlated Gaussian random variables with distributions and correlation

$$\Delta W \sim \mathcal{N}(0, h), \quad I(W) \sim \mathcal{N}\left(0, \frac{1}{3}h^3\right), \quad \mathbb{E}[I(W)\Delta W] = \frac{1}{2}h^2. \quad (14.2)$$

The proof of (14.2) follows from the Itô formula to give the identity

$$d(t W_t) = t dW_t + W_t dt$$

and the fact that the Itô integral $\int_0^h t dW_t$ is $\mathcal{N}(0, \frac{1}{3}h^3)$ -distributed, since

$$\mathbb{E}\left[\int_0^h t dW_t\right] = 0, \quad \text{and} \quad \mathbb{E}\left[\left(\int_0^h t dW_t\right)^2\right] = \int_0^h t^2 dt = \frac{1}{3}h^3$$

by basic properties of the Itô integral.

Common sample paths of $I(W)$ and ΔW can be simulated using two independent $\mathcal{N}(0, 1)$ -distributed random variable Z_1 and Z_2 via the linear relationship

$$\Delta W = h^{1/2} Z_1, \quad I(W) = \frac{1}{2} h^{3/2} Z_1 + \frac{1}{\sqrt{12}} h^{3/2} Z_2.$$

This is useful for comparing the numerical schemes with the averaged integral simulated directly or approximated as a Riemann sum.

14.3 Integral of an Ornstein–Uhlenbeck Process

An Ornstein–Uhlenbeck (OU) stationary stochastic process O_t with positive constant parameters μ and σ is a solution to the scalar Itô SODE

$$dO_t = -\mu O_t dt + \sigma dW_t, \quad (14.3)$$

which has the explicit solution

$$O_t = O_0 e^{-\mu t} + \sigma e^{-\mu t} \int_0^t e^{\mu s} dW_s. \quad (14.4)$$

Let O_t be an OU process, then the integral $I(O) = \int_0^h O_t dt$ is an $\mathcal{N}(0, \varsigma^2)$ -distributed random variable, where the variance ς^2 is given by

$$\varsigma^2 := \frac{\sigma^2}{\mu^3} (\mu h - 1 + e^{-\mu h}).$$

The desired integral $I_h = \frac{1}{h} I(O)$ in the averaged numerical schemes for RODEs can thus be simulated directly as

$$I_h = \frac{\varsigma}{h} Z,$$

by using an $\mathcal{N}(0, 1)$ -distributed random variable Z .

Define

$$\tilde{\Delta}O := O_h - O_0 e^{-\mu h} = \sigma e^{-\mu h} \int_0^h e^{\mu s} dW_s,$$

Then by (14.4), $\tilde{\Delta}O$ is Gaussian distributed with mean $\mathbb{E}[\tilde{\Delta}O] = 0$ and, by the Itô isometry, variance

$$\begin{aligned} \text{Var}[\tilde{\Delta}O] &= \mathbb{E}[(\tilde{\Delta}O)^2] - 0 = \mathbb{E}\left[\left(\sigma e^{-\mu h} \int_0^h e^{\mu s} dW_s\right)^2\right] \\ &= \sigma^2 e^{-2\mu h} \int_0^h e^{2\mu s} ds = \frac{\sigma^2}{2\mu} (1 - e^{-2\mu h}). \end{aligned}$$

Moreover, by the generalized Itô isometry,

$$\begin{aligned} \mathbb{E}[\tilde{\Delta}O \cdot \Delta W] &= \mathbb{E}\left[\sigma e^{-\mu h} \int_0^h e^{\mu s} dW_s \int_0^h dW_s\right] \\ &= \sigma e^{-\mu h} \int_0^h e^{\mu s} ds = \frac{\sigma}{\mu} (1 - e^{-\mu h}). \end{aligned}$$

On the other hand, integrating the SODE (14.3) directly for the solution O_t over the interval $[0, h]$ gives

$$I(O) := \int_0^h O_s ds = \frac{1}{\mu} \left(\sigma \int_0^h dW_s - \int_0^h dO_s \right)$$

$$\begin{aligned}
&= \frac{\sigma}{\mu} \Delta W - \frac{1}{\mu} (O_h - O_0) \\
&= \frac{\sigma}{\mu} \Delta W - \frac{1}{\mu} \tilde{\Delta} O + \frac{1}{\mu} (1 - e^{-\mu h}) O_0.
\end{aligned}$$

Thus $I(O)$ is Gaussian distributed with mean $\mathbb{E}[I(O)] = 0$.

It is known that O_t is also given explicitly by

$$O_t = \sigma e^{-\mu t} \int_{-\infty}^t e^{\mu s} dW_s,$$

which requires W_t to be a two-sided Wiener process, i.e., defined for all $t \in \mathbb{R}$. Hence

$$O_0 = \sigma \int_{-\infty}^0 e^{\mu s} dW_s$$

is $\mathcal{N}(0, \frac{\sigma^2}{2\mu})$ -distributed by the properties of the Itô integral. In fact, by the Itô isometry,

$$\begin{aligned}
\mathbb{E}[O_0] &= \sigma \mathbb{E}\left[\int_{-\infty}^0 e^{\mu s} dW_s\right] = 0 \\
\mathbb{E}[O_0^2] &= \sigma^2 \mathbb{E}\left[\left(\int_{-\infty}^0 e^{\mu s} dW_s\right)^2\right] = \sigma^2 \int_{-\infty}^0 e^{2\mu s} ds = \frac{\sigma^2}{2\mu}.
\end{aligned}$$

Since O_0 depends on the Wiener process only up to time $t = 0$, it is clearly independent of ΔW and $\tilde{\Delta} O$. Hence the variance of $I(O_0)$ is given by

$$\begin{aligned}
\text{Var}[I(O_0)] &= \mathbb{E}[I(O_0)^2] - (\mathbb{E}[I(O_0)])^2 \\
&= \frac{1}{\mu^2} \left\{ \sigma^2 \mathbb{E}[(\Delta W)^2] - 2\sigma \mathbb{E}[\Delta W \tilde{\Delta} O] + \mathbb{E}[(\tilde{\Delta} O)^2] + (1 - e^{-\mu h})^2 \mathbb{E}[O_0^2] \right\} \\
&= \frac{\sigma^2 h}{\mu^2} - \frac{2\sigma^2}{\mu^3} (1 - e^{-\mu h}) + \frac{\sigma^2}{2\mu^3} (1 - e^{-2\mu h}) + \frac{\sigma^2}{2\mu^3} (1 - e^{-\mu h})^2 \\
&= \frac{\sigma^2}{\mu^3} (\mu h - 1 + e^{-\mu h}) =: \varsigma^2.
\end{aligned}$$

14.4 Fractional Brownian Motion

A fractional Brownian motion (fBm) is a self-similar process with long range dependence given by

$$B_H(t) := \frac{1}{\Gamma(H+1/2)} \left(\int_{-\infty}^0 ((t-s)^{H-1/2} - (-s)^{H-1/2}) dW_s + \int_0^t (t-s)^{H-1/2} dW_s \right), \quad (14.5)$$

where W_t is a standard Brownian motion, i.e., a Wiener process, H is the Hurst parameter satisfying $0 < H < 1$ and Γ is the Gamma function

$$\Gamma(b) := \int_0^\infty x^{b-1} \exp(-x) dx.$$

When $H = 1/2$, the fBm (14.5) reduces to $W(t)$, a Wiener process.

The fBm is a Gaussian process with stationary increments $B_H(t) - B_H(s) \sim B_H(t-s)$ satisfying

$$\mathbb{E}[B_H(t) - B_H(s)] = 0, \quad \mathbb{E}[(B_H(t) - B_H(s))^2] = |t-s|^{2H} \sigma^2$$

and the covariance function

$$\mathbb{E}[B_H(t) B_H(s)] = \frac{1}{2} (t^{2H} + s^{2H} - |t-s|^{2H}) \sigma^2.$$

(In what follows often $\sigma = 1$ for simplicity).

Various exact and approximate methods have been used in the literature to generate fBm. Three of them, the Cholesky method, the FFT method and the RMD method will be considered here. After generating random Gaussian increments by these three methods, $B_H(t)$ is obtained by summing up the increments appropriately. These methods require the covariance matrix Σ_{11} of $B_H(t)$, which is given by

$$(\Sigma_{11})_{i,j} = \text{Cov}(X_i, X_j) = \mathbb{E}[X_i X_j],$$

where

$$X_i = B_H(t_i) - B_H(t_{i-1}), \quad X_j = B_H(t_j) - B_H(t_{j-1})$$

for $i, j = 1, \dots, n$.

For simplicity, an equidistant partition on a given interval $[t_0, t_n]$ with $h = t_i - t_{i-1}$ for $i = 1, \dots, n$, will be used here. These components are evaluated in Appendix C. In particular, the covariance matrix Σ_{11} is calculated to be

$$\Sigma_{11} = \begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_1 & a_0 & a_1 & \cdots & a_{n-2} \\ a_2 & a_1 & a_0 & \cdots & a_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n-1} & a_{n-2} & a_{n-3} & \cdots & a_0 \end{pmatrix}, \quad (14.6)$$

where a_0 is given by (C.3) and the other a_i are by (C.4) in Appendix C.

The Cholesky Method

The matrix (14.6) is always positive definite due to the local non-determinism of the fBm. It thus has a Cholesky decomposition given by

$$\Sigma_{11} = Q \Lambda Q^\top = (Q \sqrt{\Lambda}) (\sqrt{\Lambda} Q^\top) = (Q \sqrt{\Lambda}) (Q \sqrt{\Lambda})^\top = \tilde{Q} \tilde{Q}^\top,$$

where Q is a lower triangular matrix and Λ is a matrix of eigenvalues, i.e., $\Lambda = \text{diag}(\lambda_0, \dots, \lambda_{n-1})$ with the eigenvalues λ_i for $i = 0, \dots, n - 1$.

Suppose that the matrix \tilde{Q} is given by

$$\tilde{Q} = \begin{pmatrix} q_{1,1} & 0 & \cdots & 0 \\ q_{2,1} & q_{2,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ q_{n,1} & q_{n,2} & \cdots & q_{n,n} \end{pmatrix}$$

and that a vector $Z = (Z_1, Z_2, \dots, Z_n)^\top$ has independent and identically distributed (i.i.d.) $\mathcal{N}(0, 1)$ components. The product of \tilde{Q} and Z ,

$$\begin{pmatrix} q_{1,1} & 0 & \cdots & 0 \\ q_{2,1} & q_{2,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ q_{n,1} & q_{n,2} & \cdots & q_{n,n} \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{pmatrix} = \begin{pmatrix} q_{1,1} Z_1 \\ \sum_{i=1}^2 q_{2,i} Z_i \\ \vdots \\ \sum_{i=1}^n q_{n,i} Z_i \end{pmatrix} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$$

gives the increments of fBm, i.e., $X_i = B_H(t_i) - B_H(t_{i-1})$ for $i = 1, \dots, n$.

Take $B_H(t_0) = 0$. Then $B_H(t_1) - B_H(t_0) = X_1$ gives $B_H(t_1) = X_1$. Similarly, $B_H(t_2) - B_H(t_1) = X_2$ gives $B_H(t_2) = X_2 + B_H(t_1) = \sum_{j=1}^2 X_j$. Continuing in this way gives

$$B_H(t_m) = \sum_{j=1}^m X_j. \quad (14.7)$$

The Fast Fourier Transformation (FFT)

The FFT method is similar to the Cholesky decomposition in that it also determines the square root of the covariance matrix Σ_{11} . It can, however, only be applied to matrices of size $2^s \times 2^s$ where $s \in \mathbb{N}$, so it will be assumed that $n = 2^s$ for some such s .

A circulant matrix C_{11} of size $2n = 2^{s+1}$ is defined by

$$C_{11} = \begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} & b & a_{n-1} & a_{n-2} & \cdots & a_1 \\ a_1 & a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} & b & a_{n-1} & \cdots & a_2 \\ a_2 & a_1 & a_0 & \cdots & a_{n-3} & a_{n-2} & a_{n-1} & b & \cdots & a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n-1} & a_{n-2} & a_{n-3} & \cdots & a_0 & a_1 & a_2 & a_3 & \cdots & b \\ b & a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & b & a_{n-1} & \cdots & a_2 & a_1 & a_0 & a_1 & \cdots & a_{n-2} \\ a_{n-2} & a_{n-1} & b & \cdots & a_3 & a_2 & a_1 & a_0 & \cdots & a_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & a_3 & \cdots & b & a_{n-1} & a_{n-2} & a_{n-3} & \cdots & a_0 \end{pmatrix},$$

where b is arbitrary.

A random vector $X \sim \mathcal{N}(0, \Sigma_{11})$ is then determined as follows. First define $Y = U \Lambda^{1/2} U^* \mathbf{Z}$ where $\Lambda^{1/2} = \text{diag}\{\lambda_0^{1/2}, \dots, \lambda_{2n-1}^{1/2}\}$ and $\mathbf{Z} = (Z_0, \dots, Z_{2n-1})^\top$ is a vector of independent $\mathcal{N}(0, 1)$ random variables. Then $Y \sim \mathcal{N}(0, C_{11})$ because U is unitary. Finally, the subvector $X = (Y_0, \dots, Y_{n-1})$ has the desired property, i.e., $X \sim \mathcal{N}(0, \Sigma_{11})$. The components of X are then summed as in (14.7) in the Cholesky decomposition case to give the fBm at time t_m .

Step 1: The fast Fourier transform is performed on the elements of the first row of the matrix C_{11} in order to determine the eigenvalues Λ :

$$\lambda_k = \sum_{j=0}^{2n-1} a_k \exp\left(-\frac{2\pi i j k}{2n}\right),$$

for $k = 0, \dots, 2n - 1$.

Step 2: Generate random numbers $Z_j \sim \mathcal{N}(0, 1)$ and determine $W = \Lambda^{1/2} U^* \mathbf{Z}$, where $U^* \mathbf{Z}$ can be generated as follows

- Generate two standard normal random variables for $U^* Z_0$ and $U^* Z_n$, the first and n th elements.
- For $1 \leq j < n$, generate two independent standard normal random variables $V_j^{(1)}$ and $V_j^{(2)}$ and calculate

$$U^* Z_j = \frac{1}{\sqrt{2}} (V_j^{(1)} + i V_j^{(2)}), \quad U^* Z_{2n-j} = \frac{1}{\sqrt{2}} (V_j^{(1)} - i V_j^{(2)}).$$

Step 3: Generate Y by the fast Fourier transform and X can be obtained by picking up n elements from the top of Y :

$$X\left(\frac{k}{n}\right) = \sum_{j=0}^{2n-1} \frac{1}{\sqrt{2n}} W_j \exp\left(-\frac{2\pi i j k}{2n}\right),$$

for $k = 0, \dots, n - 1$.

The Random Midpoint Displacement Method

The random midpoint displacement (RMD) method is similar to the Lévy construction of a Wiener process. It can be used to construct a finer sample path of an fBm between two given values. Consider the increment $B_H(t_2) - B_H(t_1)$ of an fBm for $0 \leq t_1 < t_2 \leq 1$ with $h = t_2 - t_1$. This increment has mean 0 and variance

$$\text{Var}[B_H(t_2) - B_H(t_1)] = |t_2 - t_1|^{2H} \sigma^2 = h^{2H} \sigma^2. \quad (14.8)$$

The value $B_H(t_{1.5})$ of the fBm at the midpoint $t_{1.5} = \frac{1}{2}(t_1 + t_2)$ is then given as the sum of the average of $B_H(t_1)$ and $B_H(t_2)$ plus an independent Gaussian correction ζ_1 , i.e.,

$$B_H(t_{1.5}) = \frac{1}{2} (B_H(t_2) + B_H(t_1)) + \zeta_1, \quad (14.9)$$

where ζ_1 has mean 0 and variance σ_1^2 .

To determine the appropriate value of σ_1 subtract $B_H(t_1)$ from both sides of the Eq. (14.9) and take its variance:

$$\begin{aligned} \text{Var}[B_H(t_{1.5}) - B_H(t_1)] &= \text{Var}\left[\frac{1}{2}(B_H(t_2) - B_H(t_1)) + \zeta_1\right] \\ &= \text{Var}\left[\frac{1}{2}(B_H(t_2) - B_H(t_1))\right] + \text{Var}[\zeta_1] \\ &= \frac{1}{4}h^{2H} \sigma^2 + \sigma_1^2. \end{aligned} \quad (14.10)$$

By (14.8), the left-hand side of the Eq. (14.10) is $(1/2)^{2H}h^{2H}\sigma^2$, which implies

$$\sigma_1^2 = \left(\left(\frac{1}{2}\right)^{2H} - \frac{1}{4}\right)h^{2H} \sigma^2.$$

This can be repeated by halving the subinterval with the variance of the midpoint displacement ζ_n at the n th step as

$$\sigma_n^2 = \frac{1}{(2^n)^{2H}} (1 - 2^{2H-2}) h^{2H} \sigma^2.$$

The Cholesky decomposition or the FFT can be used to generate some initial values of the fBm and then the RMD method can be applied to generate intermediate points on each subinterval to give the desired resolution. In general, the RMD method requires order $\mathcal{O}(n)$ computational costs when n points are needed. It can be faster than using the Cholesky decomposition or FFT method alone.

14.4.1 Riemann Integral of an fBm

The Riemann integral $\int_{t_0}^{t_m} B_H(s) \, ds$ of an fBm can be generated together with the process $B_H(t)$ at time instants $t_0 < t_1 < \dots < t_n$ using the Cholesky decomposition of the full covariance matrix Σ (see (C.1) in Appendix C) for the increments of the process and integral. Let $\Sigma = \tilde{Q} \tilde{Q}^\top$, where \tilde{Q} is the lower triangular matrix

$$\tilde{Q} = \begin{pmatrix} q_{1,1} & 0 & \dots & 0 \\ q_{2,1} & q_{2,2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ q_{2n,1} & q_{2n,2} & \dots & q_{2n,2n} \end{pmatrix}.$$

Prepare a vector $Z = (Z_1, Z_2, \dots, Z_{2n})^\top$ which has i.i.d. $\mathcal{N}(0, 1)$ components. Then,

$$\begin{pmatrix} q_{1,1} & 0 & \dots & 0 \\ q_{2,1} & q_{2,2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ q_{2n,1} & q_{2n,2} & \dots & q_{2n,2n} \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_{2n} \end{pmatrix} = \begin{pmatrix} q_{1,1} Z_1 \\ \sum_{i=1}^2 q_{2,i} Z_i \\ \vdots \\ \sum_{i=1}^{2n} q_{2n,i} Z_i \end{pmatrix} = \begin{pmatrix} X \\ Y \end{pmatrix},$$

where X and Y are given by

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} q_{1,1} Z_1 \\ \sum_{i=1}^2 q_{2,i} Z_i \\ \vdots \\ \sum_{i=1}^n q_{n,i} Z_i \end{pmatrix}, \quad Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n+1} q_{n+1,i} Z_i \\ \sum_{i=1}^{n+2} q_{n+2,i} Z_i \\ \vdots \\ \sum_{i=1}^{2n} q_{2n,i} Z_i \end{pmatrix},$$

which correspond to the increments of fBm $X_i = B_H(t_i) - B_H(t_{i-1})$ and the integral $Y_i = \int_{t_{i-1}}^{t_i} (B_H(s) - B_H(t_{i-1})) \, ds$.

Take $B_H(t_0) = 0$. Then $B_H(t_1) - B_H(t_0) = X_1$ gives $X_1 = B_H(t_1)$. Similarly, $B_H(t_2) - B_H(t_1) = X_2$ gives

$$B_H(t_2) = X_2 + B_H(t_1) = \sum_{i=1}^2 X_i.$$

Finally,

$$B_H(t_m) = \sum_{i=1}^m X_i.$$

In the same way one obtains

$$Y_m = \int_{t_{m-1}}^{t_m} (B_H(s) - B_H(t_{m-1})) \, ds = \int_{t_{m-1}}^{t_m} B_H(s) \, ds - B_H(t_{m-1}) \int_{t_{m-1}}^{t_m} \, ds,$$

and hence

$$\int_{t_{m-1}}^{t_m} B_H(s) \, ds = Y_m + B_H(t_{m-1})h.$$

The Riemann integral of $B_H(t)$ on $[t_0, t_m]$ is thus given by

$$\begin{aligned} \int_{t_0}^{t_m} B_H(s) \, ds &= \int_{t_0}^{t_1} B_H(s) \, ds + \cdots + \int_{t_{m-1}}^{t_m} B_H(s) \, ds \\ &= (Y_1 + B_H(t_0)h) + \cdots + (Y_m + B_H(t_{m-1})h) \\ &= \sum_{i=1}^m (Y_i + B_H(t_{i-1})h). \end{aligned}$$

14.4.2 Riemann Sums Approximation

The Riemann integral of the fBm

$$I_{t_{i-1}}(B_H) := \int_{t_{i-1}}^{t_i} B_H(s) \, ds,$$

can be approximated by the Riemann sums

$$I_{t_{i-1}}(B_H) \approx \delta \sum_{k=1}^m B_H(t_{i-1} + k\delta),$$

where $\delta = h/m$ for $h = t_i - t_{i-1}$ and m sufficiently large to achieve a desired accuracy. This can be quite efficient when the RMD method is used to generate the

values $B_H(t_{i-1} + k\delta)$ between $B_H(t_{i-1})$ and $B_H(t_i)$. To achieve order h^2 convergence of a numerical scheme, the error in approximating the integral should be of order h^3 , i.e.,

$$\left| \int_{t_{i-1}}^{t_i} B_H(s) ds - \delta \sum_{k=1}^m B_H(t_{i-1} + k\delta) \right| \leq K h^3, \quad (14.11)$$

for some constant $K > 0$. By the Hölder continuity of the sample paths of an fBm, the left hand side of the inequality Eq.(14.11) can be evaluated as

$$\begin{aligned} & \left| \int_{t_{i-1}}^{t_i} B_H(s) ds - \delta \sum_{k=1}^m B_H(t_{i-1} + k\delta) \right| \\ & \leq \sum_{k=1}^m \int_{t_{i-1}+(k-1)\delta}^{t_{i-1}+k\delta} |B_H(s) - B_H(t_{i-1} + k\delta)| ds \\ & \leq \sum_{k=1}^m \int_{t_{i-1}+(k-1)\delta}^{t_{i-1}+k\delta} C_\varepsilon |s - (t_{i-1} + k\delta)|^{H-\varepsilon} ds \\ & = -\frac{C_\varepsilon}{H+1-\varepsilon} \sum_{k=1}^m ((t_{i-1} + k\delta) - s)^{H+1-\varepsilon} \Big|_{t_{i-1}+(k-1)\delta}^{t_{i-1}+k\delta} \\ & = \frac{C_\varepsilon}{H+1-\varepsilon} \sum_{k=1}^m \delta^{H+1-\varepsilon} = \frac{C_\varepsilon}{H+1-\varepsilon} (m\delta) \delta^{H-\varepsilon} = \frac{C_\varepsilon}{H+1-\varepsilon} h \delta^{H-\varepsilon}. \end{aligned}$$

This means that the order h^2 convergence can be attained if the following inequality holds

$$\frac{C_\varepsilon}{H+1-\varepsilon} \delta^{H-\varepsilon} \leq Ch^2.$$

The condition $\delta^{H-\varepsilon} \leq K(H+1-\varepsilon)h^2$, where $K = C/C_\varepsilon$, is thus necessary for order h^2 convergence.

14.4.3 Comparison of Computational Costs

The computational costs for generating the fBm $B_H(t)$ and its Riemann integral $I(B_H)$ on the unit interval $[0, 1]$ will be compared here for the following methods:

1. Solving the full covariance matrix Σ by Cholesky decomposition and generating $B_H(t)$ and $I(B_H)$ simultaneously.
2. Solving Σ_{11} by Cholesky decomposition, generating intermediate points by the RMD method and taking the Riemann sums for $I(B_H)$.

Table 14.1 Computational costs to generate 1 sample path

h	2^{-4}	2^{-6}	2^{-8}	2^{-10}
Method 1	0.007063	0.020983	0.514968	23.880383
Method 2	0.004789	0.025458	0.225578	12.767898
Method 3	0.004829	0.009526	0.068878	0.851457
Method 4	0.004082	0.010707	0.034688	0.740504

Table 14.2 Computational costs to generate 1000 sample path

h	2^{-4}	2^{-6}	2^{-8}	2^{-10}
Method 1	1.149970	2.806160	18.66572	170.5010
Method 2	2.882757	9.091249	52.02268	693.3296
Method 3	3.587168	9.020846	51.34058	679.4948
Method 4	3.930499	10.310124	51.16574	655.2259

3. Solving Σ_{11} by the FFT, generating intermediate points by the RMD method and taking the Riemann sums for $I(B_H)$.
4. Generating $B_H(t)$ with small step size by the RMD method and taking the Riemann sums for $I(B_H)$.

The step size for $B_H(t)$ are chosen to be $h = 2^{-4}, 2^{-6}, 2^{-8}$ and 2^{-10} and the step size for $I(B_H)$ for the methods 2–4 are set to $\delta = 2^{-8}, 2^{-12}, 2^{-16}$ and 2^{-20} , i.e., $\delta = h^2$.

Computational costs to generate 1 sample path as well as 1000 paths by the above four methods with different step sizes h are listed in Tables 14.1 and 14.2 and also illustrated in Fig. 14.1. The number of subintervals of $[0, 1]$ in each column is given by $1/h$.

In general, the FFT and the RMD method have computational costs of order $\mathcal{O}(n \log(n))$ and $\mathcal{O}(n)$, respectively, while the Cholesky decomposition has $\mathcal{O}(n^3)$ [81, 112]. Table 14.1 shows the FFT and the RMD method have a clear advantage from computational point. In particular, they generate sample paths much faster than the Cholesky decomposition. However, for additional paths it is necessary to repeat the whole process by these two methods, whereas the Cholesky decomposition reuses the square root of the covariance matrix that has already been obtained. If the number of sample paths to be generated is large, then the total computational costs by Cholesky decomposition (the method 2) are almost the same as the FFT (the method 3) and the RMD method (method 4).

The computational costs to generate one sample path by the method 1 also becomes larger when the step size h becomes smaller, i.e., as the number of intervals increases. However, the square root of the covariance matrix in the method 1 can be reused. Table 14.2 and Fig. 14.1 show that it requires the least computational cost when many sample paths are required.

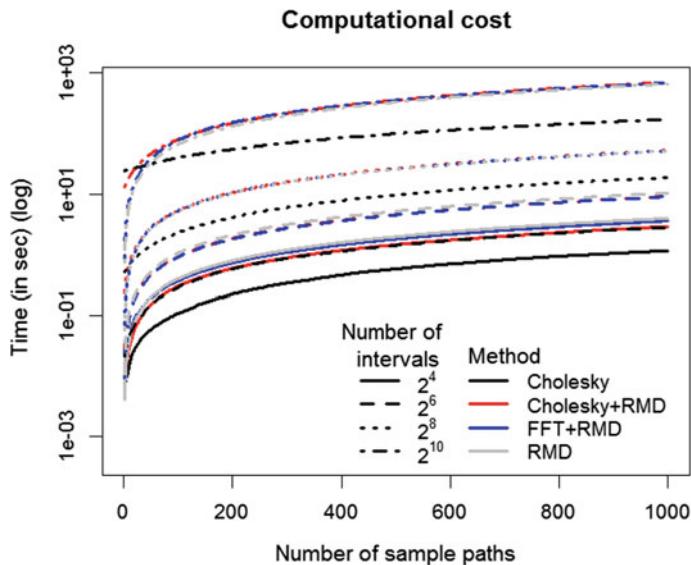


Fig. 14.1 Comparison of computational costs to generate $B_H(t)$ and $I(B_H)$

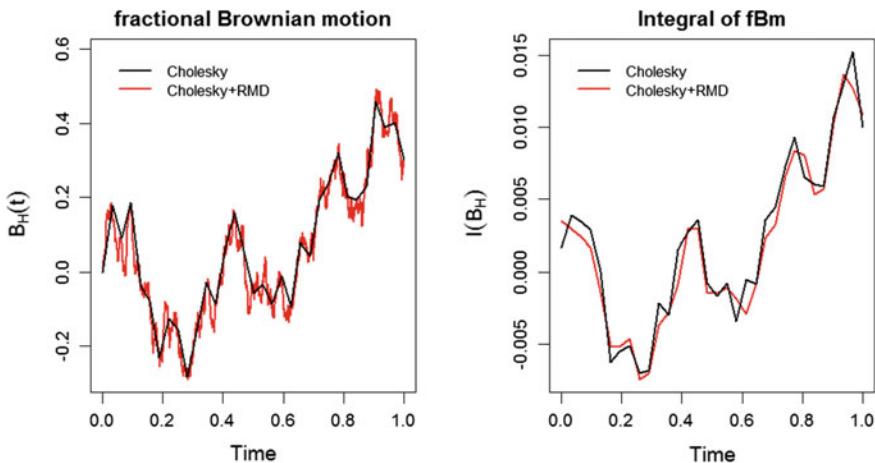


Fig. 14.2 Sample paths by the method 1 and the method 2 with $H = 0.6$, $h = 2^{-5}$ and $\delta = h^2$

Figure 14.2 illustrates the sample paths of $B_H(t)$ and $I(B_H)$ generated by the method 1 and the method 2.

14.5 Integrals of Compound Poisson Processes

The numerical schemes for affine-RODEs considered in Chap. 11 require the simulation of noise process η_t and its integrals

$$I_{(j_1, \dots, j_i), n} = \int_{t_n}^{t_{n+1}} \dots \int_{t_n}^{s_{i-1}} \eta_s^{j_i} \dots \eta_s^{j_1} ds_i \dots ds_1 \quad (14.12)$$

on each discretised subinterval $[t_n, t_{n+1}]$, where the noise process η_t^j could be a Wiener process or Ornstein–Uhlenbeck process with continuous sample paths or a compound Poisson process with piecewise constant sample paths. In general, such integrals can be calculated using Riemann sums on a much finer partition of the discretisation subinterval so that the error is dominated by local discretisation error of the scheme itself.

The Poisson process, or more generally the compound Poisson process, is a representative example of a noise process with piecewise continuous sample path. A Poisson process N_t counts the number of events, which are independent of each other, that occur during a given time interval $[0, t]$. It has the probabilities

$$\mathbb{P}\{N_t = k\} = (\lambda t)^k \frac{e^{-\lambda t}}{k!}, \quad k = 0, 1, 2, \dots,$$

where λ is a positive real number. Let T_1 be the time when the first event occurs and let N_t the number of events occurred until time t . The distribution function of T_1 is

$$F_{T_1}(t) = \mathbb{P}\{T_1 \leq t\} = \mathbb{P}\{N_t \geq 1\} = \sum_{i=1}^{\infty} (\lambda t)^k \frac{e^{-\lambda t}}{k!} = 1 - e^{-\lambda t},$$

which is continuous and monotonically increasing, hence invertible. Thus T_1 can be simulated as $T_1 = F_{T_1}^{-1}(t) = \log(1 - U)/\lambda = \log(V)/\lambda$, where U and $V = 1 - U$ are uniformly distributed on $[0, 1]$.

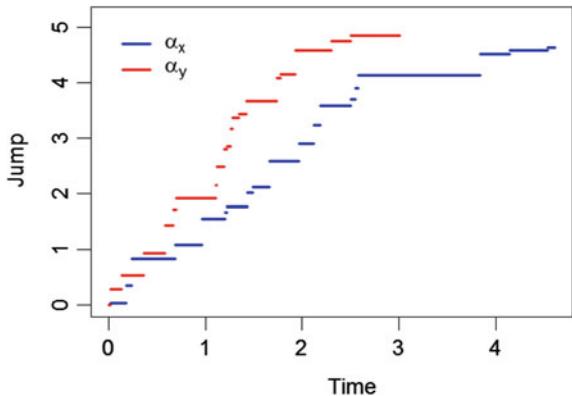
The compound Poisson process is a Poisson process with jumps of random magnitude, which satisfy a distribution function f , for example uniformly distributed on some bounded interval. It is defined by

$$P_t = \sum_{i=1}^{N_t} Y_i, \quad (14.13)$$

where Y_i is the jump magnitude at i -th jump. When the jump magnitude $Y_i \equiv 1$ for all i , the compound Poisson process is just a Poisson process. The compound Poisson process (14.13) is a left-continuous step function (see Fig. 14.3). Let

$$P_t^j = p_i^j \quad \text{for } t_i \leq t < t_{i+1}.$$

Fig. 14.3 Typical trajectories of the compound Poisson processes with $\lambda = 5$ and jumps uniformly distributed on $[0, 0.5]$



For a scalar case with $\eta_t = P_t$, if there is no jump in the interval $[t_n, t_{n+1})$, then the integration of (14.12) can be given as direct integration

$$I_{(j),n} = \int_{t_n}^{t_{n+1}} P_s^j ds = \int_{t_n}^{t_{n+1}} p_i^j ds = p_i^j h_n.$$

On the other hand if $t_n < T_i < t_{n+1}$ and there are no other jumps in this interval, then

$$I_{(j),n} = \int_{t_n}^{t_{n+1}} P_s^j ds = \int_{t_n}^{T_i} p_{i-1}^j ds + \int_{T_i}^{t_{n+1}} p_i^j ds = p_{i-1}^j(T_i - t_n) + p_i^j(t_{n+1} - T_i).$$

In case there are more jumps on $[t_n, t_{n+1})$, the given interval is divided into an appropriate number of subintervals and the result summed.

Higher order integrals can be derived in the same manner. Assume that $\eta_t^{j_1} = P_t^{j_1}$ and $\eta_t^{j_2} = P_t^{j_2}$ are two independent compound Poisson processes with jumps (only) at T_{i_1} and T_{i_2} on $[t_n, t_{n+1})$ and values at the i_1 -th and i_2 -th jumps given by $p_{i_1}^{j_1}$ and $p_{i_2}^{j_2}$, respectively. For simplicity, suppose that $t_n < T_{i_1} < T_{i_2} < t_{n+1}$. Then

$$\begin{aligned} I_{(j_1, j_2),n} &= \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_1} P_{s_1}^{j_1} P_{s_2}^{j_2} ds_2 ds_1 \\ &= \int_{t_n}^{T_{i_1}} \int_{t_n}^{s_1} p_{i_1-1}^{j_1} p_{i_2-1}^{j_2} ds_2 ds_1 + \int_{T_{i_1}}^{T_{i_2}} \int_{T_{i_1}}^{s_1} p_{i_1}^{j_1} p_{i_2-1}^{j_2} ds_2 ds_1 \\ &\quad + \int_{T_{i_2}}^{t_{n+1}} \int_{T_{i_2}}^{s_1} p_{i_1}^{j_1} p_{i_2}^{j_2} ds_2 ds_1 \\ &= \frac{1}{2} p_{i_1-1}^{j_1} p_{i_2-1}^{j_2} (T_{i_1} - t_n)^2 + \frac{1}{2} p_{i_1}^{j_1} p_{i_2-1}^{j_2} (T_{i_2} - T_{i_1})^2 + \frac{1}{2} p_{i_1}^{j_1} p_{i_2}^{j_2} (t_{n+1} - T_{i_2})^2. \end{aligned}$$

On the other hand, if $\eta_t^{j_1}$ is a stochastic process with continuous sample paths such as a Wiener process W_t and $\eta_t^{j_2} = P_t^{j_2}$ is an independent compound Poisson process which has a single jump at time T_{i_2} on $[t_n, t_{n+1})$ with value p_{i_2} , then

$$\begin{aligned} I_{(j_1, j_2), n} &= \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_1} W_{s_1} P_{s_2} ds_2 ds_1 \\ &= p_{i_2-1} \int_{t_n}^{T_{i_2}} W_{s_1}(s_1 - t_n) ds_1 + p_{i_2} \int_{T_{i_2}}^{t_{n+1}} W_{s_1}(s_1 - T_{i_2}) ds_1. \end{aligned}$$

14.6 Endnotes

See Kloeden and Platen [91] for simulating integrals of Wiener processes and Hanson [62] for compound Poisson processes. See also Asai, Herrmann and Kloeden [12] and Asai and Kloeden [11]. Rydén and Wiktorsson [118] and Wiktorsson [138] discuss the simulation of iterated Itô integrals.

SODEs with Markovian switching are considered in Mao and Yuan [100].

Fractional Brownian motion (fBm) was first mentioned by Kolmogorov and later extensively investigated by Mandelbrot and Van Ness [99]. Exact and approximate methods for generating fBm are given by Couerjolly [33] and Wood and Chan [142]. The RMD method for fBm is discussed in Norros, Mannersalo and Wang [109] and Peitgen and Saupe [112]. The idea of using the FFT for fBm was first introduced by Davies and Harte [36] and later it was generalised by Dietrich and Newsam [39] and Wood and Chan [142]. The section on fBM is based on Asai [8]. See Appendix C for the determination of the covariance matrix of an fBm and its Riemann integral.

Part IV

Random Ordinary Differential Equations

in the Life Sciences

Chapter 15

Comparative Simulations of Biological Systems

The numerical schemes introduced in the previous chapters will now be tested with representative systems in cancer modeling and epidemiology as well as a toggle switch model and one on the formation of shell patterns. They include both non-stiff and stiff systems. These systems are based on ODE models in the literature and in each case a single parameter will be randomised by the insertion of a scalar Itô diffusion process Y_t in a globally bounded function.

15.1 Tumor Inhibition Model

As an example of non-stiff system, consider the cancer model from Wodarz and Komarova [139] (p. 154), in which the angiogenesis inhibition prevents tumor cell division,

$$\begin{aligned}\frac{dC}{dt} &= \left(\frac{rC}{\varepsilon C + 1} \right) \left(\frac{P}{I + 1} \right) - \mu C \\ \frac{dP}{dt} &= a_P C - b_P P \\ \frac{dI}{dt} &= \zeta + a_I C - b_I I\end{aligned}$$

where C , P , I are the population of cancer cells, promoters of cancer cell growth and inhibitors, respectively. The parameters are all strictly positive except the inhibitor input ζ is nonnegative. Note that the coefficient functions and their partial derivatives satisfy global Lipschitz bounds on the biologically relevant region \mathbb{R}_+^3 .

Introduce the following stochastic process

$$\xi(Y_t) := \xi_0 \left(1 - 2\nu \frac{Y_t}{1 + Y_t^2} \right),$$

where ζ_0 and ν are positive constants with $\nu \in (0, 1]$ and Y_t is a given stochastic process, which will be a Wiener process W_t here. The noise here tends to peak around $\zeta_0(1 \pm \nu)$, and is thus suitable for a noisy switching scenario.

Replacing ζ by the bounded random process $\zeta(W_t)$ gives the system of RODEs:

$$\begin{aligned}\frac{dC}{dt} &= \left(\frac{rC}{\varepsilon C + 1}\right) \left(\frac{P}{I + 1}\right) - \mu C \\ \frac{dP}{dt} &= a_P C - b_P P \\ \frac{dI}{dt} &= \zeta(W_t) + a_I C - b_I I\end{aligned}\tag{15.1}$$

or the equivalent vector Itô SODE

$$d \begin{pmatrix} C_t \\ P_t \\ I_t \\ Y_t \end{pmatrix} = \begin{pmatrix} \left(\frac{rC_t}{\varepsilon C_t + 1}\right) \left(\frac{P_t}{I_t + 1}\right) - \mu C_t \\ a_P C_t - b_P P_t \\ \zeta(Y_t) + a_I C_t - b_I I_t \\ 0 \end{pmatrix} dt + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} dW_t.$$

In the following simulation examples, the time step size is fixed at $h = 0.2$ for the averaged schemes and $\delta = h^2$ for other schemes. The spread of the noise ν is fixed to be 1. The parameters in the tumor inhibition model are set to $r = 1$, $\mu = 0.1$, $a_P = 4.5$, $b_P = 0.11$, $a_I = 0.2$, $b_I = 0.01$, $\varepsilon = 0.34$, $\zeta_0 = 4$. The initial population of cancer cells is set to $C_0 = 35$ and the other initial values are obtained by a quasi-stationary approach, where the initial value of promoters P_0 is given by $P_0 = a_P/b_P C_0$ and the initial value of one of inhibitors I_0 is given by $I_0 = (\zeta_0 + a_I C_0)/b_I$.

Solutions of the system (15.1) will be approximated by

1. *the explicit Euler scheme;*
2. *the derivative-free explicit order 1.5 scheme;*
3. *the Adams–Bashforth-2 scheme;*
4. *the explicit averaged Euler scheme.*

Figure 15.1 shows a typical sample path of each scheme (four subfigures on the left) and a comparison of sample solution trajectories (subfigure on the top right) and calculation times of different schemes (subfigure on the bottom right). On the left 4 figures, the black (solid) line, the red (dashed) line and the light green (dotted) line are the population of the cancer cells, the promoters and the inhibitors, respectively. The top right subfigure shows the comparison of the solutions by the explicit Euler scheme, the derivative-free explicit order 1.5 scheme, the Adams–Bashforth-2 scheme and the explicit averaged Euler scheme. The bar plots on the bottom right show the sum of the computational times for 100 simulations of each scheme. All solution curves appear very similar in the Fig. 15.1. However, the computational time for the averaged scheme is much less than the ones by the other 3 schemes.

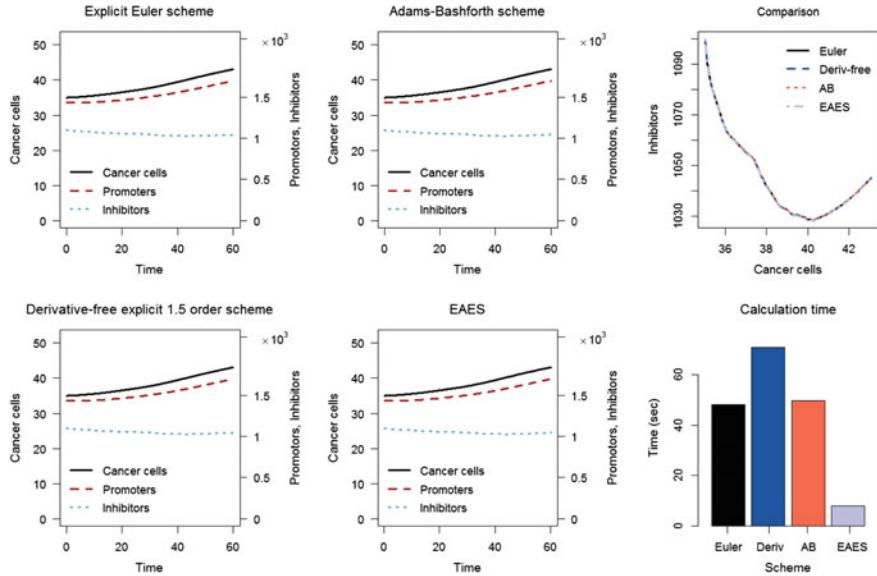


Fig. 15.1 A simulation example of the tumor-inhibition model: the step size $h = 0.2$ for the averaged scheme and $\delta = h^2$ for other schemes. $r = 1$, $\mu = 0.1$, $a_P = 4.5$, $b_P = 0.11$, $a_I = 0.2$, $b_I = 0.01$, $\varepsilon = 0.34$, $\zeta_0 = 4$ and $C_0 = 35$. The solid black line the population of the cancer cells, the dashed red line the population of the promoters and the dotted light green line the population of the inhibitors

15.2 Population Dynamics

Consider the following non-stiff nonlinear scalar RODE based on a logistic population model

$$\frac{dX_t}{dt} = r X_t(k - X_t) \cos(Y_t), \quad (15.2)$$

where Y_t is an Ornstein–Uhlenbeck process satisfying the SODE

$$dY_t = (\theta_1 - \theta_2 Y_t) dt + \theta_3 dW(t).$$

In this section various 2-step explicit SLMMs are applied to the RODE (15.2). The stochastic processes Y_t will be generated exactly. The additional initial conditions at time t_0 are provided by explicit and implicit RODE-Taylor schemes of the same order. In this section, $f(x_n, Y_{t_n})$ is denoted by f_n for simplicity.

Explicit SLMMs of order 1.0, 1.5 and 2.0 to be applied are as follows

$$(Ex1.0) \quad x_n = x_{n-1} + \frac{h}{2} (f_{n-1} + f_{n-2}),$$

$$(Ex1.5) \quad x_n = x_{n-1} + h f_{n-2} + L^1 f_{n-1} I_{(1,0),n-1} + L^1 f_{n-2} I_{(1),n-2} h \\ + L^0 f_{n-1} I_{(0,0),n-1} + L^0 f_{n-2} I_{(0),n-2} h,$$

$$(Ex2.0) \quad x_n = \frac{1}{2} x_{n-1} + \frac{1}{2} x_{n-2} + h \left(f_{n-1} + \frac{1}{2} f_{n-2} \right) \\ + L^1 f_{n-1} I_{(1,0),n-1} + \frac{1}{2} L^1 f_{n-2} I_{(1,0),n-2} + L^0 f_{n-1} I_{(0,0),n-1} \\ + \frac{1}{2} L^0 f_{n-2} I_{(0,0),n-2} + L^1 L^1 f_{n-1} I_{(1,1,0),n-1} + \frac{1}{2} L^1 L^1 f_{n-2} I_{(1,1,0),n-2},$$

where

$$I_{\alpha,n} = \int_{t_n}^{t_{n+1}} \cdots \int_{t_2}^{t_3} dW_{s_1}^{j_1} \cdots dW_{s_l}^{j_l}.$$

In particular, the Adams–Bashforth scheme reads

$$x_n = x_{n-1} + \frac{h}{2} (3 f_{n-1} - f_{n-2}),$$

which has order 1.0 convergence, will also be applied to simulate (15.2).

In the following simulations, the initial condition and the parameters are fixed to $X_0 = 0.5$, $r = 5$, $k = 3$, $\theta_1 = 0$, $\theta_2 = 4$ and $\theta_3 = 0.1$. For comparison, the solution of the order 2.0 Itô–Taylor scheme with the step size $h = 2^{-12}$ will be used as the “exact” solution and compared with the other schemes with step sizes $h = 2^{-10}$, 2^{-9} , 2^{-8} , 2^{-7} and 2^{-6} .

The step size versus mean error is shown in Fig. 15.2. The solid thick lines on Fig. 15.2 are the error by 1-step schemes and the dashed lines by multi-step schemes.

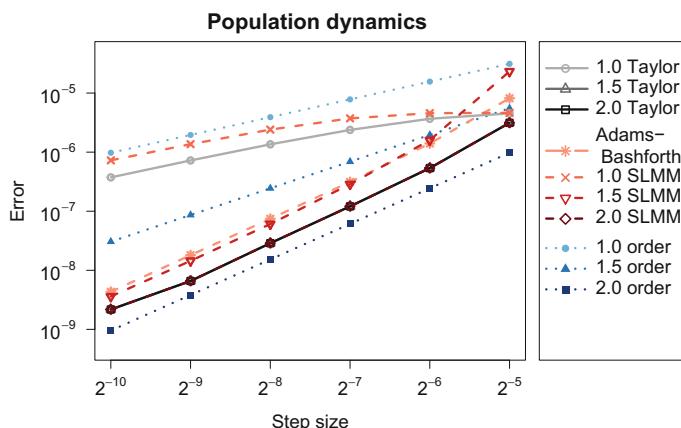


Fig. 15.2 The comparison of accuracy for simulation of (15.2)

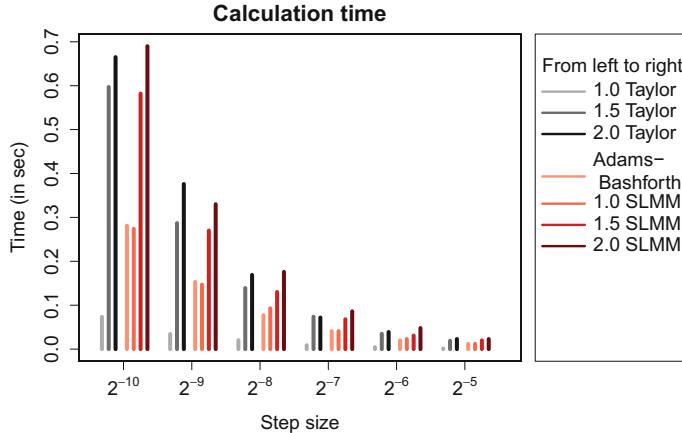


Fig. 15.3 The comparison of calculation times for simulation of (15.2)

The dotted straight lines are for reference and they have slopes of orders 1.0, 1.5 and 2.0. It can be observed in Fig. 15.2 that the order 1.5 Itô–Taylor scheme (solid triangle) almost coincides with the order 2.0 Itô–Taylor scheme (solid box). The Adams–Bashforth scheme (dashed star) has order 1.0 convergence, but its consistency condition satisfies $C_{\alpha,j}^* = 0$ for $\alpha = (1, 0)$ and $(0, 0)$. Moreover, $\cos(Y_t)$ oscillates around 0, which gives $L^1 f \approx 0$ and $R_1 \approx 0$, so the Adams–Bashforth scheme in fact shows higher order of convergence, while the other order 1.0 SLMM (dashed cross) does not.

The computational costs for 100 simulations are shown on the Fig. 15.3. At each of the step size $2^{-10}, 2^{-9}, 2^{-8}, 2^{-7}, 2^{-6}, 2^{-5}$, the calculation times of order 1.0, 1.5 and 2.5 Itô–Taylor, Adams–Bashforth, the order 1.0, 1.5 and 2.0 SLMMs are listed from the left to the right. No significant difference in computational costs are observed here, although those of the order 1.5 and 2.0 SLMMs are slightly smaller than Itô–Taylor schemes.

15.3 Toggle Switch Model

Consider two interacting genes X and Y and denote the concentrations of the corresponding protein products by x and y , respectively. A simple toggle switch model with time-dependent parameters is formulated as (see [135])

$$\frac{dx}{dt} = \left(\alpha_x + \frac{x^4}{a^4+x^4} \right) \left(\frac{b^4}{b^4+y^4} \right) - \lambda_x x, \quad (15.3)$$

$$\frac{dy}{dt} = \left(\alpha_y + \frac{y^4}{c^4+y^4} \right) \left(\frac{d^4}{d^4+x^4} \right) - \lambda_y y, \quad (15.4)$$

where the parameters α_x and α_y represent the external activation on genes X and Y respectively, a and c determine auto-activation thresholds, b and d are thresholds for mutual repression and λ_x and λ_y are protein decay rates.

Different numerical schemes will be applied to simulate solutions of the above model (15.3). In particular the order 1, 2, and 3 affine-RODE-Taylor schemes (11.4), (11.5) and (11.6), the order 1 and order 2 RODE-Taylor schemes, and the SLMM of order 1

$$x_n^k = x_{n-1}^k + \frac{1}{2}(f_{0,n-1}^k + f_{0,n-2}^k)h_n + \sum_{j=1}^m f_{j,n-1}^k I_{(j),n-1}, \quad (15.5)$$

the SLMM of order 2

$$\begin{aligned} x_n^k &= x_{n-1}^k + \frac{1}{2}(3f_{0,n-1}^k - f_{0,n-2}^k)h_n + \sum_{j=1}^m f_{j,n-1}^k I_{(j),n-1} \\ &\quad - \frac{1}{2} \sum_{j=0}^m L^j f_{0,n-2}^k I_{(j),n-2} h_n + \sum_{j_1,j_2=0}^m L^{j_1} f_{j_2,n-1}^k I_{(j_1,j_2),n-1}, \end{aligned} \quad (15.6)$$

with

$$I_{\alpha,n} = \int_{t_n}^{t_{n+1}} \cdots \int_{t_n}^{s_2} \eta_{s_1}^{j_1} \cdots \eta_{s_l}^{j_l} ds_1 \cdots ds_l,$$

are applied to simulate the model. The term $f_{j,n-1}^k$ in (15.5) and (15.6) is the k th component of function f_j evaluated at (t_{n-1}, x_{n-1}) .

For the simulations it is assumed that α_x and α_y are given by two independent compound Poisson processes (14.13). The parameters are set to $a = c = 0.25$, $b = d = 0.4$ and $\lambda_x = \lambda_y = 1.25$ with the initial values $x_0 = y_0 = 10$. The jump magnitudes of two compound Poisson processes follow a uniform distribution on $[0, 0.5]$ and the parameter λ is fixed to 5 in both cases.

The order 3 affine-RODE-Taylor scheme (11.6) with the step size $h = 2^{-9}$ is used as the “exact” solution, to compare with approximations by (11.4), (11.5), (15.5), (15.6), order-1 and order-2 RODE-Taylor schemes with different step sizes $h = 2^{-8}, 2^{-7}, 2^{-6}, 2^{-5}, 2^{-4}, 2^{-3}, 2^{-2}, 2^{-1}$. The values at time t_1 for the SLMMs were calculated by the same order affine-RODE-Taylor schemes.

The step size versus mean error of different schemes for 100 simulations are shown in Fig. 15.4.

The solid lines on the left and right plots of Fig. 15.4 are the errors by affine-RODE-Taylor schemes, the dashed lines are the errors by SLMMs and RODE-Taylor schemes, and the dotted lines are the reference lines with order 1, 2 and 3 slopes. The computational costs for 100 simulations are shown in Fig. 15.5. Both of the order 1 and 2 RODE-Taylor schemes (dashed and dotted lines) coincide closely with the 1- and 2-order affine-RODE-Taylor schemes (solid lines), respectively. No big difference in computational costs can be observed among the affine-RODE-Taylor schemes, SLMMs and RODE-Taylor schemes.

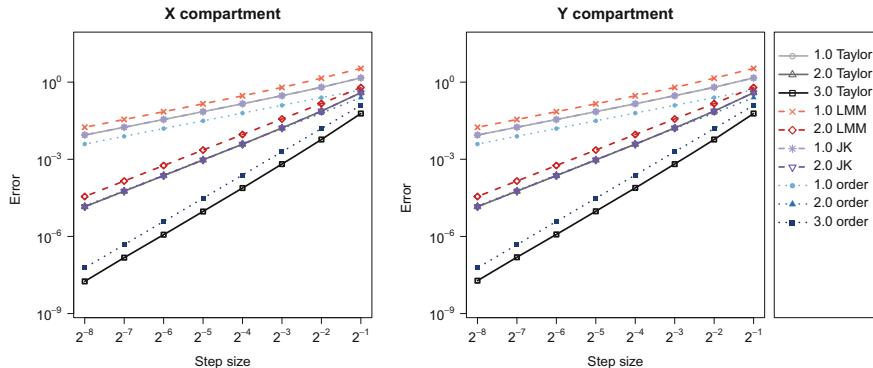


Fig. 15.4 The comparison of accuracy for the RODE (15.3)

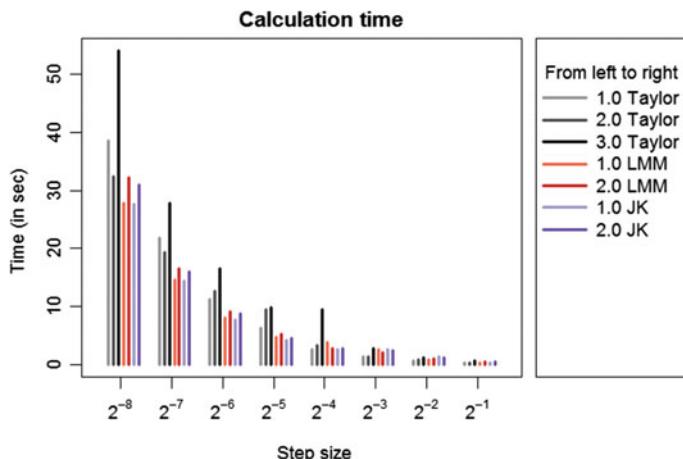


Fig. 15.5 The comparison of calculation time for the RODE (15.3)

15.4 Sea Shell Pattern Model

Consider the following system of random partial differential equations (RPDEs)

$$\begin{aligned} \frac{\partial a}{\partial t} &= s \left(\frac{a^2}{b} + \beta_a \right) - r_a a + D_a \frac{\partial^2 a}{\partial x^2}, \\ \frac{\partial b}{\partial t} &= s a^2 - r_b b + D_b \frac{\partial^2 b}{\partial x^2} + \beta_b, \end{aligned} \quad (15.7)$$

on a spatial domain given by the bounded interval $0 \leq x \leq x_f$ with Neumann boundary conditions. This system describes an interaction between the activator $a(t, x)$ and the inhibitor $b(t, x)$ in pattern formation of sea shells [104]. Here D_a and D_b are the

diffusion coefficients, r_a and r_b are the decay rates of a and b , respectively β_a and β_b are the basic activator and inhibitor production and s is the ability of the cells to perform autocatalysis. The parameter s is given by random fluctuations around r_a through the noise process

$$s(Y_t) := r_a \left(1 - \frac{2\nu}{\pi} \arctan Y_t \right),$$

where $\nu = 0.01$ and Y_t is now a Wiener process.

The system of RPDEs (15.7) is approximated by a system of RODEs by the method of lines [119] using a uniform partition of the interval $[0, x_f]$ with grid size $\Delta x = x_f/M$. In particular, the second order derivatives in (15.7) are approximated by central finite difference quotients

$$\begin{aligned}\frac{\partial^2 a_j}{\partial x^2} &= \frac{a_{j+1} - 2a_j + a_{j-1}}{\Delta x^2} + \mathcal{O}(\Delta x^2), \\ \frac{\partial^2 b_j}{\partial x^2} &= \frac{b_{j+1} - 2b_j + b_{j-1}}{\Delta x^2} + \mathcal{O}(\Delta x^2),\end{aligned}$$

where a_j and b_j are the values of a and b at j -th grid in space. Then $a(x_j, t)$ and $b(x_j, t)$ are replaced by $a_j^*(t)$ and $b_j^*(t)$ for $i = 0, 1, \dots, M$ and the local discretisation error is discarded. This results in a $2(M + 1)$ -dimensional system of RODEs for which the core blocks are

$$\begin{aligned}\frac{da_j^*}{dt} &= s(Y(t)) \left(\frac{a_j^{*2}}{b_j^*} + \beta_a \right) - r_a a_j^* + D_a \frac{a_{j+1}^* - 2a_j^* + a_{j-1}^*}{\Delta x^2} \\ \frac{db_j^*}{dt} &= s(Y(t)) a_j^{*2} - r_b b_j^* + D_b \frac{b_{j+1}^* - 2b_j^* + b_{j-1}^*}{\Delta x^2} + \beta_b\end{aligned}$$

with initial conditions

$$a_j^*(t = 0) = a(x = x_j, t = 0) \quad \text{and} \quad b_j^*(t = 0) = b(x = x_j, t = 0)$$

for $j = 0, 1, \dots, M$. The boundary blocks need to be modified to take into account the Neumann boundary conditions

$$\frac{\partial a^*(x = 0, t)}{\partial x} = \frac{\partial a^*(x = x_f, t)}{\partial x} = 0 \quad \text{and} \quad \frac{\partial b^*(x = 0, t)}{\partial x} = \frac{\partial b^*(x = x_f, t)}{\partial x} = 0.$$

The system is stiff, so the implicit averaged Euler scheme from [12], implicit Itô–Taylor schemes from [9] as well as implicit SLMMs

$$\begin{aligned}(\text{Imp1.0}) \quad x_n &= x_{n-1} + h \left(\frac{1}{2} f_n + \frac{1}{4} f_{n-1} + \frac{1}{4} f_{n-2} \right) \\ (\text{Imp1.5}) \quad x_n &= x_{n-1} + h \left(\frac{5}{12} f_n + \frac{8}{12} f_{n-1} - \frac{1}{12} f_{n-2} \right)\end{aligned}$$

$$\begin{aligned}
 & + L^1 f_{n-1} I_{(1,0),n-1} - \frac{5}{12} L^1 f_{n-1} h I_{(1),n-1} - \frac{1}{12} L^1 f_{n-2} h I_{(1),n-2} \\
 & + L^0 f_{n-1} I_{(0,0),n-1} - \frac{5}{12} L^0 f_{n-1} h I_{(0),n-1} - \frac{1}{12} L^0 f_{n-2} h I_{(0),n-2},
 \end{aligned}$$

are applied to the model. The first implicit SLMM has order 1.0 convergence and the second has order 1.5.

In the simulation, the parameters are set to $D_a = 0.01$, $D_b = 0.4$, $r_a = 0.05$, $r_b = 0.08$ and $\beta_a = 0.05$ and $\beta_b = 0$ [104] with the initial values $a_0 = 0.2$ and $b_0 = 0.1$. Moreover, $x_0 = 0$ and $x_f = 5$ with $\Delta x = 2^{-2}$, which gives a 42 dimensional system of RODEs. The “exact” solution is obtained by order 1.5 implicit Itô–Taylor scheme with the step size $h = 2^{-11}$ and compare with approximations by order 1 and 1.5 implicit Itô–Taylor schemes and the implicit SLMMs with different step sizes $h = 2^{-9}, 2^{-8}, 2^{-7}, 2^{-6}$ and 2^{-5} and implicit averaged Euler scheme with different step sizes $h = 2^{-8}, 2^{-7}, 2^{-6}$ and 2^{-5} . For the multi-step schemes, the initial conditions at time t_1 were calculated using 1-step schemes of the same orders.

The sea shell pattern simulated by different schemes are shown in Fig. 15.6.

Figures 15.7 and 15.8 illustrate the step size versus mean error and computational costs for 100 times simulation by the above schemes. The solid thick lines are the error by 1-step schemes and the dashed lines by multi-step schemes. The solid thin lines are for reference and have slopes of 1.0, 1.5 and 2.0.

The difference between the 1.5-order and 2.0-order schemes comes from $L^1 L^1 f$ terms and it depends deeply on the value of a in the inhibitor compartment, in particular a is small when time is small. This makes the 1.5-order schemes show roughly 2.0-order convergence decay.

The system is now of relatively high dimension and a difference in computational costs, especially between the order 1.5 Itô–Taylor scheme and the SLMM, is quite apparent.

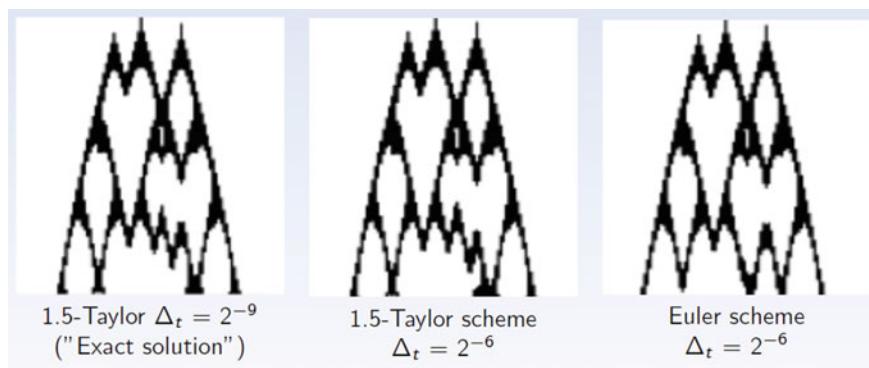


Fig. 15.6 Simulations of sea shell pattern model (15.7) by order 1 and 1.5 Itô–Taylor schemes

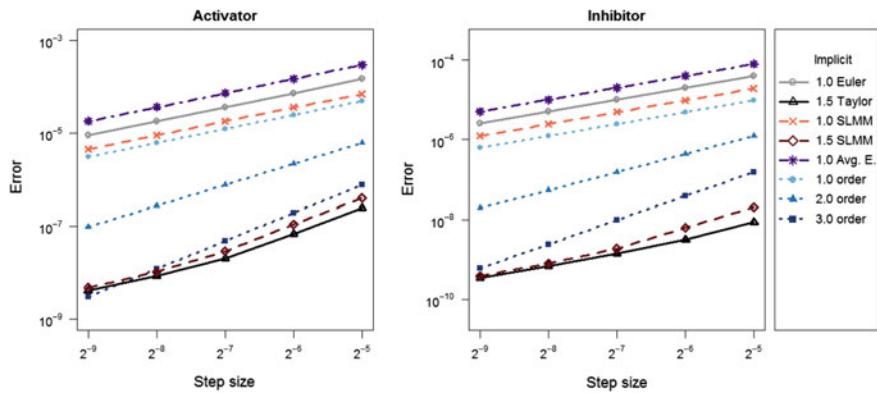


Fig. 15.7 The comparison of accuracy for (15.7)

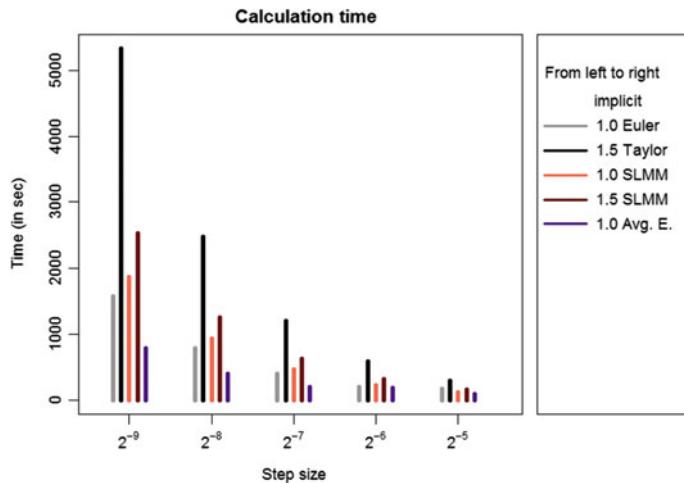


Fig. 15.8 The comparison of calculation time for (15.7)

15.5 Endnotes

This chapter is based on the simulations in Asai et al. [9–12]. The population model (15.2) is based on Higham [63].

The model on pattern formation of sea shells is taken from Meinhardt [104]. See Schiesser and Griffiths [119] for the method of lines. Riesinger, Neckel and Rupp [115] solved RODEs on GPU clusters with multiple levels of parallelism using numerical schemes introduced in this book.

Chapter 16

Chemostat

A chemostat is associated with a laboratory device which consists of three interconnected vessels and is used to grow microorganisms in a cultured environment (see Fig. 16.1). In its basic form, the outlet of the first vessel is the inlet for the second vessel and the outlet of the second vessel is the inlet for the third. The first vessel is called a feed bottle, which contains all the nutrients required to grow the microorganisms. All nutrients are assumed to be abundantly supplied except one, which is called a limiting nutrient. The contents of the first vessel are pumped into the second vessel, which is called the culture vessel, at a constant rate. The microorganisms feed on nutrients from the feed bottle and grow in the culture vessel. The culture vessel is continuously stirred so that all the organisms have equal access to the nutrients. The contents of the culture vessel are then pumped into the third vessel, which is called a collection vessel. Naturally it contains nutrients, microorganisms and the products produced by the microorganisms.

The chemostat is the simplest form of competition in biology that occurs when two or more populations compete for the same resource, e.g., a common food supply of a growth-limiting nutrient. As the best laboratory idealisation of nature to study such competitions, the chemostat plays an important role in theoretical ecology, waste water treatment and the study of recombinant problems in genetically altered organism, amongst many other areas.

Two standard assumptions for simple chemostat models are:

- (i) the availability of the nutrient and its supply rate are fixed;
- (ii) the tendency of microorganisms to adhere to surfaces is not taken into account.

Let $x(t)$ denote the concentration of the growth-limiting nutrient and $y(t)$ the concentration of the microorganism at any specific time t . Then the above assumptions lead to the following growth equations

$$\frac{dx(t)}{dt} = D(I - x(t)) - aU(x(t))y(t), \quad (16.1)$$

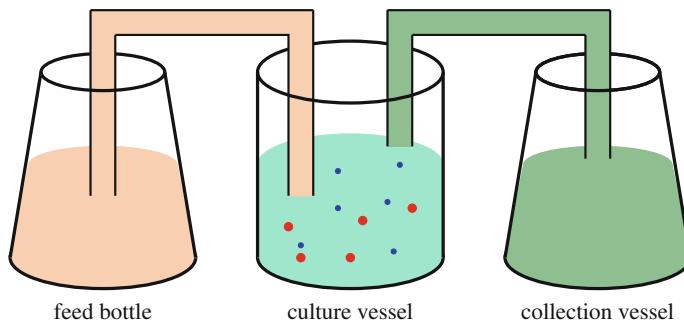


Fig. 16.1 A typical chemostat

$$\frac{dy(t)}{dt} = -Dy(t) + aU(x(t))y(t), \quad (16.2)$$

where D is the rate at which the nutrient is supplied and also the rate at which the contents of the growth medium are removed, I is the input nutrient concentration which describes the quantity of nutrient available within the system at any time, a is the maximal consumption rate of the nutrient and also the maximum specific growth rate of microorganisms, and U is the functional response of the microorganism describing how the nutrient is consumed by the species.

Typically U follows the Michaelis–Menten or Holling type-II form

$$U(x) = \frac{x}{m+x},$$

where $m > 0$ is the half-saturation constant.

16.1 Random Chemostat Models

In reality the chemostat models are neither autonomous nor deterministic: they process information provided by physical or chemical inputs with noise, which can be caused by environmental perturbations, internal variability, randomly fluctuating parameters, measurement errors, etc. This motivates the study of chemostat models with randomly fluctuating input parameters, the nutrient supplying rate D and the nutrient supplying concentration I . These will be formalized as $D(\theta_t(\omega))$ and $I(\theta_t(\omega))$, where θ_t is canonical noise process, i.e., a measure theoretic group, on an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$, see Chap. 4.

The chemostat model (16.1) and (16.2) then forms a system of RODEs,

$$\frac{dx(t)}{dt} = D(\theta_t(\omega))(I(\theta_t(\omega)) - x(t)) - a \frac{x(t)}{m + x(t)} y_1(t), \quad (16.3)$$

$$\frac{dy(t)}{dt} = -D(\theta_t(\omega))y(t) + a \frac{x(t)}{m + x(t)} y(t). \quad (16.4)$$

In particular, it will be assumed here that the inputs are perturbed by real bounded noise, i.e., $D(\theta_t(\omega))$ and $I(\theta_t(\omega))$ are continuous and essentially bounded:

$$D(\theta_t(\omega)) \in d \cdot [1 - \varepsilon, 1 + \varepsilon], \quad I(\theta_t(\omega)) \in i \cdot [1 - \varepsilon, 1 + \varepsilon], \quad d > 0, \quad i > 0, \quad \varepsilon < 1.$$

This is a natural way to model the realistic stochastic fluctuations of a biological system caused by its interaction with the external world, because the parameters in dynamical systems of biological interest are inherently positive and bounded. See Chap. 1 for some examples.

Random Chemostat with Wall Attachment and Material Recycling

Another drawback of the traditional chemostat model (16.1) and (16.2) is the ignorance of wall attachment of microorganisms. Most of the time microorganisms grow not only in the growth medium, but also along the walls of the container, either due to the ability of the microorganisms to stick on to the walls of the container or the flow rate is not fast enough to wash these organisms out of the system. The consumer population $y(t)$ can be regarded as an aggregate of two categories of populations, one in the growth medium, denoted by $y_1(t)$ and the other on the walls of the container, denoted by $y_2(t)$. Individuals may switch their categories at any time, i.e., the microorganisms on the walls may join those in the growth medium or the biomass in the medium may prefer walls.

Let r_1 and r_2 represent the rates at which the organisms stick on to and shear off from the walls, respectively. Assume that the nutrient is equally available to both of the categories, i.e., it is assumed that both categories consume the same amount of nutrient and at the same rate. When the flow rate is low, the organisms may die naturally before being washed out and thus washout is no longer the only prime factor of loss. Denote by $\mu > 0$ the collective death rate coefficient of $y(t)$ representing all the aforementioned factors such as diseases, aging, etc. On the other hand, when the flow rate is small, the dead biomass is not sent out of the system immediately and is subject to bacterial decomposition which in turn leads to regeneration of the nutrient. Let the constant $b \in (0, 1)$ describe the fraction of dead biomass that is recycled.

The evolution of concentrations of the nutrient and microorganism when the input parameters are random and wall growth is taken into account can be described by the system of RODEs

$$\frac{dx}{dt} = D(\theta_t(\omega))(I(\theta_t(\omega)) - x(t)) - a \frac{x(t)}{m + x(t)} (y_1(t) + y_2(t)) + b\mu y_1(t), \quad (16.5)$$

$$\frac{dy_1}{dt} = -(\mu + D(\theta_t(\omega))) y_1(t) + c \frac{x(t)}{m + x(t)} y_1(t) - r_1 y_1(t) + r_2 y_2(t), \quad (16.6)$$

$$\frac{dy_2}{dt} = -\mu y_2(t) + c \frac{x(t)}{m+x(t)} y_2(t) + r_1 y_1(t) - r_2 y_2(t), \quad (16.7)$$

where $0 < c \leq a$ is the growth rate coefficient of the consumer species.

16.2 RDS Generated by Random Chemostat

It will be shown here that the RODE (16.5)–(16.7) for the random chemostat with a wall generates a random dynamical system as defined in Chap. 4. A similar proof holds for the RODE (16.1) and (16.2) modeling a random chemostat without a wall.

Let

$$\mathbb{R}_+^3 = \{(x, y, z) \in \mathbb{R}^3 : x \geq 0, y \geq 0, z \geq 0\}.$$

Lemma 16.1 *For any $t_0 \in \mathbb{R}$, initial data $v_0 := (x(t_0), y_1(t_0), y_2(t_0))^T \in \mathbb{R}_+^3$ and $\omega \in \Omega$, the system (16.5)–(16.7) admits a unique bounded solution $v(\cdot; t_0, \omega, v_0) \in \mathbb{C}([t_0, \infty), \mathbb{R}_+^3)$ with $v(t_0; t_0, \omega, v_0) = v_0$. Moreover, the solution generates a random dynamical system $\varphi(t, \omega)(\cdot)$ defined by*

$$\varphi(t, \omega)v_0 := v(t; 0, \omega, v_0), \quad \forall t \geq 0, \quad v_0 \in \mathbb{R}_+^3, \quad \omega \in \Omega.$$

Proof With $v(t) = (x(t), y_1(t), y_2(t))^T$, the system (16.5)–(16.7) can be written as

$$\frac{dv}{dt} = L(\theta_t(\omega)) \cdot v + F(v, \theta_t(\omega)),$$

where

$$L(\theta_t(\omega)) = \begin{pmatrix} -D(\theta_t(\omega)) & b\mu & 0 \\ 0 & -\mu - D(\theta_t(\omega)) - r_1 & r_2 \\ 0 & r_1 & -\mu - r_2 \end{pmatrix},$$

and $F : \mathbb{R}_+^3 \times [t_0, +\infty) \rightarrow \mathbb{R}^3$ is given by

$$F(v, \theta_t(\omega)) = \begin{pmatrix} D(\theta_t(\omega))I(\theta_t(\omega)) - \frac{ax}{m+x}(y_1 + y_2) \\ \frac{cx}{m+x}y_1 \\ \frac{cx}{m+x}y_2 \end{pmatrix},$$

where $(x, y_1, y_2) \in \mathbb{R}_+^3$.

Since $D(\theta_t(\omega))$ is bounded, the operator L generates an evolution system on \mathbb{R}^3 . Also because $D(\theta_t(\omega))$ and $I(\theta_t(\omega))$ are continuous with respect to t , function

$F(\cdot, \theta_t(\omega)) \in \mathbb{C}(\mathbb{R}_+^3 \times [t_0, \infty), \mathbb{R}^3)$ and is continuously differentiable with respect to the variables (x, y_1, y_2) , which implies that it is locally Lipschitz with respect to (x, y_1, y_2) in \mathbb{R}_+^3 . Therefore, by classical existence and uniqueness theorems for ODEs, e.g., Theorem 2.1 in Chap. 2, the system (16.5)–(16.7) possesses a unique local solution.

By continuity of solutions, each solution has to take value 0 before it reaches a negative value. With $x = 0$ and $y_1, y_2 \geq 0$, Eq. (16.5) becomes

$$\frac{dx}{dt} = D(\theta_t(\omega))I(\theta_t(\omega)) + b\mu y_1(t) > 0,$$

and thus $x(t)$ is strictly increasing at $x = 0$. Similarly, with $y_1 = 0$ and $x, y_2 \geq 0$, Eq. (16.6) gives $y'_1(t) = r_2 y_2 \geq 0$, and with $y_2 = 0$ and $x, y_1 \geq 0$, Eq. (16.7) gives $y'_2(t) = r_1 y_1 \geq 0$. Therefore, $y_1(t)$ is non-decreasing at $y_1 = 0$ and $y_2(t)$ is non-decreasing at $y_2 = 0$. In conclusion, $v(t) \in \mathbb{R}_+^3$ for any $t \geq 0$. (Alternatively, one sees that the vector field is quasipositive and applies Theorem 2.4 in Chap. 2).

For $v(t) \in \mathbb{R}_+^3$, define $\|v(t)\|_1 := x(t) + y_1(t) + y_2(t)$, and let

$$s(t) = x(t) + \frac{a}{c}(y_1(t) + y_2(t)).$$

Since $a \geq c$, then

$$\|v(t)\|_1 \leq s(t) \leq \frac{a}{c} \cdot \|v(t)\|_1.$$

The time derivative of $s(t)$ along solutions to (16.5)–(16.7) satisfies

$$\begin{aligned} \frac{ds(t)}{dt} &= D(\theta_t(\omega)) [I(\theta_t(\omega)) - x(t)] - \left[\frac{a}{c} (\mu + D(\theta_t(\omega))) - b\mu \right] y_1(t) - \frac{a}{c} \mu y_2(t) \\ &\leq di(1 + \varepsilon)^2 - d(1 - \varepsilon)x(t) - \left[\frac{a}{c} (\mu + d(1 - \varepsilon)) - b\mu \right] y_1(t) - \frac{a}{c} \mu y_2(t). \end{aligned}$$

Note that, since $a \geq c$ and $0 < b < 1$,

$$\frac{a}{c}(\mu + d(1 - \varepsilon) - b\mu) = \frac{a}{c}d(1 - \varepsilon) + \left(\frac{a}{c} - b \right) \mu > \frac{a}{c}d(1 - \varepsilon).$$

Therefore, letting $\lambda := \min\{d(1 - \varepsilon), \mu\}$ gives

$$\frac{ds(t)}{dt} \leq di(1 + \varepsilon)^2 - \lambda s(t).$$

Now, $s(t)$ is non-increasing when $s(0) \geq di(1 + \varepsilon)^2/\lambda$, and thus $s(t) \leq s(0)$, while $s(t) \leq di(1 + \varepsilon)^2/\lambda$ for all $t \geq 0$ when $s(0) < di(1 + \varepsilon)^2/\lambda$. These imply that $\|u(t)\|_1$ is bounded:

$$0 \leq \|v(t)\|_1 \leq \max \left\{ di(1 + \varepsilon)^2/\lambda, x(0) + \frac{a}{c}(y_1(0) + y_2(0)) \right\}, \quad t \geq 0.$$

Therefore the local solution to system (16.5)–(16.7) can be extended to a global solution $v(\cdot; t_0, \omega, u_0) \in \mathbb{C}^1([t_0, \infty), \mathbb{R}^3)$.

It is straightforward to check that

$$v(t + t_0; t_0, \omega, v_0) = v(t; 0, \theta_{t_0}(\omega), v_0)$$

for all $t_0 \in \mathbb{R}$, $t \geq t_0$, $\omega \in \Omega$ and $u_0 \in \mathbb{R}_+^3$. Since function $F(v, \theta_t(\omega)) = F(v, t, \omega)$ is continuous in v , t and is measurable in ω , $v : [0, \infty) \times \Omega \times \mathbb{R}_+^3 \rightarrow \mathbb{R}_+^3$, defined by the mapping $(t; \omega, v_0) \mapsto v(t; 0, \omega, v_0)$ is $(\mathbb{B}[0, \infty) \times \mathbb{F}_0 \times \mathbb{B}(\mathbb{R}_+^3), \mathbb{B}(\mathbb{R}_+^3))$ -measurable.

It then follows directly that (16.5)–(16.7) generate the continuous random dynamical system with cocycle mapping $\varphi(t, \omega)(\cdot)$ defined by

$$\varphi(t, \omega)v_0 = v(t; 0, \omega, v_0), \quad \forall t \geq 0, \quad v_0 \in \mathbb{R}_+^3, \quad \omega \in \Omega.$$

□

Henceforth $v(t; 0, \omega, v_0)$ will be written $v(t; \omega, v_0)$. Also, since $\omega \in \Omega$ is usually fixed in the following calculations, it will not be mentioned explicitly and $v(t; \omega, v_0)$ will often be written simply as $v(t)$, and similarly for its components.

16.3 Existence of a Random Attractor

The system of RODEs (16.1)–(16.2) a chemostat without a wall and with both random nutrient input concentration and random nutrient supplying rate generates a random dynamical system.

Let $u(t) := x(t) + y_1(t)$. Then $u(t)$ satisfies

$$\frac{du(t)}{dt} = D(\theta_t(\omega)) [I(\theta_t(\omega)) - u(t)] \quad (16.8)$$

This has a nontrivial random solution which is both forward and pullback attracting. In fact, for any initial value $u_0 = x(0) + y_1(0)$, its explicit solution is

$$u(t; \omega, u_0) = u_0 e^{- \int_0^t D(\theta_s(\omega)) ds} + \int_0^t D(\theta_s(\omega)) I(\theta_s(\omega)) e^{- \int_s^t D(\theta_\tau(\omega)) d\tau} ds. \quad (16.9)$$

Replacing ω by $\theta_{-t}(\omega)$ in (16.9) gives

$$u(t; \theta_{-t}(\omega), u_0) = u_0 e^{- \int_{-t}^0 D(\theta_s(\omega)) ds} + \int_{-t}^0 D(\theta_s(\omega)) I(\theta_s(\omega)) e^{- \int_s^0 D(\theta_\tau(\omega)) d\tau} ds,$$

which is pullback convergent (i.e., when $t \rightarrow +\infty$ here) to

$$u^*(\omega) := \int_{-\infty}^0 D(\theta_s(\omega)) I(\theta_s(\omega)) e^{-\int_s^0 D(\theta_\tau(\omega)) d\tau} ds.$$

Since $d(1 - \varepsilon) \leq D(\theta_t(\omega)) \leq d(1 + \varepsilon)$ and $i(1 - \varepsilon) \leq I(\theta_t(\omega)) \leq i(1 + \varepsilon)$,

$$di(1 - \varepsilon)^2 \int_{-\infty}^0 e^{-\int_s^0 d(1+\varepsilon) d\tau} ds \leq u^*(\omega) \leq di(1 + \varepsilon)^2 \int_{-\infty}^0 e^{-\int_s^0 d(1-\varepsilon) d\tau} ds,$$

i.e.,

$$\frac{i(1 - \varepsilon)^2}{1 + \varepsilon} \leq u^*(\omega) \leq \frac{i(1 + \varepsilon)^2}{1 - \varepsilon}.$$

For a fixed and small enough $\delta > 0$, consider the nonempty compact set

$$K(\omega) := \left\{ (x, y_1) \in \mathbb{R}_+^2 : \frac{i(1 - \varepsilon)^2}{1 + \varepsilon} - \delta \leq x + y_1 \leq \frac{i(1 + \varepsilon)^2}{1 - \varepsilon} + \delta \right\}.$$

Then for any tempered bounded random set $B = \{B(\omega) : \omega \in \Omega\} \in \mathcal{F}(\mathbb{R}_+^2)$ (see Chap. 4), there exists $T_B(\omega) > 0$ such that

$$\varphi(t, \theta_{-t}(\omega)) B(\theta_{-t}(\omega)) \subset K(\omega), \quad t \geq T_B(\omega),$$

i.e., $K(\omega)$ is positively invariant and absorbing in \mathbb{R}_+^2 for system (16.8).

The random dynamical system generated by system (16.3) and (16.4) thus has a random attractor $\mathbb{A} = \{A(\omega) : \omega \in \Omega\}$ consisting of non-empty compact subsets of $K(\omega)$.

The next theorem provides sufficient conditions for the extinction and persistence of microorganism y_1 .

Theorem 16.1 *The random pullback attractor $\mathbb{A} = \{A(\omega) : \omega \in \Omega\}$ for the random dynamical system generated by the RODEs (16.3) and (16.4)*

(i) *has singleton component sets $A(\omega) = \{(u^*(\omega), 0)\}$ for every $\omega \in \Omega$, when*

$$d(1 - \varepsilon) \geq a;$$

(ii) *has nontrivial component sets which include $(u^*(\omega), 0)$ and strictly positive points, when*

$$d(1 + \varepsilon) \leq \frac{ai(1 - \varepsilon)^3}{m(1 - \varepsilon^2) + i(1 + \varepsilon)^3};$$

(iii) *contains a nontrivial entire solution that attracts all other strictly positive entire solutions, when*

$$d(1 - \varepsilon) > a - \frac{ami(1 - \varepsilon)^2 / (1 + \varepsilon)}{(m + i(1 + \varepsilon)^2 / (1 - \varepsilon))^2}.$$

Proof First note that since a random attractor contains all bounded entire solutions, then $(u^*(\omega), 0) \in A(\omega)$ for every $\omega \in \Omega$.

(i) Equation (16.4) can be written as

$$\begin{aligned}\frac{dy_1(t)}{dt} &= \left(-D(\theta_t(\omega)) + \frac{ax(t)}{m+x(t)} \right) y_1(t) \\ &< (-d(1-\varepsilon) + a)y_1(t).\end{aligned}$$

Thus $y'_1(t) < 0$ when $d(1-\varepsilon) > a$. This implies that $y_1(t) \rightarrow 0$ as $t \rightarrow \infty$ and $(x(t, \omega), y_1(t)) = (u^*(\theta_t(\omega)), 0)$ is asymptotically stable in \mathbb{R}_+^2 .

(ii) Besides $x + y_1 = i(1+\varepsilon)^2/(1-\varepsilon) - \delta$ and $x + y_1 = i(1-\varepsilon)^2/(1+\varepsilon) + \delta$, the absorbing set $K(\omega)$ has two other edges $x = 0$ and $y_1 = 0$.

On the edge $x = 0$, the derivative $x'(t)|_{x=0} = D(\theta_t(\omega))I(\theta_t(\omega)) > 0$, so $x(t)$ is increasing towards the interior of $K(\omega)$ on this edge. The edge $y_1 = 0$ is invariant as $y'_1|_{y_1=0} = 0$. But for $y = \epsilon \ll i(1-\varepsilon)^2/(1+\varepsilon)$, $x(t)$ satisfies

$$\frac{i(1-\varepsilon)^2}{1+\varepsilon} - \delta \leq x(t) \leq \frac{i(1+\varepsilon)^2}{1-\varepsilon} + \delta.$$

Hence, when δ is small enough and

$$d(1+\varepsilon) \leq \frac{ai(1-\varepsilon)^3}{m(1-\varepsilon^2) + i(1+\varepsilon)^3},$$

Equation (16.4) gives

$$\begin{aligned}\frac{dy_1}{dt} &= \left(-D(\theta_t(\omega)) + \frac{ax(t)}{m+x(t)} \right) y_1 \\ &> \left(\frac{ai(1-\varepsilon)^3 - \delta(1+\varepsilon)}{m(1-\varepsilon^2) + i(1+\varepsilon)^3 - \delta(1+\varepsilon)} - d(1+\varepsilon) \right) y_1 \geq 0.\end{aligned}$$

This implies that the positive interior of the absorbing set also contains points of the random attractor.

(iii) Now consider Eq.(16.3) restricted to the stable manifold $x(t) + y_1(t) = u^*(\theta_t(\omega))$, on which

$$\frac{dx(t)}{dt} = D(\theta_t(\omega))(I(\theta_t(\omega)) - x(t)) - \frac{ax(t)}{m+x(t)}(u^*(\theta_t(\omega)) - x(t)). \quad (16.10)$$

For any two solutions $x_1(t)$ and $x_2(t)$ to Eq. (16.10), define $\Delta(t) = x_1(t) - x_2(t)$. Then $\Delta(t)$ satisfies

$$\begin{aligned}\frac{d\Delta(t)}{dt} &= -D(\theta_t(\omega))\Delta(t) - au^*(\theta_t(\omega))\left(\frac{x_1}{m+x_1} - \frac{x_2}{m+x_2}\right) \\ &\quad + a\left(\frac{x_1^2}{m+x_1} - \frac{x_2^2}{m+x_2}\right) \\ &= -D(\theta_t(\omega))\Delta(t) - \frac{amu^*(\theta_t(\omega))}{(m+x_1)(m+x_2)}\Delta(t) \\ &\quad + a\frac{m(x_1+x_2)+x_1x_2}{(m+x_1)(m+x_2)}\Delta(t).\end{aligned}$$

Since, for t large enough,

$$x_1(t), x_2(t) \leq u^*(\theta_t(\omega)) \leq i(1+\varepsilon)^2/(1-\varepsilon) + \delta,$$

and $u^*(\theta_t(\omega)) \geq i(1-\varepsilon)^2/(1+\varepsilon) - \delta$, it follows that

$$\frac{d\Delta(t)}{dt} < -d(1-\varepsilon)\Delta(t) - \frac{am(i(1-\varepsilon)^2/(1+\varepsilon) - \delta)}{(m+i(1+\varepsilon)^2/(1-\varepsilon) + \delta)^2}\Delta(t) + a\Delta(t).$$

Hence for δ small enough, $\Delta(t) \rightarrow 0$ as $t \rightarrow \infty$ when

$$d(1-\varepsilon) + \frac{ami(1-\varepsilon)^2/(1+\varepsilon)}{(m+i(1+\varepsilon)^2/(1-\varepsilon))^2} > a.$$

This always holds if $d(1-\varepsilon) \geq a$, which is scenario (i) of the theorem.

This limit can still hold if a is slightly larger. In fact,

$$\frac{mi(1-\varepsilon)^2/(1+\varepsilon)}{(m+i(1+\varepsilon)^2/(1-\varepsilon))^2} < \frac{mi}{(m+i)^2} < 1.$$

In the above estimates if neither $x_1(t)$ nor $x_2(t)$ equals $u^*(\theta_t(\omega))$, the system is strictly uniformly contracting in the positive quadrant and thus has a unique entire solution $u^*(\theta_t(\omega))$, the summands of which form the components of the pullback attractor inside the positive quadrant.

□

Random Chemostat with Wall Attachment

The random chemostat with a wall (16.5)–(16.7) also has a random attractor. The results are formulated here and proved in a similar way. The reader is referred to Caraballo, Han and Kloeden [27] for proofs. More detailed information on the internal structure of the pullback attractor can be obtained with the following change of variables

$$y(t) = y_1(t) + y_2(t); \quad \gamma(t) = \frac{y_1(t)}{y(t)}.$$

The system of RODEs (16.5)–(16.7) then becomes

$$\frac{dx(t)}{dt} = D(\theta_t(\omega))(I(\theta_t(\omega)) - x(t)) - \frac{ax(t)}{m + x(t)}y(t) + b\mu\gamma(t)y(t), \quad (16.11)$$

$$\frac{dy(t)}{dt} = -\mu y(t) - D(\theta_t(\omega))\gamma(t)y(t) + \frac{cx(t)}{m + x(t)}y(t), \quad (16.12)$$

$$\frac{d\gamma(t)}{dt} = -D(\theta_t(\omega))\gamma(t)(1 - \gamma(t)) - r_1\gamma(t) + r_2(1 - \gamma(t)). \quad (16.13)$$

By definition, $\gamma(t)$ represents the portion of microorganism that attaches to the wall. The dynamics of $\gamma(t) = \gamma(t; \omega, \gamma_0)$ are uncoupled with $x(t)$ and $y(t)$.

Theorem 16.2 *The pullback attractor \mathbb{A} in \mathbb{R}^+ associated to the random dynamical system $\gamma(t, \omega, \cdot)$ generated by (16.13) consists of a single entire solution denoted by $\gamma^*(\theta_t(\omega))$, provided*

$$2r_2d(1 + \varepsilon) < (r_1 + r_2 + d(1 - \varepsilon))(r_1 + r_2).$$

Note that the RODE (16.13) is a Riccati equation.

Theorem 16.3 *Given $a \geq c$, $0 < b < 1$, $\mu > 0$ and assume that $D(\theta_t(\omega))$ and $I(\theta_t(\omega))$ are continuous and essentially bounded, with*

$$d(1 - \varepsilon) \leq D(\theta_t(\omega)) \leq d(1 + \varepsilon), \quad i(1 - \varepsilon) \leq I(\theta_t(\omega)) \leq i(1 + \varepsilon).$$

Then, system (16.11)–(16.12) has a random attractor $\mathbb{A} = \{A(\omega) : \omega \in \Omega\}$ inside the nonnegative quadrant. Moreover, defining

$$x^*(\omega) = \int_{-\infty}^0 D(\theta_s(\omega))I(\theta_s(\omega))e^{-\int_s^0 D(\theta_\tau(\omega))d\tau}ds,$$

the random attractor \mathbb{A}

(i) *has a singleton component subset $A(\omega) = \{(x^*(\omega), 0)\}$ provided*

$$\mu + d(1 - \varepsilon)\alpha^* \geq c,$$

$$\text{where } \alpha^* = \frac{r_2}{r_1 + r_2 + d(1 + \varepsilon)}$$

(ii) *also contains points strictly inside the positive quadrant in addition to the singleton solution $\{(x^*(\omega), 0)\}$ provided*

$$ac^2di(1 - \varepsilon)^2 > \left(mc(a\mu + ad(1 + \varepsilon) - cbv\beta^*) + acdi(1 - \varepsilon)^2\right) \cdot (v + d(1 + \varepsilon)\beta^*),$$

where $\beta^* = \frac{r_2}{r_1+r_2}$

16.4 Endnotes

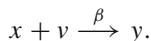
This chapter is based on Caraballo, Han and Kloeden [28]. The proofs of Theorems 16.2 and 16.3 are given in [28]. See also Smith and Waltman [122], Sree and Rao [125] and Caraballo, Han, Kloeden and Rappaport [23].

Chapter 17

Immune System Virus Model

Basic models for virus dynamics are discussed in the classic text by May and Nowak [103]. The assumption for the simplest models is that the body is modeled as a “well stirred” chemostat containing the virus and two kinds of cells, uninfected but susceptible cells and cells infected by virus. As seen in Chap. 16 modeling chemostats by systems of nonautonomous or random differential equations is fully justified, because the environment for a chemostat usually varies in time, either deterministically or randomly. Similarly, the human body also varies in time, so the virus dynamics will be modeled here by a system of RODEs.

Denote by v the population size of free virus, x the population size of uninfected cells (food for virus), and y the population size of infected cells. First, uninfected cells are produced by cell proliferation at a constant rate Λ , live for an average lifetime and die at an average death rate γ_1 . Second, the virus infects susceptible cells to produce infected cells, with an “efficiency” β . Since cells are infected by contact with the virus, the infection can be modeled as a simple mass action reaction



Third, infected cells die at an average rate γ_2 and release new virus at a rate κ . At the same time these viruses are cleared by the immune system at a rate α . This yields a basic model of virus dynamics:

$$\frac{dx(t)}{dt} = \Lambda - \gamma_1 x - \beta xv, \quad (17.1)$$

$$\frac{dy(t)}{dt} = \beta xv - \gamma_2 y, \quad (17.2)$$

$$\frac{dv(t)}{dt} = \kappa y - \alpha v. \quad (17.3)$$

The ordinary differential equation system (17.1)–(17.3) can be used to describe the dynamics of various types of virus, healthy and infected cells, but with limitations. First, the model assumes that the contribution of the immune response (to the death of infected cells or free virus and to reducing the rate of infection of new cells) is constant over time. Second, the dynamics of the susceptible cell population assumes a constant production rate from a pool of precursors. These assumptions may be justified for stationary environments, within a short term of time. However, in the long term, the human body is never a stationary environment – it varies over time in principle. Hence system (17.1)–(17.3) is not adequate to explain the real dynamics of virus and the immune response.

The human body can be considered to be a random environment which varies randomly with respect to time. Due to this random variation, the contribution of the immune response and the production rate of susceptible cells from cell proliferation will also fluctuate randomly with respect to time. More precisely, the parameters Λ and α are supposed here to be perturbed by real noise, i.e., $\Lambda = \Lambda(\theta_t(\omega))$ and $\alpha = \alpha(\theta_t(\omega))$ are continuous and essentially bounded:

$$\Lambda(\theta_t(\omega)) \in \lambda \cdot [1 - \delta_1, 1 + \delta_1], \quad \lambda > 0, \quad 0 < \delta_1 < 1, \quad (17.4)$$

$$\alpha(\theta_t(\omega)) \in a \cdot [1 - \delta_2, 1 + \delta_2], \quad a > 0, \quad 0 < \delta_2 < 1, \quad (17.5)$$

where $\theta_t(\omega)$ is a canonical representation of the driving noise system. Bounded noise can be modeled in various ways as seen in Sect. 1.3 of Chap. 1.

The system (17.1)–(17.3) then becomes

$$\frac{dx(t, \omega)}{dt} = \Lambda(\theta_t(\omega)) - \gamma_1 x - \beta xv, \quad (17.6)$$

$$\frac{dy(t, \omega)}{dt} = \beta xv - \gamma_2 y, \quad (17.7)$$

$$\frac{dv(t, \omega)}{dt} = \kappa y - \alpha(\theta_t(\omega))v, \quad (17.8)$$

where $\gamma_1, \gamma_2, \beta, \kappa$ are positive constants, and $\Lambda(\theta_t(\omega))$ and $\alpha(\theta_t(\omega))$ satisfy (17.4) and (17.5), respectively.

17.1 Properties of Solutions

The existence, uniqueness and boundedness of positive solutions to (17.6)–(17.8) will be established here. In addition, it will be shown that the solution mapping generates a random dynamical system.

Denote by

$$\mathbb{R}_+^3 = \{(x, y, v) \in \mathbb{R}^3 : x \geq 0, y \geq 0, v \geq 0\},$$

and for simplicity write $\mathbf{u}(t, \omega) = (x(t, \omega), y(t, \omega), v(t, \omega))^\top$.

Theorem 17.1 For any $\omega \in \Omega$, $t_0 \in \mathbb{R}$ and initial data $\mathbf{u}_0 = (x(t_0), y(t_0), v(t_0))^\top \in \mathbb{R}_+^3$, system (17.6)–(17.8) has a unique nonnegative bounded solution $\mathbf{u}(\cdot; t_0, \omega, \mathbf{u}_0) \in \mathcal{C}([t_0, \infty), \mathbb{R}_+^3)$, with $\mathbf{u}(t_0; t_0, \omega, \mathbf{u}_0) = \mathbf{u}_0$. Moreover, the solution generates a random dynamical system $\varphi(t, \omega)(\cdot)$ defined as

$$\varphi(t, \omega)\mathbf{u}_0 = \mathbf{u}(t; 0, \omega, \mathbf{u}_0), \quad \forall t \geq 0, \mathbf{u}_0 \in \mathbb{R}_+^3, \omega \in \Omega.$$

Proof Write

$$L(\theta_t(\omega)) = \begin{pmatrix} -\gamma_1 & 0 & 0 \\ 0 & -\gamma_2 & 0 \\ 0 & \kappa & -\alpha(\theta_t(\omega)) \end{pmatrix} \quad \text{and} \quad f(\theta_t(\omega), \mathbf{u}) = \begin{pmatrix} \Lambda(\theta_t(\omega)) - \beta xv \\ \beta xv \\ 0 \end{pmatrix}.$$

Then Eqs. (17.6)–(17.8) become

$$\frac{d\mathbf{u}(t, \omega)}{dt} = L(\theta_t(\omega))\mathbf{u} + f(\theta_t(\omega), \mathbf{u}). \quad (17.9)$$

Since $\alpha(\theta_t(\omega))$ is bounded, the operator L generates an evolution system on \mathbb{R}^3 . Second, since $\Lambda(\theta_t(\omega))$ is continuous with respect to t , function f is continuous with respect to t and locally Lipschitz with respect to \mathbf{u} . Hence system (17.9) has a unique local solution $\mathbf{u}(\cdot; t_0, \omega, \mathbf{u}_0) \in \mathcal{C}([t_0, T], \mathbb{R}^3)$.

By continuity of solutions, each solution has to take value 0 before it reaches a negative value. Notice that

$$\begin{aligned} \left. \frac{dx(t, \omega)}{dt} \right|_{x=0, y \geq 0, v \geq 0} &= \Lambda(\theta_t(\omega)) > 0, \\ \left. \frac{dy(t, \omega)}{dt} \right|_{x \geq 0, y=0, v \geq 0} &= \beta xv \geq 0, \\ \left. \frac{dv(t, \omega)}{dt} \right|_{x \geq 0, y \geq 0, v=0} &= \kappa y \geq 0, \end{aligned}$$

so $x(t)$ is strictly increasing at $x = 0$, while $y(t)$ and $v(t)$ are non-decreasing at $y = 0$ and $v = 0$, respectively. This implies that $\mathbf{u}(t) \in \mathbb{R}_+^3$ for $t \in [t_0, T]$.

For $\mathbf{u}(t) \in \mathbb{R}_+^3$, define

$$\|\mathbf{u}(t)\|_1 := x(t) + y(t) + v(t).$$

Let $s(t) = 2\kappa x(t) + 2\kappa y(t) + \gamma_2 v(t)$, then

$$\|\mathbf{u}(t)\|_1 \leq \frac{s(t)}{\min\{2\kappa, \gamma_2\}}.$$

On the other hand by (17.6)–(17.8),

$$\begin{aligned}
\frac{ds(t, \omega)}{dt} &= 2\kappa A(\theta_t(\omega)) - 2\kappa\gamma_1 x - \kappa\gamma_2 y - \gamma_2\alpha(\theta_t(\omega))v \\
&\leq 2\kappa\lambda(1 + \delta_1) - 2\kappa\gamma_1 x - \kappa\gamma_2 y - \gamma_2a(1 - \delta_2)v \\
&\leq 2\kappa\lambda(1 + \delta_1) - \mu_1 s(t),
\end{aligned} \tag{17.10}$$

where

$$\mu_1 = \min\{\gamma_1, \gamma_2/2, a(1 - \delta_2)\} > 0. \tag{17.11}$$

For $s(t_0) \geq 2\kappa\lambda(1 + \delta_1)/\mu_1$, $s(t)$ will be non-increasing for $t \geq t_0$ and thus $s(t) \leq s(t_0)$. Otherwise, for $s(t_0) \leq 2\kappa\lambda(1 + \delta_1)/\mu_1$, $s(t)$ will stay $\leq 2\kappa\lambda(1 + \delta_1)/\mu_1$. In summary,

$$0 \leq \|\mathbf{u}\|_1 \leq \frac{s(t)}{\min\{2\kappa, \gamma_2\}} \leq \frac{\max\{2\kappa x(t_0) + 2\kappa y(t_0) + \gamma_2 v(t_0), 2\kappa\lambda(1 + \delta_1)/\mu_1\}}{\mu_2},$$

where

$$\mu_2 = \min\{2\kappa, \gamma_2\}. \tag{17.12}$$

This implies that system (17.9) has a unique positive and bounded global solution $\mathbf{u}(\cdot; t_0, \omega, \mathbf{u}_0) \in \mathbb{R}_+^3$.

It is straightforward to check that

$$\mathbf{u}(t + t_0; t_0, \omega, \mathbf{u}_0) = \mathbf{u}(t; 0, \theta_{t_0}(\omega), \mathbf{u}_0)$$

for all $t_0 \in \mathbb{R}$, $t \geq t_0$, $\omega \in \Omega$ and $\mathbf{u}_0 \in \mathbb{R}_+^3$. This allows us to define a mapping $\varphi(t, \omega)(\cdot)$:

$$\varphi(t, \omega)\mathbf{u}_0 = \mathbf{u}(t; 0, \omega, \mathbf{u}_0), \quad \forall t \geq 0, \mathbf{u}_0 \in \mathbb{R}_+^3, \omega \in \Omega. \tag{17.13}$$

Henceforth, $\mathbf{u}(t; 0, \omega, \mathbf{u}_0)$ will be simplified to $\mathbf{u}(t; \omega, \mathbf{u}_0)$.

For any $\mathbf{u}_0 \in \mathbb{R}_+^3$, solution $\mathbf{u}(\cdot; \omega, \mathbf{u}_0) \in \mathbb{R}_+^3$ for $t \in [0, \infty)$. Since the function $f(\mathbf{u}, \theta_t(\omega)) = f(\mathbf{u}, t, \omega)$ is continuous in \mathbf{u} , t , and is measurable in ω , $\mathbf{u} : [0, \infty) \times \Omega \times \mathbb{R}_+^3 \rightarrow \mathbb{R}_+^3$, $(t, \omega, \mathbf{u}_0) \mapsto \mathbf{u}(t; \omega, \mathbf{u}_0)$ is $(\mathcal{B}[0, \infty) \times \mathcal{F}_0 \times \mathcal{B}(\mathbb{R}_+^3), \mathcal{B}(\mathbb{R}_+^3))$ -measurable. It then follows directly that (17.9) generates a continuous random dynamical system $\varphi(t, \omega)(\cdot)$ defined by (17.13). \square

17.2 Existence of Global Random Attractors

The existence of a global random attractor for the random dynamical system (θ, φ) will be established first and then its geometric structures will be investigated.

Theorem 17.2 *The random dynamical system generated by system (17.9) possesses a unique global random attractor $\mathcal{A} = \{A(\omega) : \omega \in \Omega\}$.*

Proof It will first be proved that for $\omega \in \Omega$, there exists a tempered bounded closed random absorbing set $K(\omega) \in \mathcal{K}(\mathbb{R}_+^3)$ of the random dynamical system (θ, φ) such that for any $B \in \mathcal{K}(\mathbb{R}_+^3)$ and each $\omega \in \Omega$, there exists $T_B(\omega) > 0$ for which

$$\varphi(t, \theta_{-t}(\omega))B(\theta_{-t}(\omega)) \subset K(\omega) \quad \forall t \geq T_B(\omega).$$

In fact, recall that $\mathbf{u}(t; \omega, \mathbf{u}_0) = \varphi(t, \omega)\mathbf{u}_0$ denotes the solution of system (17.9) satisfying $\mathbf{u}(0; \omega, \mathbf{u}_0) = \mathbf{u}_0$. Then for any $\mathbf{u}_0 := \mathbf{u}_0(\theta_{-t}(\omega)) \in B(\theta_{-t}(\omega))$,

$$\|\varphi(t, \theta_{-t}(\omega))\mathbf{u}_0\|_1 = \|\mathbf{u}(t; \theta_{-t}(\omega), \mathbf{u}_0(\theta_{-t}(\omega)))\|_1 \leq \frac{1}{\mu_2} \cdot s(t; \theta_{-t}(\omega), s_0(\theta_{-t}(\omega))).$$

Using inequality (17.10) and substituting ω by $\theta_{-t}(\omega)$,

$$\begin{aligned} s(t; \theta_{-t}(\omega), s_0(\theta_{-t}(\omega))) s_0 &\leq e^{-\mu_1 t} + \frac{2\kappa\lambda(1 + \delta_1)}{\mu_1} \\ &\leq e^{-\mu_1 t} \sup_{(x, y, v) \in B(\theta_{-t}(\omega))} (2\kappa x + 2\kappa y + \gamma_2 v) + \frac{2\kappa\lambda(1 + \delta_1)}{\mu_1}. \end{aligned}$$

Therefore for any $\varepsilon > 0$, and $\mathbf{u}_0 \in B(\theta_{-t}(\omega))$, there exists $T_B(\omega)$ such that when $t > T_B$,

$$\begin{aligned} \|\varphi(t, \theta_{-t}(\omega))\mathbf{u}_0\|_1 &\leq \frac{1}{\mu_2} \cdot s(t; \theta_{-t}(\omega), s_0(\theta_{-t}(\omega))) \\ &\leq \frac{1}{\mu_2} \cdot \frac{2\kappa\lambda(1 + \delta_1)}{\mu_1} + \varepsilon, \end{aligned}$$

Define

$$K_\varepsilon(\omega) = \left\{ (x, y, v) \in \mathbb{R}_+^3 : x + y + v \leq \frac{1}{\mu_2} \cdot \frac{2\kappa\lambda(1 + \delta_1)}{\mu_1} + \varepsilon \right\}. \quad (17.14)$$

Then $K_\varepsilon(\omega)$ is positively invariant and absorbing in \mathbb{R}_+^3 .

It follows directly from Theorem 4.2 of Chap. 4 that the random dynamical system generated by system (17.6)–(17.8) possesses a random attractor $\mathcal{A} = \{A(\omega) : \omega \in \Omega\}$, consisting of nonempty compact random subsets of \mathbb{R}_+^3 contained in $K_\varepsilon(\omega)$.

The details of the random attractor \mathcal{A} will now be investigated.

Theorem 17.3 *The random attractor $\mathcal{A} = \{A(\omega) : \omega \in \Omega\}$ for the random dynamical system generated by system (17.6)–(17.8) has singleton component sets $A(\omega) = \{(x^*(\omega), 0, 0)\}$ for every $\omega \in \Omega$, provided that*

$$\frac{\kappa}{\gamma_2} \leq 1 \quad \text{and} \quad \frac{\beta\lambda(1+\delta_1)}{\mu_1 a(1-\delta_2)} < 1. \quad (17.15)$$

Proof Summing (17.7) and (17.8) yields

$$\frac{d(y+v)}{dt} = -(\gamma_2 - \kappa)y - (\alpha(\theta_t(\omega)) - \beta x)v.$$

Recall that due to (17.14), for any $\varepsilon > 0$, there exists $T_B(\omega)$ such that when $t > T_B$,

$$x(t) \leq \|\mathbf{u}(t)\|_1 \leq \frac{1}{\mu_2} \cdot \frac{2\kappa\lambda(1+\delta_1)}{\mu_1} + \varepsilon.$$

By the definition of μ_2 in (17.12) it follows that $2\kappa/\mu_2 \leq 1$. Then, picking ε small enough,

$$\begin{aligned} \alpha(\theta_t(\omega)) - \beta x &> \alpha(1 - \delta_2) - \beta \cdot \frac{1}{\mu_2} \cdot \frac{2\kappa\lambda(1+\delta_1)}{\mu_1} \\ &\geq \alpha(1 - \delta_2) - \beta \cdot \frac{\lambda(1+\delta_1)}{\mu_1} > 0, \end{aligned}$$

which implies that $y + v$ decreases to 0 as t approaches ∞ .

Letting $y = v = 0$ in Eq. (17.6) gives

$$\frac{dx}{dt} = \Lambda(\theta_t(\omega)) - \gamma_1 x. \quad (17.16)$$

The solution of Eq. (17.16) is

$$x(t; \omega, x_0) = x_0 e^{-\gamma_1 t} + \int_0^T \Lambda(\theta_s(\omega)) e^{\gamma_1(t-s)} ds,$$

so

$$\begin{aligned} x(t; \theta_{-t}(\omega), x_0) &= x_0 e^{-\gamma_1 t} + \int_{-t}^0 \Lambda(\theta_s(\omega)) e^{-\gamma_1 s} ds \\ &\xrightarrow{t \rightarrow \infty} \int_{-\infty}^0 \Lambda(\theta_s(\omega)) e^{-\gamma_1 s} ds := x^*(\omega). \end{aligned}$$

This completes the proof. \square

Theorem 17.3 implies that $(x^*(\theta_t(\omega)), 0, 0)$ is asymptotically stable as $t \rightarrow \infty$, i.e., an endemic occurs when the parameters satisfy (17.15).

Conditions under which an epidemic occurs will now be established.

Theorem 17.4 *The random pullback attractor $\mathcal{A} = \{A(\omega) : \omega \in \Omega\}$ for the random dynamical system generated by system (17.6)–(17.8) possesses nontrivial component*

sets which include $(x^*(\omega), 0, 0)$ and strictly positive points provided that

$$\frac{\beta\lambda(1 + \delta_1)}{\mu_1 a(1 + \delta_2)} > \frac{\gamma_2}{\kappa}. \quad (17.17)$$

Proof First notice that the Eq.(17.7) is deterministic, and implies that the surface $y = \frac{\beta}{\gamma_2}xv$ is invariant. The dynamics of x and v restricted on this invariant surface satisfy

$$\begin{aligned} \frac{dx(t, \omega)}{dt} &= \Lambda(\theta_t(\omega)) - \gamma_1 x - \beta xv, \\ \frac{dv(t, \omega)}{dt} &= \frac{\kappa\beta}{\gamma_2}xv - \alpha(\theta_t(\omega))v. \end{aligned}$$

Define the region Γ_ε by

$$\Gamma_\varepsilon := \left\{ (x, v) \in \mathbb{R}_+^2 : x \geq \frac{a(1 + \delta_2)\gamma_2}{\kappa\beta} + \varepsilon, v \geq \varepsilon, \frac{\kappa}{\gamma_2}x(t) + v(t) \leq \frac{\kappa\lambda}{\mu_1\gamma_2}(1 + \delta_1) + \varepsilon \right\}.$$

For any $(x, v) \in \Gamma_\varepsilon$,

$$\frac{dv}{dt} = \left(\frac{\kappa\beta}{\gamma_2}x - \alpha(\theta_t(\omega)) \right) v > \left(\frac{\kappa\beta}{\gamma_2} \cdot \frac{a(1 + \delta_2)\gamma_2}{\kappa\beta} - a(1 + \delta_2) \right) v \geq 0.$$

On the other hand,

$$\begin{aligned} \frac{d}{dt} \left(\frac{\kappa}{\gamma_2}x(t) + v(t) \right) &= \frac{\kappa}{\gamma_2}\Lambda(\theta_t(\omega)) - \gamma_1 \frac{\kappa}{\gamma_2}x - \alpha(\theta_t(\omega))v \\ &\leq \frac{\kappa\lambda}{\gamma_2}(1 + \delta_1) - \gamma_1 \frac{\kappa}{\gamma_2}x - a(1 + \delta_2)v \\ &\leq \frac{\kappa\lambda}{\gamma_2}(1 + \delta_1) - \mu_1 \left(\frac{\kappa}{\gamma_2}x(t) + v(t) \right), \end{aligned}$$

where μ_1 is as defined in (17.11). This implies that

$$\frac{\kappa}{\gamma_2}x(t) + v(t) \leq \frac{\kappa\lambda}{\mu_1\gamma_2}(1 + \delta_1) + \varepsilon$$

for t large enough. Assumption (17.17) ensures that Γ_ε is a nonempty compact positive invariant absorbing set, which then ensures the existence of a nontrivial random attractor $\mathcal{A}_\varepsilon = \{A_\varepsilon(\omega) : \omega \in \Omega\}$ in Γ_ε . \square

17.3 Numerical Simulations

The virus system (17.6)–(17.8) will be simulated numerically here and it will be verified that conditions (17.15) and (17.17) give rise to an endemic state (all infected cells and viruses are cleared) and a pandemic state (susceptible cells, infected cells, and viruses co-exist) of system (17.6)–(17.8), respectively.

First, the system (17.6)–(17.8) with two OU processes $O_1(t)$ and $O_2(t)$ is transformed into a RODE-SODE pair:

$$d \begin{pmatrix} x(t) \\ y(t) \\ v(t) \\ O_1(t) \\ O_2(t) \end{pmatrix} = \begin{pmatrix} \Lambda(O_1) - \gamma_1 x - \beta xv \\ \beta xv - \gamma_2 y \\ \kappa y - \alpha(O_2) \\ \theta_{11} - \theta_{12} O_1 \\ \theta_{21} - \theta_{22} O_2 \end{pmatrix} dt + \begin{pmatrix} 0 \\ 0 \\ 0 \\ \theta_{13} \\ \theta_{23} \end{pmatrix} dW_t.$$

The OU processes $O_1(t)$ and $O_2(t)$ can be generated independently, so only the RODE part, i.e., x , y and v compartments, of the RODE-SODE system, needs to be discretised. The system is usually stiff, so the implicit order 1.5 RODE-Taylor scheme (10.14) in Chap. 10 is applied here.

In the following simulation, the cell proliferation rate $\Lambda(O_1)$ is assumed to have a switching effect and the loss rate of viruses $\alpha(O_2)$ is distributed over a finite interval. They are randomised by the equations (1.6) and (1.7) of Chap. 1, respectively, and given by

$$\Lambda(O_1) = \lambda \left(1 - 2\delta_1 \frac{O_1}{1 + O_1^2} \right), \quad \alpha(O_2) = a \left(1 - \frac{2\delta_2}{\pi} \arctan O_2 \right).$$

They satisfy (17.4) and (17.5).

Initial conditions for the x , y and v compartments are set as $x_0 = 2 \times 10^5$, $y_0 = 1 \times 10^5$ and $v_0 = 3 \times 10^5$. The coefficients for the OU processes are fixed to $\theta_{11} = 1$, $\theta_{12} = 3$, $\theta_{13} = 0.8$, $\theta_{21} = 0$, $\theta_{22} = 1$ and $\theta_{23} = 0.5$ for all examples. A different set of parameters that satisfy assumption (17.15) or assumption (17.17) will be chosen.

Example 1

In this example the parameters are set to be $\gamma_1 = 0.25$, $\gamma_2 = 0.5$, $\beta = 1 \times 10^{-5}$, $\lambda = 4 \times 10^4$, $a = 3$, $\delta_1 = 0.45$, $\delta_2 = 0.2$, and $\kappa = 0.2$. Assumptions in (17.15) are satisfied by this set of parameters. Figure 17.1 shows that the y and v compartments go to zero after enough amount of time and only x compartment remains non zero, which means that the endemic state is achieved for parameters satisfying (17.15).

Example 2

In this example the parameters are set to be $\gamma_1 = 0.25$, $\gamma_2 = 0.5$, $\beta = 1 \times 10^{-5}$, $\lambda = 4 \times 10^4$, $a = 3$, $\delta_1 = 0.45$, $\delta_2 = 0.2$, and $\kappa = 2$. Assumptions in (17.17) are satisfied by this set of parameters. Figure 17.2 shows that x , y and v all remain nonzero for a

Fig. 17.1 With parameters $\gamma_1 = 0.25$, $\gamma_2 = 0.5$, $\beta = 1 \times 10^{-5}$, $\lambda = 4 \times 10^4$, $a = 3$, $\delta_1 = 0.45$, $\delta_2 = 0.2$, and $\kappa = 0.2$ satisfying assumption (17.15), both infect cells and viruses are cleared; only healthy cells remain

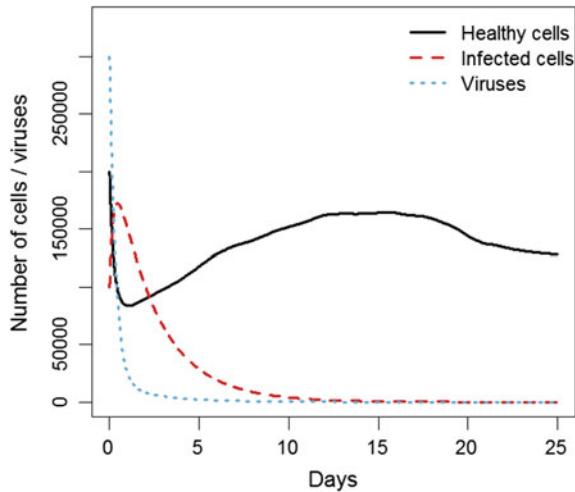
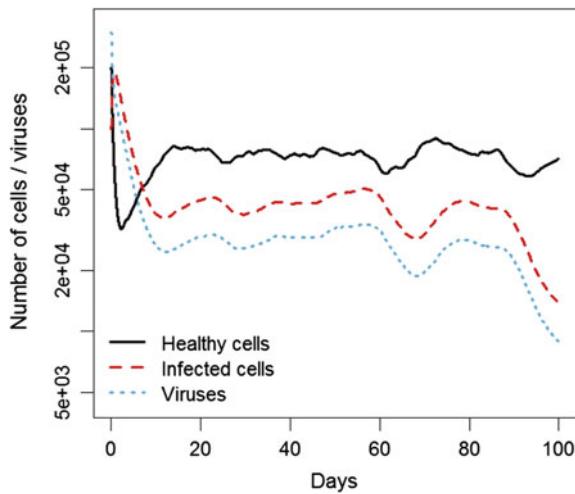


Fig. 17.2 With parameters $\gamma_1 = 0.25$, $\gamma_2 = 0.5$, $\beta = 1 \times 10^{-5}$, $\lambda = 4 \times 10^4$, $a = 3$, $\delta_1 = 0.45$, $\delta_2 = 0.2$, and $\kappa = 2$ satisfying assumption (17.17), infected cells, susceptible cells and viruses coexist



time long enough, which means that the pandemic state is achieved for parameters satisfying (17.17).

Example 3

Notice that the only parameter that has different values in Example 1 and Example 2 is κ . This implies that the rate at which virus is generated by dead susceptible cells is critical. A series of numerical simulations with different parameters were done to support this argument. One more example is presented below in which the parameters are chosen to be $\gamma_1 = 0.4$, $\gamma_2 = 0.5$, $\beta = 5 \times 10^{-5}$, $\lambda = 10^5$, $a = 5$, $\delta_1 = 0.4$, $\delta_2 = 0.2$. When $\kappa = 0.3$, assumption (17.15) is satisfied and an endemic

Fig. 17.3 With parameters $\gamma_1 = 0.4$, $\gamma_2 = 0.5$, $\beta = 5 \times 10^{-5}$, $\lambda = 10^5$, $a = 5$, $\delta_1 = 0.4$, $\delta_2 = 0.2$ and $\kappa = 0.1$ satisfying assumption (17.15), both infect cells and viruses are cleared; only healthy cells remain

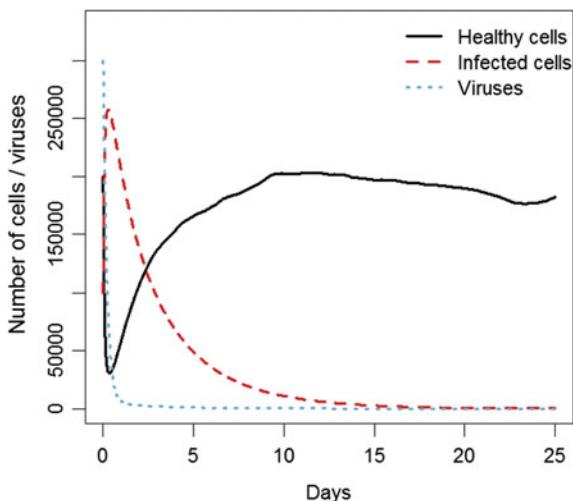
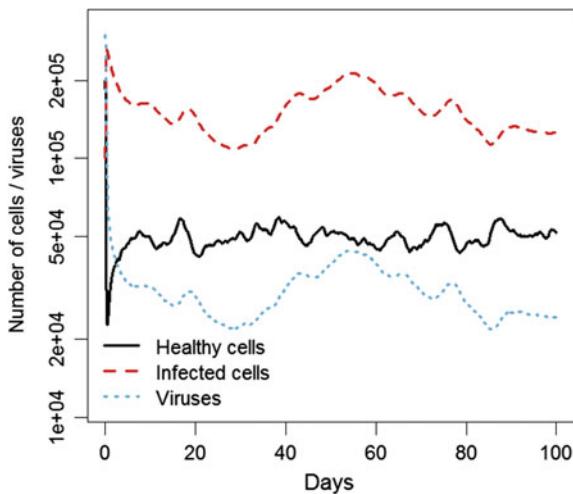


Fig. 17.4 With parameters $\gamma_1 = 0.4$, $\gamma_2 = 0.5$, $\beta = 5 \times 10^{-5}$, $\lambda = 10^5$, $a = 5$, $\delta_1 = 0.4$, $\delta_2 = 0.2$ and $\kappa = 1$ satisfying assumption (17.17), infected cells, susceptible cells and viruses coexist



state is obtained (see Fig. 17.3). When $\kappa = 3$, assumption (17.17) is satisfied and a pandemic state is obtained (see Fig. 17.4).

17.4 Endnotes

This chapter is based on Asai, Caraballo, Han and Kloeden [7]. Virus dynamics models are discussed in May and Nowak [103] and an HIV model is presented by Perelson and Ribeiro [113].

Chapter 18

Random Markov Chains

Markov chains can be used to model the random environments and other random factors in biological systems, as often the switching between different environments is memoryless and the waiting time for the next switch is exponentially distributed [145]. In particular, Markov chains with tridiagonal transition matrices are common in biological models, for example, birth-and-death processes, cell-cell communication and cancer dynamics, to name just a few. As an illustrative example, consider the distance $d(t_n)$ between two cells at time $t_n = nh$, which is supposed to take discrete values in $\{1, \dots, N\}$, essentially the distance that they can move in one unit of time, where $d(t_n)$ can stay unchanged or change to $d(t_n) \pm 1$ with certain probabilities. This can be formulated as an N state discrete-time Markov chain with states $\{1, \dots, N\}$ corresponding to the value of $d(t_n)$.

Let $\mathbf{p}(t_n) = (p_1(t_n), \dots, p_N(t_n))^\top$ be the probability vector for the state of the system at time t_n . The dynamics are described by the system of difference equations

$$\begin{aligned} p_1(t_{n+1}) &= [1 - q_1 h] p_1(t_n) + q_2 p_2(t_n) h, \\ p_j(t_{n+1}) &= q_{2j-3} p_{j-1}(t_n) h + [1 - (q_{2j-2} + q_{2j-1}) h] p_j(t_n) + q_{2j} p_{j+1}(t_n) h, \\ &\quad j = 2, \dots, N-1, \\ p_N(t_{n+1}) &= q_{2N-3} p_{N-1}(t_n) h + [1 - q_{2N-2} h] p_N(t_n), \end{aligned}$$

where the $q_i \in [0, 1]$ for $i = 1, \dots, 2N-2$, $h > 0$ and p_j 's satisfy the probability constraints

$$\sum_{j=1}^N p_j = 1, \quad p_j \geq 0, \quad j = 1, \dots, N.$$

This is a vector-valued difference equation

$$\mathbf{p}(t_{n+1}) = [I_N + h Q_N] \mathbf{p}(t_n)$$

on the simplex Σ^N in \mathbb{R}^N defined by

$$\Sigma^N = \left\{ \mathbf{p} = (p_1, \dots, p_N)^\top : \sum_{j=1}^N p_j = 1, p_1, \dots, p_N \in [0, 1] \right\},$$

where I_N is the $N \times N$ identity matrix and Q_N is the tridiagonal $N \times N$ -matrix

$$Q_N = \begin{pmatrix} -q_1 & q_2 & & & & & & \circlearrowleft \\ q_1 & -(q_2 + q_3) & q_4 & & & & & \\ & \ddots & \ddots & \ddots & \ddots & & & \\ & & & & & \ddots & & \\ & & & & & & -q_{2N-4} - (q_{2N-3} + q_{2N-2}) & q_{2N-2} \\ & & & & & & q_{2N-3} & -q_{2N-2} \end{pmatrix} \quad (18.1)$$

Note that $\mathbf{1}_N^\top Q_N = \mathbf{0}$, were $\mathbf{1}_N$ is the vector in \mathbb{R}^N with all components equal to 1.

This generates a discrete-time finite-state Markov chain

$$\mathbf{p}^{(n+1)} = L_N \mathbf{p}^{(n)} \quad (18.2)$$

with the transition matrix $L_N := [I_N + h Q_N]$. It is a first order linear difference equation on Σ^N and corresponds to the Euler numerical scheme for the ordinary algebraic-differential equation

$$\frac{d\mathbf{p}}{dt} = Q_N \mathbf{p}, \quad \mathbf{p} \in \Sigma^N, \quad (18.3)$$

with the constant time step $h > 0$.

It can be shown that each of the Markov chains (18.2) and (18.3) has a unique nontrivial equilibrium solution, which is globally asymptotically stable provided $h < 1/(2b)$ in the discrete-time case, where b is the maximum of the q_i .

18.1 Random Environment

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and suppose now that the coefficients in the Q_N matrix are random, i.e., the $q_j : \Omega \rightarrow \mathbb{R}$ are \mathcal{F} -measurable mappings or, equivalently, $Q_N : \Omega \rightarrow \mathbb{R}^{N \times N}$ is an \mathcal{F} -measurable $N \times N$ matrix-valued mapping.

This corresponds to a random environment, which is supposed to vary or be driven by a stochastic process modeled by a metrical (i.e., measurable) dynamical system $\theta = \{\theta_t, t \in \mathbb{T}\}$ ¹ on Ω generated by a bi-measurable invertible mapping $\theta : \Omega \rightarrow \Omega$. In particular, θ satisfies $\theta_0(\omega) = \omega$ and

¹ $\mathbb{T} = \mathbb{Z}$ in the discrete-time case and \mathbb{R} in the continuous-time case.

$$\theta_{s+t}\omega \equiv \theta_s(\theta_t(\omega)), \quad \forall s, t \in \mathbb{T}, \omega \in \Omega.$$

In the continuous-time case $(t, \omega) \mapsto \theta_t(\omega)$ is also assumed to be measurable. It forms the driving system of the random dynamical system, see Chap. 4.

Define $L_N(\omega) := I_N + h Q_N(\omega)$. This gives the *random Markov chain*

$$\mathbf{p}^{(n+1)} = L_{\theta_n(\omega)} \mathbf{p}^{(n)}, \quad (18.4)$$

which is a random linear difference equation on Σ^N . The iterates of (18.4) are random probability vectors in Σ^N , i.e., \mathcal{F} -measurable mappings $\mathbf{p} : \Omega \rightarrow \Sigma^N$.

The continuous-time case yields the linear RODE

$$\frac{d\mathbf{p}}{dt} = Q_N(\theta_t(\omega)) \mathbf{p}, \quad \mathbf{p}(0) = \mathbf{p}_0 \in \mathbb{R}^N. \quad (18.5)$$

The iterates of the linear random difference equation (18.4) and the solution mapping of the linear RODE (18.5) each form a linear cocycle mapping φ on \mathbb{R}^N . In the continuous-time case this is defined by

$$\varphi(t, \omega, \mathbf{p}_0) = \mathbf{p}(t, \omega; \mathbf{p}_0) \quad \text{for all } t \geq 0, \omega \in \Omega, \mathbf{p}_0 \in \mathbb{R}^N.$$

It will be shown that the probability simplex Σ^N is positively invariant under these linear cocycle mappings, provided the coefficients of the random matrix $Q(\omega)$ are strictly positive and bounded. In particular, they are assumed to satisfy:

Assumption 18.1.1

$$a \leq q_i(\omega) \leq b, \quad \omega \in \Omega, \quad i = 1, 2, \dots, 2N - 2,$$

hold for some $a, b > 0$.

In the discrete-time Markov chain case, it is also assumed that the parameter $h < 1/(2b)$. This suffices to ensure the positivity of the solutions of discrete-time Markov chain (18.4) and the positive invariance of the probability simplex Σ^N under it.

18.2 Positivity of Solutions of Linear RODEs

Unfortunately, the quasipositive property of Chap. 4 is rather restrictive and does not apply to the deterministic linear ODE (18.3) or the linear RODE (18.5) with a tridiagonal matrix Q_N given by (18.1).

There is, however, another useful notion that can be used in combination with it. An $N \times N$ matrix A with non-negative off-diagonal elements is said to have a *path of nonsingularity* if there is a set of indices i_1, i_2, \dots, i_n with $i_n = i_1$ such that all the

elements $a_{i_j, i_{j+1}}$ are strictly positive and this set of indices contains all the numbers $1, 2, \dots, N$. In particular, the matrix A defined by (18.1) has the path of nonsingularity:

$$\{i_1, i_2, \dots, i_{2N-1}\} = \{1, 2, \dots, N-1, N, N-1, \dots, 2, 1\}.$$

Recall that a set K in a Banach space is called a *cone* if it is convex, closed with $tK \subseteq K$ for any real $t \geq 0$ and $K \cap (-K) = \{0\}$. Fix a norm $\|\cdot\|$ in \mathbb{R}^N and denote by \mathbb{K}_+^N the cone of elements $x = (x_1, x_2, \dots, x_N)^\top \in \mathbb{R}^N$ with nonnegative components and by $\overset{\circ}{\mathbb{K}}_+^N$ the interior of \mathbb{K}_+^N , which is clearly non-empty. In addition, denote by \mathbb{B}^N the unit ball in the norm $\|\cdot\|$.

Theorem 18.1 *Suppose that the matrix Q_N in the linear ODE (18.3) or RODE (18.5) with nonnegative off-diagonal entries has a path of nonsingularity. Then, for any non-zero initial condition $\mathbf{p}(0) = \mathbf{p}_0 \in \mathbb{K}_+^N$, the solution $\mathbf{p}(t, \mathbf{p}_0)$ is strongly positive for all $t > 0$.*

Moreover, for every bounded interval $[T_1, T_2] \subset (0, \infty)$, there is a number $c(T_1, T_2) > 0$ such that

$$\mathbf{p}(t; \mathbf{p}_0) + c(T_1, T_2) \|\mathbf{p}_0\| \mathbb{B}^N \subseteq \overset{\circ}{\mathbb{K}}_+^N \quad (18.6)$$

for all $\mathbf{p}_0 \in \mathbb{K}_+^N \setminus \{0\}$ and $t \in [T_1, T_2]$.

Proof The first part of Theorem 18.1 follows from [95, Theorem 4.7], while the second part is clear. \square

Remark 18.1 The inclusion (18.6) means that the solution operator $\mathbf{p}_0 \mapsto \mathbf{p}(t, \mathbf{p}_0)$ maps the set $\mathbb{K}_+^N \setminus \{0\}$ into its interior $\overset{\circ}{\mathbb{K}}_+^N$ for every $t > 0$. In addition, if the matrix Q_N satisfies the conditions of Theorem 18.1 as well as $\mathbf{1}_N^\top Q_N = \mathbf{0}$, then $\mathbf{p}(t; \Sigma^N) \subseteq \Sigma^N$ for all $t \geq 0$. Moreover, the simplex Σ^N is mapped by $\mathbf{p}(t; \cdot)$ into its interior $\overset{\circ}{\Sigma}_N$ for every $t > 0$.

In order to apply Theorem 18.1 to the linear RODE (18.5) define a (deterministic) tridiagonal matrix

$$\bar{Q} = \begin{pmatrix} \tilde{q}_1 & \tilde{q}_2 & & & & & \bigcirc & \\ \tilde{q}_1 & \tilde{q}_2 & \tilde{q}_4 & & & & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ & & & \ddots & \ddots & \ddots & & \\ & & & & \tilde{q}_{2N-5} & \tilde{q}_{N-1} & \tilde{q}_{2N-2} & \\ & & & & & \tilde{q}_{2N-3} & \tilde{q}_N & \\ \bigcirc & & & & & & & \end{pmatrix}$$

with the off-diagonal elements \tilde{q}_i given by

$$\tilde{q}_i = \inf_{\omega \in \Omega} q_i(\omega), \quad i = 1, 2, \dots, 2N-2,$$

and the diagonal elements \tilde{q}_i satisfying

$$\begin{aligned}\tilde{q}_1 &= \inf_{\omega \in \Omega} \{-q_1(\omega)\}, & \tilde{q}_N &= \inf_{\omega \in \Omega} \{-q_{2N-2}(\omega)\} \\ \tilde{q}_i &= \inf_{\omega \in \Omega} \{-(q_{2i-2}(\omega) + q_{2i-1}(\omega))\}, & i &= 2, \dots, 2N-1.\end{aligned}$$

Then

$$Q_N(\omega) \geq \bar{Q}, \quad \text{for all } \omega \in \Omega,$$

where the inequality between the matrices is interpreted componentwise. The off-diagonal elements \tilde{q}_i of the matrix \bar{Q} are strictly positive since

$$0 < c_* \leq \tilde{q}_i, \quad i = 1, 2, \dots, 2N-2. \quad (18.7)$$

Now consider the autonomous ordinary differential equation

$$\frac{d\pi}{dt} = \bar{Q}\pi \quad (18.8)$$

and denote the solution of this equation with the initial condition $\pi(0) = \pi_0 \in \mathbb{R}^N$ by $\pi(t; \pi_0)$. It follows from (18.7) that \bar{Q} is a matrix with non-negative off-diagonal elements that has a path of nonsingularity. Then by Theorem 18.1 and Remark 18.1, the solution operator $\pi(t; \cdot)$ maps the set $\mathbb{K}_+^N \setminus \{0\}$ into its interior $\overset{\circ}{\mathbb{K}}_+^N$ for each $t > 0$ maps. Moreover, properties like (18.6) hold for it with appropriate parameters.

Theorem 18.2 *Let θ a metric dynamical system and let $Q_N(\omega)$ be a matrix of the tridiagonal form (18.1), which satisfies Assumption 18.1.1. Then, the solution $p(t; \omega, p_0)$ of the linear RODE (18.5) satisfies*

$$p(t; \omega, p_0) \geq \pi(t; p_0) \quad \text{for all } \omega \in \Omega, t \geq 0, p_0 \in \mathbb{K}_+^N,$$

where the inequality is meant componentwise.

Proof Fix an $\omega \in \Omega$ and fix the initial conditions $p(0) = \pi(0) = p_0 \in \mathbb{K}_+^N$ for the differential equations (18.5) and (18.8), respectively. Then the function

$$x(t) := p(t; \omega, p_0) - \pi(t; p_0)$$

satisfies the differential equation

$$\frac{dx}{dt} = Q_N(\theta_t(\omega))p - \bar{Q}\pi = Q_N(\theta_t(\omega))x + (Q_N(\theta_t(\omega)) - \bar{Q})\pi$$

with the initial condition $x(0) = 0$. Denoting

$$w(t) := (Q_N(\theta_t(\omega)) - \bar{Q})\pi(t; p_0),$$

this differential equation can be written as

$$\frac{dx}{dt} = Q_N(\theta_t(\omega))x + w(t), \quad x(0) = 0. \quad (18.9)$$

The matrix $(Q_N(\theta_t(\omega)) - \bar{Q})$ has, by definition, non-negative components for every t and ω . Now the function $\pi(t; p_0)$ with $p_0 \in \mathbb{K}_+^N$ has non-negative components for all $t \geq 0$. Hence the function $w(t)$ has non-negative components. Moreover, the matrix $Q_N(\omega)$ as a matrix with non-negative off-diagonal components has a path of nonsingularity, so it follows that the differential equation (18.9) satisfies all the conditions of Theorem 18.1. Hence, the solution function $x(t)$ is positive, i.e., its components are non-negative. \square

Theorem 18.2 has some important consequences.

Corollary 18.1 *The following statements are valid:*

- (i) *for any $\omega \in \Omega$ and $t > 0$, the solution operator $p(t; \omega, \cdot)$ maps the set $\mathbb{K}_+^N \setminus \{0\}$ into its interior $\overset{\circ}{\mathbb{K}}_+^N$;*
- (ii) *for any $\omega \in \Omega$ and $t > 0$, the solution operator $p(t; \omega, \cdot)$ maps the simplex Σ^N into its interior $\overset{\circ}{\Sigma}_N$;*
- (iii) *for any bounded interval $[T_1, T_2] \subset (0, \infty)$ there is a number $C(T_1, T_2) > 0$ such that*

$$p(t; \omega, p_0) + C(T_1, T_2) \|p_0\| (\mathbb{B}^N \cap \Sigma^N) \subseteq \Sigma^N,$$

for all $\omega \in \Omega$, $p_0 \in \mathbb{K}_+^N \setminus \{0\}$, $t \in [T_1, T_2]$.

18.3 Linear Random Dynamical Systems

The RODE (18.5) generates a random dynamical system (θ, φ) on \mathbb{R}^N with a linear cocycle mapping φ . It follows from Theorem 18.2 and Corollary 18.1 that the solution operator $p(t; \omega, \cdot)$ of the RODE (18.5) maps the simplex Σ^N into its interior $\overset{\circ}{\Sigma}_N$ uniformly for t in bounded intervals from $(0, \infty)$.

The existence of a random attractor with singleton component sets is established by showing that the linear cocycle mapping is uniformly dissipative and uniformly contracting (see 4.18 in Chap. 4) in an appropriate metric. This metric is given by the Hilbert projective metric, which is sometimes called the Birkhoff metric.

The *Hilbert projective metric* ρ_H on a cone \mathbb{K}_+^N is defined as

$$\rho_H(x, y) = \left| \ln \left(\frac{\max_i y_i / x_i}{\max_i x_i / y_i} \right) \right|$$

for vectors $x = (x_1, x_2, \dots, x_N)^\top$ and $y = (y_1, y_2, \dots, y_N)^\top$ in \mathbb{K}_+^N . It is in fact only a semi-metric on \mathbb{K}_+^N , but becomes a metric on a projective space. Important here is that the interior $\overset{\circ}{\Sigma}_N$ of the probability simplex Σ^N is the complete metric space with the Hilbert projective metric.

The next theorem is proved in Kloeden and Kozyakin [85, 88]. The proof follows by an application of the proof of the corresponding result for the discrete-time case in [85] to the time-one mapping of the RODE, i.e., the cocycle mapping at integer time values. The uniform dissipativity of the cocycle ensures the existence of a random attractor, while the uniform contractivity implies that the components subsets are singleton sets.

Theorem 18.3 *The restriction of $p(t; \omega, \cdot)$ to the set Σ^N is a uniformly dissipative and uniformly contractive cocycle (with respect to the Hilbert projective metric), which has a random attractor $\mathcal{A} = \{A(\omega), \omega \in \Omega\}$ such that set $A(\omega) = \{a(\omega)\}$ consists of a single point for each $\omega \in \Omega$. Moreover, the random attractor is asymptotically stable with respect to the Hilbert projective metric., i.e.,*

$$\rho_H(p(t; \omega, p_0), a(\theta_t(\omega))) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

for all $p_0 \in \overset{\circ}{\Sigma}_N$ and $\omega \in \Omega$.

Since convergence in the Hilbert projective metric implies convergence in any norm on \mathbb{R}^N the random attractor is, asymptotically stable with respect to any norm $\|\cdot\|$ on \mathbb{R}^N , i.e.,

$$\|p(t; \omega, p_0) - a(\theta_t(\omega))\| \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

for all $p^{(0)} \in \Sigma^N$ and $\omega \in \Omega$. The random attractor is, in fact, asymptotic Lyapunov stable in the conventional forward sense.

18.4 Random Attractor Under Discretisation

A general result in numerical dynamics says that the attractor of a numerical scheme converges upper semi continuously in the Hausdorff metric to the attractor of the differential equation. This means that the numerical attractors could collapse into a subset of the limiting attractor. However, this cannot happen when the latter attractor is a singleton set, in which case the convergence is continuous in the Hausdorff metric. This is what happens for the above random Markov chains, where the discrete-time Markov chain is, in fact, the Euler numerical approximation of the corresponding RODE in the continuous-time case. A direct proof will be given here.

Theorem 18.4 *Let $\mathcal{A}^h = \{\{a^h(\omega)\}, \omega \in \Omega\}$ be the random attractor for the discrete-time random Markov chain (18.4) and let $\mathcal{A} = \{\{a(\omega)\}, \omega \in \Omega\}$ be the random attractor for the continuous-time random Markov chain RODE (18.5). Then*

$$\|a^h(\omega) - a(\omega)\| \rightarrow 0 \quad \text{as } h \rightarrow 0, \quad \forall \omega \in \Omega. \quad (18.10)$$

Proof Suppose that the convergence (18.10) does not hold. Then there is an $\omega \in \Omega$, an $\varepsilon_0 > 0$, a sequence $h_n \rightarrow 0$ as $n \rightarrow \infty$ and an $N > 0$ such that

$$\|a^{h_n}(\omega) - a(\omega)\| \geq \varepsilon_0 \quad \text{for } n \geq N. \quad (18.11)$$

Since the continuous-time cocycle uniformly attracts points in Σ^N , there is $T = T(\varepsilon_0)$, which is independent of n , such that

$$\|\mathbf{p}(T, a^{h_n}(\theta_{-T}(\omega)) - a(\omega)\| \leq \frac{1}{4}\varepsilon_0.$$

Taking this T and N_n so that $N_n h_n = T$, the global discretisation error of the Euler scheme (see Chap. 9) applied on the time interval $[-T, 0]$ gives

$$\begin{aligned} \left\| \mathbf{p}_{N_n}^{h_n} (a^{h_n}(\theta_{-T}(\omega))) - \mathbf{p}(T, a^{h_n}(\theta_{-T}(\omega))) \right\| &= \|a^{h_n}(\omega) - \mathbf{p}(T, a^{h_n}(\theta_{-T}(\omega)))\| \\ &\leq \frac{1}{4}\varepsilon_0 \end{aligned}$$

for all n large enough, where $N_n h_n = T$.

Combining the two results gives

$$\begin{aligned} \|a^{h_n}(\omega) - a(\omega)\| &\leq \|a^{h_n}(\omega) - \mathbf{p}(T, a^{h_n}(\theta_{-T}(\omega)))\| + \|\mathbf{p}(T, a^{h_n}(\theta_{-T}(\omega))) - a(\omega)\| \\ &\leq \frac{1}{4}\varepsilon_0 + \frac{1}{4}\varepsilon_0 < \varepsilon_0, \end{aligned}$$

which contradicts the assumption (18.11). □

18.5 Endnotes

This chapter is based on Kloeden and Kozyakin [85, 88]. See [4, 31, 69] for random difference equations and Chueshov [31] for monotone random dynamical systems. Biological applications with tridiagonal Markov chains arise in, for example, birth-and-death processes [2], cell-cell communication [66] and cancer dynamics [139]. Cones and the Hilbert projective metric are discussed in Krasnosel'skij, Lifshits and Sobolev [96] and Nussbaum [110]. See Krasnosel'skij [95] and Smith [121] for properties of monotone ODE.

Stuart and Humphries [128] is the standard reference book on numerical dynamics. See also Kloeden and Lorenz [89] and Chap. 5.

Appendix A

Probability Spaces

Let Ω denote a sample space, i.e., a set of all possible events in an experiment. A collection \mathcal{F} of subsets of Ω is called a σ -algebra (or σ -field) if

- (i) $\emptyset \in \mathcal{F}$,
- (ii) if $F_1, F_2, \dots \in \mathcal{F}$ then $\cup_{i=1}^{\infty} F_i \in \mathcal{F}$,
- (iii) if $F \in \mathcal{F}$ then $F^C := \Omega \setminus F \in \mathcal{F}$.

The pair (Ω, \mathcal{F}) is called a measurable space. A probability measure \mathbb{P} on a measurable space (Ω, \mathcal{F}) is a function $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ such that

- (i) $\mathbb{P}(\emptyset) = 0, \mathbb{P}(\Omega) = 1$,
- (ii) if $F_1, F_2, \dots \in \mathcal{F}$ and $\{F_i\}_{i=1}^{\infty}$ is disjoint, i.e., $F_i \cap F_j = \emptyset$ for any $i \neq j$, then

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} F_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(F_i).$$

The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a probability space.

A.1 Convergence Almost Surely Implies Convergence in Probability

Suppose that a sequence of random variables X_n converges almost surely to a random variable X . Then the set $A = \{\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) \neq X(\omega)\}$ has probability zero.

Fix $\varepsilon > 0$ and consider the event $E_n^{(\varepsilon)} := \bigcup_{m \geq n} \{|X_m(\omega) - X(\omega)| > \varepsilon\}$ for each $n \in \mathbb{N}$. These form a decreasing sequence of events, i.e., $E_n^{(\varepsilon)} \supseteq E_{n+1}^{(\varepsilon)} \supseteq \dots$, which decreases towards the event $E_{\infty}^{(\varepsilon)} := \bigcap_{n \geq 1} E_n^{(\varepsilon)}$. Moreover, the probabilities $\mathbb{P}(E_n^{(\varepsilon)})$ decrease towards the probability $\mathbb{P}(E_{\infty}^{(\varepsilon)}) \geq 0$.

Consider $\omega \in A^c$, so $\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)$, which implies that $|X_n(\omega) - X(\omega)| < \varepsilon$ for all n larger than a certain number N . Thus $\omega \notin E_n^{(\varepsilon)}$ for all $n \geq N$, from which it follows that $\omega \notin E_\infty^{(\varepsilon)}$. Hence $E_\infty^{(\varepsilon)}$ and A^c are disjoint, so $E_\infty^{(\varepsilon)} \subset A$, which means $\mathbb{P}(E_\infty^{(\varepsilon)}) = 0$. Finally, convergence in probability follows because

$$\mathbb{P}(|X_n(\omega) - X(\omega)| > \varepsilon) \leq \mathbb{P}(E_\infty^{(\varepsilon)}) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

A.2 Borel-Cantelli Lemma

Let $\{E_n\}_{n \in \mathbb{N}}$ be a sequence of events in sample space Ω . Then

$$E^S := \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} E_m$$

is the *limsup* event of the infinite sequence. Event E^S occurs if and only if

- (i) for all $n \geq 1$, there exists an $m \geq n$ such that E_m occurs;
- (ii) infinitely many of the E_n occur.

Similarly, let

$$E^I := \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} E_m$$

be the *liminf* event of the infinite sequence $\{E_n\}_{n \in \mathbb{N}}$. Event E^I occurs if and only if

- (i) there exists $n \geq 1$, such that for all $m \geq n$ E_m occurs;
- (ii) only finitely many of the E_n do not occur.

Theorem A.1 (The Borel-Cantelli Lemma) *Let $\{E_n\}_{n \in \mathbb{N}}$ be a sequence of events in sample space Ω . Then*

- (a) $\mathbb{P}(E^S) = 0$, i.e., $\mathbb{P}(E_n \text{ occurs infinitely often}) = 0$, provided that

$$\sum_{n=1}^{\infty} \mathbb{P}(E_n) < \infty;$$

- (b) $\mathbb{P}(E^S) = 1$, i.e., $\mathbb{P}(E_n \text{ occurs infinitely often}) = 1$, provided that the events $\{E_n\}$ are independent and

$$\sum_{n=1}^{\infty} \mathbb{P}(E_n) = \infty.$$

A.3 Burkholder-Davis-Gundy Inequalities

For a continuous local martingale $M := (M_t)_{t \geq 0}$, denote by $[M]$ the quadratic variation process and write $[M]_\infty$ for the pointwise limit of $[M]_t$ as $t \rightarrow \infty$. Note that this pointwise limit always exists in $\mathbb{R}^+ \cup \{\infty\}$.

Theorem A.2 (Burkholder-Davis-Gundy Inequalities) *For each $p > 0$ there exist constants $c_p, C_p \in (0, \infty)$, such that for any continuous local martingale M with $M_0 = 0$,*

$$c_p \mathbb{E}[M]_\infty^{p/2} \leq \mathbb{E} \left[\sup_{t \geq 0} |M_t| \right]^p \leq C_p \mathbb{E}[M]_\infty^{p/2}.$$

The Burkholder-Davis-Gundy inequalities were first proved for discrete martingales and $p > 1$ by Burkholder in 1966. In 1968 Millar extended the result to continuous martingales. In 1970, Davis extended the result for discrete martingales to $p = 1$. The extension to $p > 0$ was obtained independently by Burkholder and Gundy in 1970 and Novikov in 1971.

Endnotes

See Gut [57] for background material on probability theory. Wikipedia has an easily accessible article on the comparison of different kinds of convergences for random variables.

Appendix B

Chain Rule for Affine RODEs

Affine RODEs were introduced in Chap. 7. They are pathwise Carathéodory ODEs when the noise process has Lebesgue integrable rather than continuous sample paths. A d -dimensional affine RODE with m -dimensional affine noise has the structure

$$\frac{dx}{dt} = f_0(t, x) + \sum_{j=1}^m f_j(t, x) \eta_t^j, \quad (\text{B.1})$$

where $x = (x^1, \dots, x^m) \in \mathbb{R}^d$ and the noise process $\eta_t = (\eta_t^1, \dots, \eta_t^m)$ takes values in \mathbb{R}^m . It will be assumed here that the coefficient functions $f_0, f_1, \dots, f_m : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ are at least continuously differentiable and that the sample paths of noise processes are assumed to be Lebesgue integrable.

The differential equation (B.1) is interpreted pathwise as a Carathéodory ODE with weak derivatives and has the equivalent integral equation representation

$$x(t) = x(t_0) + \int_{t_0}^t f_0(s, x(s)) ds + \sum_{j=1}^m \int_{t_0}^t f_j(s, x(s)) \eta_s^j ds.$$

The following results on absolutely continuous functions and their properties are needed below.

Theorem B.1 [52, Theorem 4.14] *If $F : [a, b] \rightarrow \mathbb{R}$ is an absolutely continuous function on $[a, b]$, then the weak derivative F' is Lebesgue integrable on $[a, b]$ and*

$$\int_a^t F'(s) ds = F(t) - F(a)$$

for each $t \in [a, b]$.

Let $AC_{loc}(I; \mathbb{R}^d)$ be the space of functions $x : I \rightarrow \mathbb{R}^d$ which are locally absolutely continuous.

Theorem B.2 *Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a continuously differentiable function and let $I \subset \mathbb{R}$ be an interval. Then $f \circ x \in AC_{loc}(I; \mathbb{R})$ for every $x \in AC_{loc}(I; \mathbb{R}^d)$ and*

$$(f \circ x)'(t) = \sum_{i=1}^d \frac{\partial f}{\partial x_i}(x(t)) x'_i(t)$$

for \mathcal{L}^1 -a.e. $t \in I$.

Theorem B.2 is from Marcus and Mizel [101] for the more general case of a locally Lipschitz function f (which requires some additional assumptions about the set of points where f is not differentiable).

Define the partial differential operators L^0 and L^1, \dots, L^m by

$$L^0 = \frac{\partial}{\partial t} + \sum_{k=1}^d f_0^k \frac{\partial}{\partial x^k}, \quad L^j = \sum_{k=1}^d f_j^k \frac{\partial}{\partial x^k}, \quad j = 1, \dots, m,$$

where f_j^k is the k th component of the vector valued functions f_j for $k = 1, \dots, d$ and $j = 0, 1, \dots, m$.

Theorem B.3 *Let $U : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ be continuously differentiable. Then the chain rule for $U(t, x(t))$, where $x(t)$ is a solution of the RODE with affine noise (B.1) is given by*

$$U(t, x(t)) = U(t_0, x(t_0)) + \int_{t_0}^t L^0 U(s, x(s)) \, ds + \sum_{j=1}^m \int_{t_0}^t L^j U(s, x(s)) \eta_s^j \, ds$$

for $t \in [t_0, T]$.

Proof First consider a continuously differentiable function $U : \mathbb{R}^d \rightarrow \mathbb{R}$ and let $x(t)$ be the solution of the affine RODE (B.1), i.e., satisfies a.e.

$$x'(t) = f_0(t, x(t)) + \sum_{j=1}^m f_j(t, x(t)) \eta_t^j.$$

Then Theorem B.2 holds with $f = U$ and $I = [t_0, T]$. It follows that $U(x(t))$ is absolutely continuous on $[t_0, T]$ and satisfies a.e.

$$U'(x(t)) = \sum_{i=1}^d \frac{\partial U}{\partial x_i}(x(t)) x'_i(t).$$

Hence

$$\begin{aligned}
U'(x(t)) &= \sum_{i=1}^d \frac{\partial U}{\partial x_i} \left(x(t) \right) \left(f_0^i(t, x(t)) + \sum_{j=1}^m f_j^i(t, x(t)) \eta_t^j \right) \\
&= \sum_{i=1}^d f_0^i \left(t, x(t) \right) \frac{\partial U}{\partial x_i} \left(x(t) \right) + \sum_{j=1}^m \sum_{i=1}^d f_j^i \left(t, x(t) \right) \frac{\partial U}{\partial x_i} \left(x(t) \right) \eta_t^j \\
&= L^0 U((x(t))) + \sum_{j=1}^m L^j \left(x(t) \right) \eta_t^j. \tag{B.2}
\end{aligned}$$

For a continuously differentiable function $U : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$, the affine RODE (B.1) is extended to include an extra component $x_0(t) \equiv t$, which satisfies

$$x'_0(t) = f_0^0(x_0(t), x(t)) + f_0^i(t, x(t)) \eta_t^0 \equiv 1$$

with $f_0^0(t, x) \equiv 1$, $f_0^i(t, x) \equiv 0$ and $\eta_t^0 \equiv 1$. The first summation in (B.2) becomes

$$\begin{aligned}
\sum_{i=0}^d f_0^i \left(t, x(t) \right) \frac{\partial U}{\partial x_i} \left(x_0(t), x(t) \right) &= f_0^0 \left(t, x(t) \right) \frac{\partial U}{\partial x_0} \left(x_0(t), x(t) \right) \\
&\quad + \sum_{i=1}^d f_0^i \left(t, x(t) \right) \frac{\partial U}{\partial x_i} \left(x_0(t), x(t) \right) \\
&= 1 \cdot \frac{\partial U}{\partial t} \left(t, x(t) \right) + \sum_{i=1}^d f_0^i \left(t, x(t) \right) \frac{\partial U}{\partial x_i} \left(t, x(t) \right) \\
&= L^0 U(t, x(t)),
\end{aligned}$$

while the other summations remain unchanged.

Finally, Theorem B.1 gives the chain rule in integral form, i.e.,

$$U(t, x(t)) = U(t_0, x(t_0)) + \int_{t_0}^t L^0 U(s, x(s)) \, ds + \sum_{j=1}^m \int_{t_0}^t L^j U(s, x(s)) \eta_s^j \, ds. \quad \square$$

Endnotes

Various versions of Theorem B.2 can be found in Leoni [98, Theorem 4.45] and Evans and Gariepy [47, Theorem 4.4b]. See also Ziemer [146] for related results.

Appendix C

Covariance Matrix of a Fractional Brownian Motion and Its Integral

The generation of the same sample paths of a fBm $B_H(t)$ and its Riemann integral $I(B_H)$ simultaneously requires the covariance matrix:

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} = \begin{pmatrix} (\mathbb{E}[X_i X_j])_{i,j} & (\mathbb{E}[X_i Y_j])_{i,j} \\ (\mathbb{E}[Y_i X_j])_{i,j} & (\mathbb{E}[Y_i Y_j])_{i,j} \end{pmatrix}, \quad (\text{C.1})$$

where for $i, j = 1, \dots, n$,

$$\begin{aligned} X_i X_j &= (B_H(t_i) - B_H(t_{i-1}), B_H(t_j) - B_H(t_{j-1})) , \\ X_i Y_j &= \left(B_H(t_i) - B_H(t_{i-1}), \int_{t_{j-1}}^{t_j} (B_H(t) - B_H(t_{j-1})) dt \right) , \\ Y_i X_j &= \left(\int_{t_{i-1}}^{t_i} (B_H(t) - B_H(t_{i-1})) dt, B_H(t_j) - B_H(t_{j-1}) \right) , \\ Y_i Y_j &= \left(\int_{t_{i-1}}^{t_i} (B_H(t) - B_H(t_{i-1})) dt, \int_{t_{j-1}}^{t_j} (B_H(t) - B_H(t_{j-1})) dt \right) . \end{aligned}$$

Each block of Σ will be calculated below. For simplicity, an equidistant partition is considered with step size $h = t_i - t_{i-1}$ for $i = 1, \dots, n$.

I. $\Sigma_{11} = (\mathbb{E}[X_i X_j])_{i,j}$

Rearranging terms in $X_i X_j$ gives

$$\Sigma_{11} = \mathbb{E} [B_H(t_i)B_H(t_j) - B_H(t_i)B_H(t_{j-1}) - B_H(t_{i-1})B_H(t_j) + B_H(t_{i-1})B_H(t_{j-1})]$$

$$= \frac{1}{2} \left((t_i^{2H} + t_j^{2H} - |t_i - t_j|^{2H}) - (t_i^{2H} + t_{j-1}^{2H} - |t_i - t_{j-1}|^{2H}) \right. \\ \left. - (t_{i-1}^{2H} + t_j^{2H} - |t_{i-1} - t_j|^{2H}) + (t_{i-1}^{2H} + t_{j-1}^{2H} - |t_{i-1} - t_{j-1}|^{2H}) \right),$$

i.e.,

$$\Sigma_{11} = \frac{1}{2} \left(-\underbrace{|t_i - t_j|^{2H}}_a + \underbrace{|t_i - t_{j-1}|^{2H}}_b + \underbrace{|t_{i-1} - t_j|^{2H}}_c - \underbrace{|t_{i-1} - t_{j-1}|^{2H}}_d \right). \quad (\text{C.2})$$

(i) Case $t_i = t_j$: the two terms a and d in (C.2) vanish and give

$$(\text{C.2}) = (t_i - t_{i-1})^{2H} = h^{2H}. \quad (\text{C.3})$$

(ii) Case $t_i > t_j$: with $t_i - t_j = kh$

$$(\text{C.2}) = -\frac{1}{2} \left((t_i - t_j)^{2H} - (t_{i-1} - t_j)^{2H} - (t_i - t_{j-1})^{2H} + (t_{i-1} - t_{j-1})^{2H} \right) \\ = \frac{1}{2} h^{2H} \left(-2k^{2H} + (k-1)^{2H} + (k+1)^{2H} \right). \quad (\text{C.4})$$

(iii) Case $t_i < t_j$: with $t_i - t_j = kh$

$$(\text{C.2}) = -\frac{1}{2} \left((t_j - t_i)^{2H} - (t_{j-1} - t_i)^{2H} - (t_j - t_{i-1})^{2H} + (t_{j-1} - t_{i-1})^{2H} \right) \\ = \frac{1}{2} h^{2H} \left(-2k^{2H} + (k-1)^{2H} + (k+1)^{2H} \right).$$

II. $\Sigma_{21} = (\mathbb{E}[Y_i X_j])_{i,j}$

$$\Sigma_{21} = \int_{t_{i-1}}^{t_i} \mathbb{E}[(B_H(t) - B_H(t_{i-1})) (B_H(t_j) - B_H(t_{j-1}))] dt \\ = \int_{t_{i-1}}^{t_i} \mathbb{E}[B_H(t)B_H(t_j) - B_H(t)B_H(t_{j-1}) - B_H(t_{i-1})B_H(t_j) + B_H(t_{i-1})B_H(t_{j-1})] dt \\ = \frac{1}{2} \int_{t_{i-1}}^{t_i} \left((t^{2H} + t_j^{2H} - |t - t_j|^{2H}) - (t^{2H} + t_{j-1}^{2H} - |t - t_{j-1}|^{2H}) \right. \\ \left. - (t_{i-1}^{2H} + t_j^{2H} - |t_{i-1} - t_j|^{2H}) + (t_{i-1}^{2H} + t_{j-1}^{2H} - |t_{i-1} - t_{j-1}|^{2H}) \right) dt$$

$$= \frac{1}{2} \int_{t_{i-1}}^{t_i} \left(-\underbrace{|t-t_j|^{2H}}_e + \underbrace{|t-t_{j-1}|^{2H}}_f + \underbrace{|t_{i-1}-t_j|^{2H}}_g - \underbrace{|t_{i-1}-t_{j-1}|^{2H}}_h \right) dt. \quad (\text{C.5})$$

Now compute the terms e , f , g and h separately.

(i) Case $t_i = t_j$: then $t \in [t_{i-1}, t_i]$ gives $t - t_j \leq 0$, $t - t_{j-1} \geq 0$, $t_{i-1} - t_j < 0$ and $t_{i-1} - t_{j-1} = 0$.

$$\text{Term } e = \frac{1}{2} \int_{t_{i-1}}^{t_i} (-|t-t_j|^{2H}) dt = -\frac{1}{2} \int_{t_{i-1}}^{t_i} (t_i - t)^{2H} dt = -\frac{1}{2(2H+1)} h^{2H+1},$$

$$\text{Term } f = \frac{1}{2} \int_{t_{i-1}}^{t_i} |t-t_{j-1}|^{2H} dt = \frac{1}{2} \int_{t_{i-1}}^{t_i} (t - t_{i-1})^{2H} dt = \frac{1}{2(2H+1)} h^{2H+1},$$

$$\text{Term } g = \frac{1}{2} \int_{t_{i-1}}^{t_i} |t_{i-1}-t_j|^{2H} dt = \frac{1}{2} \int_{t_{i-1}}^{t_i} (t_i - t_{i-1})^{2H} dt = \frac{1}{2} h^{2H+1}.$$

Substituting terms e , f and g in (C.5) then yields

$$(C.5) = -\frac{1}{2} h^{2H+1} + \frac{1}{2} h^{2H+1} + \frac{1}{2} h^{2H+1} = \frac{1}{2} h^{2H+1}.$$

(ii) Case $t_i > t_j$: then $t \in [t_{i-1}, t_i]$ gives $t - t_j \geq 0$, $t - t_{j-1} > 0$, $t_{i-1} - t_j \geq 0$ and $t_{i-1} - t_{j-1} > 0$.

$$\begin{aligned} \text{Term } e &= \frac{1}{2} \int_{t_{i-1}}^{t_i} (-|t-t_j|^{2H}) dt = -\frac{1}{2} \int_{t_{i-1}}^{t_i} (t - t_j)^{2H} dt \\ &= -\frac{1}{2(2H+1)} \left((t_i - t_j)^{2H+1} - (t_{i-1} - t_j)^{2H+1} \right), \end{aligned}$$

$$\begin{aligned} \text{Term } f &= \frac{1}{2} \int_{t_{i-1}}^{t_i} |t-t_{j-1}|^{2H} dt = \frac{1}{2} \int_{t_{i-1}}^{t_i} (t - t_{j-1})^{2H} dt \\ &= \frac{1}{2(2H+1)} \left((t_i - t_{j-1})^{2H+1} - (t_{i-1} - t_{j-1})^{2H+1} \right), \end{aligned}$$

$$\begin{aligned} \text{Term } g &= \frac{1}{2} \int_{t_{i-1}}^{t_i} |t_{i-1}-t_j|^{2H} dt = \frac{1}{2} \int_{t_{i-1}}^{t_i} (t_{i-1} - t_j)^{2H} dt \\ &= \frac{1}{2} \left((t_i - t_{i-1})(t_{i-1} - t_j) \right)^{2H}, \end{aligned}$$

$$\begin{aligned} \text{Term } h &= \frac{1}{2} \int_{t_{i-1}}^{t_i} (-|t_{i-1} - t_{j-1}|^{2H}) dt = -\frac{1}{2} \int_{t_{i-1}}^{t_i} (t_{i-1} - t_{j-1})^{2H} dt \\ &= -\frac{1}{2} \left(t_i - t_{i-1} \right) (t_{i-1} - t_{j-1})^{2H}. \end{aligned}$$

Substituting terms e , f , g and h in (C.5) gives

$$\begin{aligned} (C.5) &= -\frac{1}{2} \left\{ \frac{1}{2H+1} \left((t_i - t_j)^{2H+1} - (t_{i-1} - t_j)^{2H+1} \right. \right. \\ &\quad \left. \left. - (t_i - t_{j-1})^{2H+1} + (t_{i-1} - t_{j-1})^{2H+1} \right) \right. \\ &\quad \left. - (t_i - t_{i-1})(t_{i-1} - t_j)^{2H} + (t_i - t_{i-1})(t_{i-1} - t_{j-1})^{2H} \right\} \\ &= -\frac{1}{2} h^{2H+1} \left\{ \frac{1}{2H+1} \left(2k^{2H+1} - (k-1)^{2H+1} - (k+1)^{2H+1} \right) \right. \\ &\quad \left. - (k-1)^{2H} + k^{2H} \right\}, \end{aligned}$$

where $t_i - t_j = kh$.

(iii) Case $t_i < t_j$: then $t \in [t_{i-1}, t_i]$ gives $t - t_j < 0$, $t - t_{j-1} \leq 0$, $t_{i-1} - t_j < 0$ and $t_{i-1} - t_{j-1} < 0$. The equation (C.5) can be estimated in the same manner as for the case $t_i > t_j$ to obtain

$$\begin{aligned} (C.5) &= \frac{1}{2} \left\{ \frac{1}{2H+1} \left((t_j - t_i)^{2H+1} - (t_{j-1} - t_i)^{2H+1} \right. \right. \\ &\quad \left. \left. - (t_j - t_{i-1})^{2H+1} + (t_{j-1} - t_{i-1})^{2H+1} \right) \right. \\ &\quad \left. + (t_i - t_{i-1})(t_j - t_{i-1})^{2H} - (t_i - t_{i-1})(t_{j-1} - t_{i-1})^{2H} \right\} \\ &= \frac{1}{2} h^{2H+1} \left\{ \frac{1}{2H+1} \left(2k^{2H+1} - (k-1)^{2H+1} - (k+1)^{2H+1} \right) \right. \\ &\quad \left. + (k+1)^{2H} - k^{2H} \right\}, \end{aligned} \tag{C.6}$$

where $t_j - t_i = kh$.

III. $\Sigma_{12} = (\mathbb{E}[X_i Y_j])_{i,j}$

The arguments for Σ_{21} can also be applied to $\Sigma_{12} = (\mathbb{E}[X_i Y_j])_{i,j}$ to obtain

$$\begin{aligned}
\Sigma_{12} &= \int_{t_{j-1}}^{t_j} \mathbb{E}[(B_H(t_i) - B_H(t_{i-1})) (B_H(t) - B_H(t_{j-1}))] dt \\
&= \int_{t_{j-1}}^{t_j} \mathbb{E}[B_H(t_i)B_H(t) - B_H(t_i)B_H(t_{j-1}) - B_H(t_{i-1})B_H(t) + B_H(t_{i-1})B_H(t_{j-1})] dt \\
&= \frac{1}{2} \int_{t_{j-1}}^{t_j} \left((t_i^{2H} + t^{2H} - |t_i - t|^{2H}) - (t_i^{2H} + t_{j-1}^{2H} - |t_i - t_{j-1}|^{2H}) \right. \\
&\quad \left. - (t_{i-1}^{2H} + t^{2H} - |t_{i-1} - t|^{2H}) + (t_{i-1}^{2H} + t_{j-1}^{2H} - |t_{i-1} - t_{j-1}|^{2H}) \right) dt \\
&= \frac{1}{2} \int_{t_{j-1}}^{t_j} \left(-|t_i - s|^{2H} + |t_i - t_{j-1}|^{2H} + |t_{i-1} - s|^{2H} - |t_{i-1} - t_{j-1}|^{2H} \right) dt. \quad (\text{C.7})
\end{aligned}$$

(i) Case $t_i = t_j$: obviously

$$(\text{C.7}) = \frac{1}{2} \left(t_i - t_{i-1} \right)^{2H+1} = \frac{1}{2} h^{2H+1}. \quad (\text{C.8})$$

(ii) Case $t_i > t_j$: the equation (C.7) can be estimated in the same manner as (C.5) to obtain

$$\begin{aligned}
(\text{C.7}) &= \frac{1}{2} \left\{ \frac{1}{2H+1} \left((t_i - t_j)^{2H+1} - (t_{i-1} - t_j)^{2H+1} \right. \right. \\
&\quad \left. \left. - (t_i - t_{j-1})^{2H+1} + (t_{i-1} - t_{j-1})^{2H+1} \right) \right. \\
&\quad \left. + (t_j - t_{j-1})(t_i - t_{j-1})^{2H} - (t_j - t_{j-1})(t_{i-1} - t_{j-1})^{2H} \right\} \\
&= \frac{1}{2} h^{2H+1} \left(\frac{1}{2H+1} \left(2k^{2H+1} - (k-1)^{2H+1} - (k+1)^{2H+1} \right) \right. \\
&\quad \left. + (k+1)^{2H} - k^{2H} \right), \quad (\text{C.9})
\end{aligned}$$

where $t_i - t_j = kh$.

(iii) Case $t_i < t_j$: with $t_j - t_i = kh$

$$\begin{aligned}
(\text{C.7}) &= -\frac{1}{2} \left\{ \frac{1}{2H+1} \left((t_j - t_i)^{2H+1} - (t_{j-1} - t_i)^{2H+1} \right. \right. \\
&\quad \left. \left. - (t_j - t_{i-1})^{2H+1} + (t_{j-1} - t_{i-1})^{2H+1} \right) \right\}
\end{aligned}$$

$$\begin{aligned}
& - (t_j - t_{j-1})(t_{j-1} - t_i)^{2H} + (t_j - t_{j-1})(t_{j-1} - t_{i-1})^{2H} \Big\} \\
= & -\frac{1}{2} h^{2H+1} \left\{ \frac{1}{2H+1} \left(2k^{2H+1} - (k-1)^{2H+1} - (k+1)^{2H+1} \right) \right. \\
& \quad \left. - (k-1)^{2H} + k^{2H} \right\}, \tag{C.10}
\end{aligned}$$

$$\mathbf{IV. } \Sigma_{22} = (\mathbb{E}[Y_i Y_j])_{i,j}$$

$$\begin{aligned}
\Sigma_{22} &= \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_j} \mathbb{E}[(B_H(t) - B_H(t_{i-1})) (B_H(s) - B_H(t_{j-1}))] ds dt \\
&= \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_j} \mathbb{E}[B_H(t)B_H(s) - B_H(t)B_H(t_{j-1}) - B_H(t_{i-1})B_H(s) \\
&\quad + B_H(t_{i-1})B_H(t_{j-1})] ds dt \\
\Sigma_{22} &= \frac{1}{2} \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_j} \left((t^{2H} + s^{2H} - |t-s|^{2H}) - (t^{2H} + t_{j-1}^{2H} - |t-t_{j-1}|^{2H}) \right. \\
&\quad \left. - (t_{i-1}^{2H} + s^{2H} - |t_{i-1}-s|^{2H}) + (t_{i-1}^{2H} + t_{j-1}^{2H} - |t_{i-1}-t_{j-1}|^{2H}) \right) ds dt \\
&= \frac{1}{2} \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_j} \left(- \underbrace{|t-s|^{2H}}_m + \underbrace{|t-t_{j-1}|^{2H}}_n + \underbrace{|t_{i-1}-s|^{2H}}_o - \underbrace{|t_{i-1}-t_{j-1}|^{2H}}_p \right) ds dt. \tag{C.11}
\end{aligned}$$

Next m , n , o and p are computed separately for each different case.

(i) Case $t_i = t_j$: then $t, s \in [t_{i-1}, t_i]$ gives $t - t_{j-1} \geq 0$ and $t_{i-1} - s \leq 0$. In addition, $t_{i-1} - t_{j-1} = 0$.

$$\begin{aligned}
\text{Term } m &= \frac{1}{2} \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_j} (-|t-s|^{2H}) ds dt = -\frac{1}{2} \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t_i} |t-s|^{2H} ds dt \\
&= -\frac{1}{2} \int_{t_{i-1}}^{t_i} \left(\int_{t_{i-1}}^t (t-s)^{2H} ds + \int_t^{t_i} (s-t)^{2H} ds \right) dt \\
&= -\frac{1}{2(2H+1)} \int_{t_{i-1}}^{t_i} \left([-(t-s)^{2H+1}]_{t_{i-1}}^t + [(s-t)^{2H+1}]_t^{t_i} \right) dt
\end{aligned}$$

$$\begin{aligned}
&= -\frac{1}{2(2H+1)} \int_{t_{i-1}}^{t_i} ((t-t_{i-1})^{2H+1} + (t_i-t)^{2H+1}) dt \\
&= -\frac{1}{2(2H+1)(2H+2)} \left[(t-t_{i-1})^{2H+2} - (t_i-t)^{2H+2} \right] \Big|_{t=t_{i-1}}^{t_i} \\
&= -\frac{1}{(2H+1)(2H+2)} \left(t_i - t_{i-1} \right)^{2H+2},
\end{aligned}$$

$$\begin{aligned}
\text{Term } n &= \frac{1}{2} \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_j} |t-t_{j-1}|^{2H} ds dt = \frac{1}{2} \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t_i} (t-t_{i-1})^{2H} ds dt \\
&= \frac{1}{2} \left(t_i - t_{i-1} \right) \int_{t_{i-1}}^{t_i} (t-t_{i-1})^{2H} dt = \frac{1}{2(2H+1)} \left(t_i - t_{i-1} \right)^{2H+2},
\end{aligned}$$

$$\begin{aligned}
\text{Term } o &= \frac{1}{2} \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_j} |t_{i-1}-s|^{2H} ds dt = \frac{1}{2} \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t_i} (s-t_{i-1})^{2H} ds dt \\
&= \frac{1}{2} \int_{t_{i-1}}^{t_i} \frac{1}{2H+1} \left(t_i - t_{i-1} \right)^{2H+1} dt = \frac{1}{2(2H+1)} \left(t_i - t_{i-1} \right)^{2H+2}.
\end{aligned}$$

Substituting terms m , n and o in (C.11) gives

$$\begin{aligned}
(\text{C.11}) &= -\frac{1}{(2H+1)(2H+2)} \left(t_i - t_{i-1} \right)^{2H+2} + \frac{1}{(2H+1)} \left(t_i - t_{i-1} \right)^{2H+2} \\
&= \frac{1}{2H+2} \left(t_i - t_{i-1} \right)^{2H+2} = \frac{1}{2H+2} h^{2H+2}. \tag{C.12}
\end{aligned}$$

(ii) Case $t_i > t_j$: then $t \in [t_{i-1}, t_i]$ and $s \in [t_{j-1}, t_j]$ yield $t-s \geq 0$, $t-t_{j-1} > 0$, $t_{i-1}-s \geq 0$ and $t_{i-1}-t_{j-1} > 0$.

$$\begin{aligned}
\text{Term } m &= \frac{1}{2} \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_j} (-|t-s|^{2H}) ds dt = -\frac{1}{2} \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_j} (t-s)^{2H} ds dt \\
&= -\frac{1}{2} \int_{t_{i-1}}^{t_i} \left[-\frac{1}{2H+1} (t-s)^{2H+1} \right] \Big|_{s=t_{j-1}}^{t_j} dt \\
&= \frac{1}{2(2H+1)} \int_{t_{i-1}}^{t_i} ((t-t_j)^{2H+1} - (t-t_{j-1})^{2H+1}) dt \\
&= \frac{1}{2(2H+1)(2H+2)} \left((t_i-t_j)^{2H+2} - (t_{i-1}-t_j)^{2H+2} \right)
\end{aligned}$$

$$- (t_i - t_{j-1})^{2H+2} + (t_{i-1} - t_{j-1})^{2H+2} \Big),$$

$$\begin{aligned} \text{Term } n &= \frac{1}{2} \int_{t_i}^{t_{i+1}} \int_{t_j}^{t_{j+1}} |t - t_j|^{2H} ds dt = \frac{1}{2} \int_{t_i}^{t_{i+1}} \int_{t_j}^{t_{j+1}} (t - t_j)^{2H} ds dt \\ &= \frac{1}{2} \left(t_{j+1} - t_j \right) \int_{t_i}^{t_{i+1}} (t - t_j)^{2H} dt \\ &= \frac{1}{2(2H+1)} \left(t_{j+1} - t_j \right) \left((t_{i+1} - t_j)^{2H+1} - (t_i - t_j)^{2H+1} \right), \end{aligned}$$

$$\begin{aligned} \text{Term } o &= \frac{1}{2} \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_j} |t_{i-1} - s|^{2H} ds dt = \frac{1}{2} \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_j} (t_{i-1} - s)^{2H} ds dt \\ &= \frac{1}{2} \int_{t_{i-1}}^{t_i} -\frac{1}{2H+1} \left((t_{i-1} - t_j)^{2H+1} - (t_{i-1} - t_{j-1})^{2H+1} \right) dt \\ &= -\frac{1}{2(2H+1)} \left(t_i - t_{i-1} \right) \left((t_{i-1} - t_j)^{2H+1} - (t_{i-1} - t_{j-1})^{2H+1} \right), \end{aligned}$$

$$\begin{aligned} \text{Term } p &= \frac{1}{2} \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_j} (-|t_{i-1} - t_{j-1}|^{2H}) ds dt = -\frac{1}{2} \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_j} (t_{i-1} - t_{j-1})^{2H} ds dt \\ &= -\frac{1}{2} \left(t_i - t_{i-1} \right) \left(t_j - t_{j-1} \right) \left(t_{i-1} - t_{j-1} \right)^{2H}. \end{aligned}$$

Substituting terms m, n, o and p in (C.11) gives

$$\begin{aligned} (\text{C.11}) &= \frac{1}{2(2H+1)(2H+2)} \left((t_i - t_j)^{2H+2} - (t_{i-1} - t_j)^{2H+2} \right. \\ &\quad \left. - (t_i - t_{j-1})^{2H+2} + (t_{i-1} - t_{j-1})^{2H+2} \right) \\ &\quad + \frac{1}{2(2H+1)} \left(t_j - t_{j-1} \right) \left((t_i - t_{j-1})^{2H+1} - (t_{i-1} - t_{j-1})^{2H+1} \right) \\ &\quad - \frac{1}{2(2H+1)} \left(t_i - t_{i-1} \right) \left((t_{i-1} - t_j)^{2H+1} - (t_{i-1} - t_{j-1})^{2H+1} \right) \\ &\quad - \frac{1}{2} \left(t_i - t_{i-1} \right) \left(t_j - t_{j-1} \right) \left(t_{i-1} - t_{j-1} \right)^{2H}. \end{aligned}$$

Suppose that $t_i - t_j = kh$. Then $t_{i-1} - t_j = (k-1)h$, $t_i - t_{j-1} = (k+1)h$ and

$$(C.11) = \frac{1}{2} h^{2H+2} \left(\frac{1}{2H+1} ((k+1)^{2H+1} - (k-1)^{2H+1}) - k^{2H} + \frac{1}{(2H+1)(2H+2)} \left(2k^{2H+2} - (k-1)^{2H+2} - (k+1)^{2H+2} \right) \right). \quad (C.13)$$

(iii) Case $t_i < t_j$: then $t \in [t_{i-1}, t_i]$ and $s \in [t_{j-1}, t_j]$ give $t - s \leq 0$, $t - t_{j-1} \leq 0$, $t_{i-1} - s < 0$ and $t_{i-1} - t_{j-1} < 0$. The terms m , n , o and p can be estimated in the same manner to obtain

$$\begin{aligned} (C.11) = & \frac{1}{2(2H+1)(2H+2)} \left((t_j - t_i)^{2H+2} - (t_{j-1} - t_i)^{2H+2} \right. \\ & \left. - (t_j - t_{i-1})^{2H+2} + (t_{j-1} - t_{i-1})^{2H+2} \right) \\ & + \frac{1}{2(2H+1)} \left(t_i - t_{i-1} \right) \left((t_j - t_{i-1})^{2H+1} - (t_{j-1} - t_{i-1})^{2H+1} \right) \\ & - \frac{1}{2(2H+1)} \left(t_j - t_{j-1} \right) \left((t_{j-1} - t_i)^{2H+1} - (t_{j-1} - t_{i-1})^{2H+1} \right) \\ & - \frac{1}{2} \left(t_i - t_{i-1} \right) \left(t_j - t_{j-1} \right) \left(t_{j-1} - t_{i-1} \right)^{2H}. \end{aligned}$$

Suppose that $t_j - t_i = kh$. Then $t_{j-1} - t_i = (k-1)h$, $t_j - t_{i-1} = (k+1)h$ and

$$(C.11) = \frac{1}{2} h^{2H+2} \left(\frac{1}{2H+1} ((k+1)^{2H+1} - (k-1)^{2H+1}) - k^{2H} + \frac{1}{(2H+1)(2H+2)} \left(2k^{2H+2} - (k-1)^{2H+2} - (k+1)^{2H+2} \right) \right), \quad (C.14)$$

which is equal to (C.13).

In conclusion, to see that the covariance matrix Σ given by (C.1) is a symmetric matrix with elements that are real numbers, write

$$\Sigma_{11} = \begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_1 & a_0 & a_1 & \cdots & a_{n-2} \\ a_2 & a_1 & a_0 & \cdots & a_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n-1} & a_{n-2} & a_{n-3} & \cdots & a_0 \end{pmatrix}, \quad \Sigma_{12} = \begin{pmatrix} b_0 & b_1 & b_2 & \cdots & b_{n-1} \\ c_1 & b_0 & b_1 & \cdots & b_{n-2} \\ c_2 & c_1 & b_0 & \cdots & b_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{n-1} & c_{n-1} & c_{n-3} & \cdots & b_0 \end{pmatrix},$$

$$\Sigma_{21} = \begin{pmatrix} b_0 & c_1 & c_2 & \cdots & c_{n-1} \\ b_1 & b_0 & c_1 & \cdots & c_{n-2} \\ b_2 & b_1 & b_0 & \cdots & c_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{n-1} & b_{n-2} & b_{n-3} & \cdots & b_0 \end{pmatrix}, \quad \Sigma_{22} = \begin{pmatrix} d_0 & d_1 & d_2 & \cdots & d_{n-1} \\ d_1 & d_0 & d_1 & \cdots & d_{n-2} \\ d_2 & d_1 & d_0 & \cdots & d_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{n-1} & d_{n-2} & d_{n-3} & \cdots & d_0 \end{pmatrix},$$

where $a_0, a_i, b_0, b_i, c_i, d_0$, and d_i are given by (C.3), (C.4), (C.8), (C.9), (C.10), (C.12) and (C.13), respectively, for $i = 1, \dots, n - 1$.

Endnotes

The covariance matrix of the fractional Brownian motion is taken from Asai [8].

References

1. E. Allen, *Modeling with Itô Stochastic Differential Equations* (Springer, Dordrecht, 2007)
2. L.J.S. Allen, *An Introduction to Stochastic Processes with Applications to Biology*, 2nd edn. (CRC Press, Taylor & Francis Group, 2010)
3. L. Arnold, *Stochastic Differential Equations* (Wiley & Sons, New York, 1974)
4. L. Arnold, *Random Dynamical Systems* (Springer, Berlin, 1998)
5. L. Arnold, P.E. Kloeden, Lyapunov exponents and rotation number of two-dimensional systems with telegraphic noise. *SIAM J. Appl. Math.* **49**, 1242–1274 (1989)
6. L. Arnold, P.E. Kloeden, Discretization of a random dynamical system near a hyperbolic point. *Mathematische Nachrichten*. **181**, 43–72 (1996)
7. Y. Asai, T. Caraballo, X. Han, P. Kloeden, *Advances in Dynamical Systems and Control, A random model for immune response to virus in fluctuating environments* (Springer, Berlin, 2016), pp. 211–225
8. Y. Asai, *Numerical Methods for Random Ordinary Differential Equations and their Applications in Biology and Medicine*. Dissertation, Institut für Mathematik, Goethe Universität Frankfurt am Main (2016)
9. Y. Asai, P.E. Kloeden, Numerical schemes for random ODEs via stochastic differential equations. *Commun. Appl. Anal.* **17**(3 & 4), 511–528 (2013)
10. Y. Asai, P.E. Kloeden, Multi-step schemes for random ODEs driven by an Itô diffusion process. *Comput. Appl. Math.* **294**, 210–224 (2016)
11. Y. Asai, P.E. Kloeden, Numerical schemes for random ODEs with affine noise. *Numer. Algorithms* **72**, 155–171 (2016)
12. Y. Asai, E. Herrmann, P.E. Kloeden, Stable integration of stiff random ordinary differential equations. *Stoch. Anal. Appl.* **31**, 293–313 (2013)
13. L. Bachelier, Théorie de la spéculation. *Annales Scientifiques de L' École Normale Supérieure* **17**, 21–86 (1900)
14. S. Becker, *Numerical Simulation of Random Ordinary Differential Equations and Stochastic Partial Differential Equations* (Diplomarbeit, Institut für Mathematik, Johann Wolfgang Goethe Universität, 2009)
15. W.-J. Beyn, On the numerical approximation of phase portraits near stationary points. *SIAM J. Numer. Anal.* **14**, 1095–1113 (1987)
16. A.T. Bharucha-Reid, *Random Integral Equations* (Academic Press, New York, 1972)
17. D.C. Biles, P.A. Binding, On Carathéodory's conditions for the initial value problem. *Proc. Am. Math. Soc.* **125**, 1371–1376 (1997)

18. D. Bobrowski, *Wstęp do losowych równań różniczkowych zwyczajnych* (Państwowe Wydawnictwo Naukowe, Warszawa, 1987)
19. E. Buckwar, R. Winkler, Multi-step methods for SODEs and their application to problems with small noise. *SIAM J. Numer. Anal.* **44**(2), 779–803 (2006)
20. H. Bunke, *Gewöhnliche Differentialgleichungen mit zufälligen Parametern* (Akademie, Berlin, 1972)
21. K. Burrage, P.M. Burrage, High strong order explicit Runge-Kutta methods for stochastic ordinary differential equations. *Appl. Numer. Math.* **22**, 81–101 (1996)
22. J.C. Butcher, *The Numerical Analysis of Ordinary Differential Equations: Runge-Kutta and General Linear Methods* (John Wiley & Sons, New York, 1986)
23. T. Caraballo, X. Han, P.E. Kloeden, A. Rappaport.: Dynamics of non-autonomous chemostat model. In: V.A. Sadovnichiy, M.Z. Zgurovsky (eds.) *Continuous and Distributed Systems II, Studies in Systems, Decision and Controls*, vol. 30, (Springer, Berlin, 2015) pp. 103–120
24. T. Caraballo, X. Han, *Applied Nonautonomous and Random Dynamical Systems* (Springer, BCAM SpringerBrief, 2016)
25. T. Caraballo, P.E. Kloeden, B. Schmalfuß, Exponentially stable stationary solutions for stochastic evolution equations and their perturbation. *Appl. Math. Optim.* **50**, 183–207 (2004)
26. T. Caraballo, P.E. Kloeden, J. Real, Discretization of asymptotically stable stationary solutions of delay differential equations with a random stationary delay. *J. Dyn. Differ. Equ.* **18**, 863–880 (2006)
27. T. Caraballo, X. Han, P.E. Kloeden, Chemostats with random inputs and wall growth. *Math. Methods Appl. Sci.* **38**, 3538–3550 (2015)
28. T. Caraballo, X. Han, P.E. Kloeden, Nonautonomous chemostats with variable delays. *SIAM J. Math. Anal.* **47**, 2178–2199 (2015)
29. C. Carathéodory, *Vorlesungen über reelle Funktionen* (Teubner, 1918)
30. F. Carbonell, J.C. Jimenez, R.J. Biscay, H. de la Cruz, The local linearization method for numerical integration of random differential equations. *BIT* **45**, 1–14 (2005)
31. I. Chueshov, *Monotone random systems theory and applications*, Lecture Notes in Mathematics (Springer, Berlin, 2002)
32. E.A. Coddington, N. Levinson, *Theory of Ordinary Differential Equations* (McGraw-Hill, New York, 1955)
33. J.-F. Coeurjolly, Simulation and identification of the fractional Brownian motion: a bibliographical and comparative study. *J. Stat. Softw.* **5**, (2000). doi:[10.18637/jss.v005.i07](https://doi.org/10.18637/jss.v005.i07)
34. H. Crauel, P.E. Kloeden, Nonautonomous and random Attractors. *Jahresbericht der Deutschen Mathematiker-Vereinigung* **117**, 173–206 (2015)
35. H. Crauel, P.E. Kloeden, Meihua Yang, Random attractors of stochastic reaction-diffusion equations on variable domains. *Stoch. Dyn.* **11**, 301–314 (2011)
36. R.B. Davies, D.S. Harte, Tests for hurst effect. *Biometrika* **74**, 95–102 (1987)
37. weight functions of b-series expressed as multiple integrals, K. Debrabant, A. Kværnø Stochastic taylor expansions. *J. Stoch. Anal. Appl.* **2**, 293–302 (2010)
38. P. Deuflhard, V. Bornemann, *Scientific Computing with Ordinary Differential Equations* (Springer, New York, 2002)
39. C.R. Dietrich, G.N. Newsam, Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix. *SIAM J. Sci. Comput.* **18**, 1088–1107 (1997)
40. W. Döblin, *Sur l'équation de Kolmogoroff*, Pli cacheté à l'Académie des Sciences, édité par B. Bru et M. Yor, C.R. Acad. Sci. vol. 331. Paris (2000)
41. A. d'Onofrio (ed.), *Random Bounded Noises in Physics, Biology, and Engineering* (Birkhäuser, 2013)
42. J.L. Doob, The Brownian movement and stochastic equations. *Ann. Math. Second Series* **43**, 351–369 (1942)
43. H. Doss, Liens entre équations différentielles stochastiques et ordinaires, *Ann. Inst. Henri Poincaré, Nouv. Sér., Sect. B* **13**, 99–125 (1977)

44. J. Duan, *An Introduction to Stochastic Dynamics* (Cambridge University Press, Cambridge, 2015)
45. A. Einstein, *Investigations on the Theory of the Brownian Movement* (Dover Publications). (1956 republication of the original 1926 translation)
46. L.C. Evans, *An Introduction to Stochastic Differential Equations*, Am. Math. Soc. (Providence, RI, 2013)
47. L.C. Evans, R.F. Gariepy, *Measure Theory and Fine Properties of Functions*, Revised edn. (C.R.C, Boca Raton, 2015)
48. P.K. Friz, M. Hairer, *A Course on Rough Paths: With an Introduction to Regularity Structures* (Springer, Heidelberg, 2014)
49. M.J. Garrido-Atienza, P.E. Kloeden, A. Neuenkirch, Discretization of stationary solutions of stochastic systems driven by fractional Brownian motion. *J. Appl. Math. Optim.* **60**, 151–172 (2009)
50. I.I. Gikhman, A.V. Skorokhod, *Stochastic Differential Equations* (Springer, Berlin, 1972)
51. G.S. Goodman, Subfunctions and the initial value problem for differential equations satisfying Carathéodory's hypotheses. *J. Differ. Equ.* **7**, 232–242 (1970)
52. R.A. Gordon, *The Integrals of Lebesgue, Denjoy, Perron, and Henstock* (Am. Math. Soc, Providence, RI, 1994)
53. D.F. Griffiths, D.J. Higham, *Numerical Methods for Ordinary Differential Equations: Initial Value Problems* (Springer, Berlin, 2010)
54. L. Grüne, P.E. Kloeden, Pathwise approximation of random ordinary differential equations. *BIT* **41**, 710–721 (2001)
55. L. Grüne, P.E. Kloeden, Higher order numerical schemes for affinely controlled nonlinear systems. *Numer. Math.* **89**, 669–690 (2001)
56. L. Grüne, P.E. Kloeden, Higher order numerical approximation of switching systems. *Syst. Control Lett.* **55**, 746–754 (2006)
57. A. Gut, *Probability: A Graduate Course*, 2nd edn. (Springer, New York, 2013)
58. I. Gyöngy, A note on Euler's approximations. *Potential Anal.* **8**, 205–216 (1998)
59. E. Hairer, G. Wanner, *Solving Ordinary Differential Equations I* (Springer, Heidelberg, 1991)
60. E. Hairer, G. Wanner, *Solving Ordinary Differential Equations II* (Springer, Heidelberg, 1991)
61. E. Hairer, S.P. Nørsett, G. Wanner, *Solving Ordinary Differential Equations I* (Second edn, Nonstiff Problems) (Springer, Berlin, 1993)
62. F.B. Hanson, *Applied Stochastic Processes and Control for Jump-Diffusions Modeling, Analysis, and Computation* (SIAM, Philadelphia, 2007)
63. D.J. Higham, An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Rev.* **43**(3), 525–546 (2001)
64. D.J. Higham, X. Mao, A.M. Stuart, Strong convergence of Euler-type methods for nonlinear stochastic differential equations. *SIAM J. Numer. Anal.* **40**(3), 1041–1063 (2002)
65. Y. Hu, Y. Li, D. Nualart, Taylor schemes for rough differential equations and fractional diffusions. *Discret. Conts. Dyn. Syst. Ser. B* **21**, 3115–3162 (2016)
66. A.E. Hutzenthaler, *Mathematical models for cell-cell communication on different time scales*, Ph.D. thesis, Zentrum Mathematik, Technische Universität München (2009)
67. M. Hutzenthaler, A. Jentzen, *Numerical Approximations of Stochastic Differential Equations with Non-globally Lipschitz Continuous Coefficients* (Mem. Am. Math. Soc, Providence, 2015)
68. M. Hutzenthaler, A. Jentzen, P.E. Kloeden, Strong convergence of an explicit numerical method for SDEs with non-globally Lipschitz continuous coefficients. *Ann. Appl. Probab.* **22**, 1611–1641 (2012)
69. P. Imkeller, P.E. Kloeden, On the computation of invariant measures in random dynamical systems. *Stoch. Dyn.* **3**, 247–265 (2003)
70. P. Imkeller, C. Lederer, On the cohomology of flows of stochastic and random differential equations. *Probab. Theory Relat. Fields* **120**, 209–235 (2001)
71. P. Imkeller, C. Lederer, The cohomology of stochastic and random differential equations, and local linearization of stochastic flows. *Stoch. Dyn.* **2**, 131–159 (2002)

72. P. Imkeller, B. Schmalfuß, The conjugacy of stochastic and random differential equations and the existence of global attractors. *J. Dyn. Differ. Equ.* **13**, 215–249 (2001)
73. K. Itô, On a formula concerning stochastic differentials. *Nagoya Math. J.* **3**, 55–65 (1951)
74. A. Jentzen, P.E. Kloeden, A. Neuenkirch, Pathwise convergence of numerical schemes for random and stochastic differential equations. In: Cucker, F., Pinkus, A., Todd, M. (eds.) *Foundations of Computational Mathematics*, pp. 410–161. Cambridge University Press, Hong Kong (2009)
75. A. Jentzen, P.E. Kloeden, *Taylor Expansions of Stochastic Partial Differential Equations*. CBMS Lecture ser. SIAM, Philadelphia (2011)
76. A. Jentzen, *Numerische Verfahren hoher Ordnung für zufällige Differentialgleichungen* (Johann Wolfgang Goethe-University, Diplomarbeit, 2009)
77. A. Jentzen, P.E. Kloeden, The numerical approximation of stochastic partial differential equations. *Milan J. Math.* **77**, 205–244 (2009)
78. A. Jentzen, P.E. Kloeden, Pathwise taylor schemes for random ordinary differential equations. *BIT* **49**(1), 113–140 (2009)
79. A. Jentzen, A. Neuenkirch, A random Euler scheme for Carathéodory differential equations. *J. Comput. Appl. Math.* **224**, 346–359 (2009)
80. A. Jentzen, P.E. Kloeden, A. Neuenkirch, Pathwise approximation of stochastic differential equations on domains: Higher order convergence rates without global Lipschitz coefficients. *Numerische Mathematik* **112**(1), 41–64 (2009)
81. H.-D. J. Jeong, D. McNickle, K. Pawlikowski, A Search for Computationally Efficient Generators of Synthetic Self-similar Teletraffic, *Proceedings Australasian Computer Science Conference ASCS'99*, Auckland, 21 (Springer, Singapore, 1999), pp. 75–86
82. I. Karatzas, S.E. Shreve, *Brownian Motion and Stochastic Calculus* (Springer, New York, 1988)
83. R. Khasminskii, *Stochastic Stability of Differential Equations*, 2nd edn. (Springer, Berlin, 2012)
84. P.E. Kloeden, H. Keller, B. Schmalfuß, Towards a theory of random numerical dynamics, in *Stochastic Dynamics*, ed. by H. Crauel, V.M. Gundlach (Springer, Berlin, 1999), pp. 259–282
85. P.E. Kloeden, V. Kozyakin, Asymptotic behaviour of random tridiagonal Markov chains in biological applications. *Discret. Conts. Dyn. Syst. Ser. B* **18**, 453–465 (2012)
86. P.E. Kloeden, T. Lorenz, Pullback incremental attraction. *Nonautonomous Random Dyn. Syst.* 53–60 (2013). doi:[10.2478/msds-2013-000453-60](https://doi.org/10.2478/msds-2013-000453-60)
87. P.E. Kloeden, A. Jentzen, Pathwise convergent higher order numerical schemes for random ordinary differential equations. *Proc. Roy. Soc. London A* **463**, 2929–2944 (2007)
88. P.E. Kloeden, V. Kozyakin, Asymptotic behaviour of random Markov chains with tridiagonal generators. *Bull. Aust. Math. Soc.* **87**, 27–36 (2013)
89. P.E. Kloeden, J. Lorenz, Stable attracting sets in dynamical systems and in their one-step discretizations. *SIAM J. Numer. Anal.* **23**, 986–995 (1986)
90. P.E. Kloeden, A. Neuenkirch, The pathwise convergence of approximation schemes for stochastic differential equations. *LMS J. Comp. Math.* **10**, 235–253 (2007)
91. P.E. Kloeden, E. Platen, *Numerical Solutions of Stochastic Differential Equations* (Springer, Berlin, 1992)
92. P.E. Kloeden, M. Rasmussen, *Nonautonomous Dynamical Systems Amer* (Math. Soc, Providence, 2011)
93. P.E. Kloeden, A. Rößler, Runge-Kutta methods for affinely controlled nonlinear systems. *J. Appl. Comp. Math.* **205**, 957–968 (2007)
94. P.E. Kloeden, B. Schmalfuß, Nonautonomous systems, cocycle attractors and variable time-step discretization. *Numer. Algorithms* **14**, 141–152 (1997)
95. M.A. Krasnosel'skij, *The Operator of Translation along The Trajectories of Differential Equations* (Transl. Math. Monogr. Am. Math.Soc., Providence, R.I, 1968)
96. M.A. Krasnosel'skij, J.A. Lifshits, A.V. Sobolev, *Positive Linear Systems* (Heldermann, Berlin, 1989)
97. P. Langevin, Sur la théorie du mouvement brownien. *C.R. Acad. Sci.* **146** (1908)

98. G. Leoni, *A First Course in Sobolev Spaces* (Amer. Math. Soc, Providence, R.I., 2009)
99. B.B. Mandelbrot, J.W. Van Ness, Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* **10**(4), 422–437 (1968)
100. X. Mao, C. Yuan, *Stochastic Differential Equations with Markovian Switching* (World Sci. Publ, Singapore, 2006)
101. M. Marcus, V.J. Mizel, Absolute continuity on tracks and mappings of Sobolev spaces. *Arch. Ration. Mech. Anal.* **45**, 294–320 (1972)
102. J.C. Mattingly, A.M. Stuart, D.J. Higham, Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stoch. Process. Appl.* **101**, 185–232 (2002)
103. R. May, M. Nowak, *Virus Dynamics: Mathematical Principles of Immunology and Virology* (Oxford University Press, Oxford, 2001)
104. H. Meinhardt, *The Algorithmic Beauty of Sea Shells*, 4th edn. (Springer, Berlin, 2009)
105. G.N. Milstein, *Numerical Integration of Stochastic Differential Equations* (Kluwer, 1995)
106. G.N. Milstein, M.V. Tretyakov, *Stochastic Numerics for Mathematical Physics* (Springer, Heidelberg, 2004)
107. G.N. Milstein, M.V. Tretyakov, Numerical integration of stochastic differential equations with nonglobally Lipschitz coefficients. *SIAM J. Numer. Anal.* **43**(3), 1139–1154 (2005)
108. T. Neckel, F. Rupp, *Random Differential Equations in Scientific Computing* (Versita de Gruyter-Verlag, 2014)
109. I. Norros, P. Mannersalo, J.L. Wang, Simulation of fractional Brownian motion with conditionalized random midpoint displacement. *Adv. Perform. Anal.* **2**, 77–101 (1999)
110. R.D. Nussbaum, Finsler structures for the part metric and Hilbert's projective metric and applications to ordinary differential equations. *Differ. Integral Equ.* **7**, 1649–1707 (1994)
111. G. Ochs, Bremen Preprint with Rotating Saddle Points, Universität Bremen (2000)
112. H.-O. Peitgen, D. Saupe, *The Science of Fractal Images* (Springer, New York, 1988)
113. A.S. Perelson, R.M. Ribeiro, Modeling the within host dynamics of HIV infection. *BMC Biol.* **11**, 96,10 (2013)
114. H. Qian, Nonlinear stochastic dynamics of mesoscopic homogeneous biochemical reaction systems - an analytical theory. *Nonlinearity* **24**, R19–R49 (2011)
115. C. Riesinger, T. Neckel, F. Rupp, Solving random ordinary differential equations on GPU clusters using multiple levels of parallelism. *SIAM J. Sci. Comput.* **38**(4), C372–C402
116. L.C.G. Rogers, D. Williams, *Diffusions, Markov Processes, and Martingales*, vol. 2 (Itô Calculus, Wiley, Chichester, 1987)
117. A. Rößler, Rooted tree analysis for order conditions of stochastic Runge-Kutta methods for the weak approximation of stochastic differential equations. *Stoch. Anal. Appl.* **24**, 97–134 (2006)
118. T. Rydén, M. Wiktorsson, On the simulation of iterated Itô integrals. *Stoch. Process. Appl.* **91**(1), 151–168 (2001)
119. W.E. Schiesser, G.W. Griffiths, *A Compendium of Partial Differential Equation Models Method of Lines Analysis with Matlab* (Cambridge University Press, Cambridge, 2009)
120. T. Shardlow, P. Taylor, On the pathwise approximation of stochastic differential equations. *BIT* **56**, 1572–9125 (2016)
121. H.L. Smith, *Monotone Dynamical Systems* (Amer. Math. Soc, Providence, R.I., 1995)
122. H.L. Smith, P. Waltman, *The Theory of the Chemostat: Dynamics of Microbial Competition* (Cambridge University Press, Cambridge, 1995)
123. K. Sobczyk, *Stochastic Differential Equations with Applications to Physics and Engineering* (Kluwer, Dordrecht, 1991)
124. T.T. Soong, *Random Differential Equations in Science and Engineering* (Academic Press, New York, 1973)
125. V. Sree Hari Rao, P. Raja Sekhara Rao, *Dynamic Models and Control of Biological Systems* (Springer, Heidelberg, 2009)
126. S.K. Srinivasan, R. Vasudevan, *Introduction to Random Differential Equations and Their Applications* (North-Holland, New York, 1971)

127. R.L. Stratonovich, *Conditional Markov processes and their application to the theory of optimal control* (Elsevier, Amsterdam, 1968)
128. A.M. Stuart, A.R. Humphries, *Dynamical Systems and Numerical Analysis* (Cambridge University Press, Cambridge, 1996)
129. H.J. Sussmann, On the gap between deterministic and stochastic differential equations. *Ann. Probab.* **6**, 590–603 (1977)
130. A. Swishchuk, *Wu Jianhong, Evolution of Biological Systems in Random Media: Limit Theorems and Stability* (Kluwer Academic Publishers, Dordrecht, 2003)
131. J. Szarski, Differential inequalities. *Monografie Matematyczne*, Tom 43 Państwowe Wydawnictwo Naukowe, Warsaw (1965)
132. D. Talay, Résolution trajectorielle et analyse numérique des équations différentielles stochastiques. *Stochastics* **9**, 275–306 (1983)
133. J.L. Tiwari, J.E. Hobbie, Random differential equations as models of ecosystems: Monte Carlo simulation approach. *Math. Biosci.* **28**, 25–44 (1976)
134. C.P. Tsokos, W.J. Padgett, *Random Integral Equations with Applications to Life Sciences and Engineering* (Academic Press, New York, 1974)
135. B. Verd, A. Crombach, J. Jaeger, Classification of transient behaviours in a time-dependent toggle switch model. *BMC Syst. Biol.* **8**, 1–19 (2014)
136. J. vom Scheidt, H.J. Starkloff, R. Wunderlich, Random transverse vibrations of a one-sided fixed beam and model reduction. *Z. Angew. Math. Mech.* **82**, 847–859 (2002)
137. W. Walter, Differential inequalities, *Lecture Notes in Pure & Appl.*, vol. 129, Math (Dekker, New York, 1991), pp. 249–283
138. M. Wiktorsson, Joint characteristic function and simultaneous simulation of iterated Itô integrals for multiple independent Brownian motions. *Ann. Appl. Probab.* **11**(2), 470–487 (2001)
139. D. Wodarz, N. Komarova, *Computational Biology of Cancer: Lecture Notes and Mathematical Modeling* (World Scientific Publishing, Singapore, 2005)
140. E.W. Wong, M. Zakai, On the relation between ordinary and stochastic differential equations. *Int. J. Eng. Sci.* **3**, 213–229 (1965)
141. J.L. Wonham, Random differential equations in control theory. *Probab. Methods Appl. Math.* **2**, 31–212 (1971)
142. A.T.A. Wood, G. Chan, Simulation of stationary Gaussian processes in $[0, 1]^d$. *J. Comput. Graph. Stat.* **3**(4), 409–432 (1994)
143. I. Ya. Kac, N.N.Krasovskii Kac, On stability of systems with random parameters. *J. Appl. Math. Mech.* **24**, 1225–1246 (1960)
144. H. Zhao, Z.H. Zheng, Random periodic solutions of random dynamical systems. *J. Differ. Equ.* **246**, 2020–2038 (2009)
145. C. Zhu, G. Yin, On hybrid competitive Lotka-Volterra ecosystems. *Nonlinear Anal.* **71**, 1370–1379 (2009)
146. W.P. Ziemer, *Weakly Differentiable Functions* (Springer, New York, 1989)

Index

A

Affinely nonlinear system, 79
Affine structure, 11, 29
Asymptotically stable
 exponentially, 57
 globally, 42
Attracting universe, 49, 52

B

Borel-Cantelli Lemma, 72, 224
B-stability, 155
 implicit averaged Euler, 156
 implicit averaged midpoint, 157
 stochastic implicit multi-step, 158
Burkholder-Davis-Gundy inequality, 71, 225

C

Carathéodory
 conditions, 22, 74
 ODE, 11–13, 74
Chain rule
 affine RODEs, 228
 deterministic, 7
 integral form, 73
 stochastic, 7, 33, 63
 Stratonovich, 11, 34

Chebyshev-Markov inequality, 71
Cocycle
 contractive, 46
 mapping, 42, 50, 53
 strictly uniformly contracting, 46

Coefficient function, 65, 129

Cone, 218

Continuous

© Springer Nature Singapore Pte Ltd. 2017

X. Han and P.E. Kloeden, *Random Ordinary Differential Equations*

and Their Numerical Solution, Probability Theory and Stochastic Modelling 85,

DOI 10.1007/978-981-10-6265-0

absolutely, 12, 22
global Lipschitz, 6
Hölder, 49
uniformly, 22
uniformly Lipschitz, 18

Convergence
 almost surely, 44, 223
 forward, 38
 in probability, 44, 223
 pathwise, 70, 107, 113
 pullback, 38
 strong, 71
 uniformly, 18
 upper semi continuous, 51

D

Derivative
 total, 33, 62
 weak, 22
Differential operator, 63, 110, 112
 iterated, 105
 partial, 65, 73, 129
 total, 102

Discretisation error
 global, 53, 103, 145
 linear, 56
 local, 103
 nonlinear, 56

Doss–Sussmann transformation, 35

Dynamical system
 autonomous, 25
 discrete time autonomous, 52
 discrete time nonautonomous, 53
 driving, 43, 52
 measure-preserving, 49
 metric, 54, 216

- M**
- nonautonomous, 47
 - numerical RDS, 50, 55
 - partial random, 53
 - random, 4, 42
- E**
- Entire solution, 41
 - Equilibrium
 - constant, 37
 - hyperbolic, 53
 - random, 38, 42, 54
 - Ergodic, 54
 - Euler scheme, 5, 9, 102
 - averaged, 5
 - implicit, 50, 53
 - implicit averaged, 156
 - random, 13
- F**
- Filtration, 33
- H**
- Hausdorff semi-distance, 43, 51
 - Hierarchical set, 66, 105
 - Hyperbolic
 - random hyperbolic point, 54
 - steady state, 53
- I**
- Identity
 - function, 65, 86, 111, 130
 - operator, 105
 - Initial value problem, 16, 101
 - integral equation representation, 17
 - solution, 17
 - Integral equation expansion, 86
 - Iterated integral, 84
 - Itô isometry, 31, 32
 - Itô's formula, 7, 33, 63
 - Itô stochastic calculus, 4
- L**
- Lipschitz condition
 - dissipative one-sided, 39, 155
 - global, 19, 32
 - local, 32
 - Lyapunov exponents, 55, 57
- M**
- Manifold
 - stable, 53
 - stable invariant, 57
 - unstable, 53
 - unstable invariant, 57
 - Markov chain, 13, 215
 - discrete-time finite-state, 216
 - Measurable mapping, 20
 - Multi-index, 64, 105
 - concatenated, 67
 - length, 64
 - matrix-valued, 83
 - number of component, 64
 - Multiple integral, 64, 76
 - affine RODE, 130
 - stochastic, 105
- N**
- Noise
 - additive, 35, 133
 - affine, 11
 - bounded, 37
 - canonical, 25, 49
 - centered, 12
 - commutative, 133
 - fictitious, 64, 76, 129
 - Gaussian white, 3, 29
 - linear multiplicative, 36
 - switching, 11
 - Noisy differential equation, 29
 - Non-anticipative, 33
 - Nonautonomous ODE, 37
 - Norm
 - Euclidean, 81, 103
 - exponential, 19
 - Hölder, 82
 - random, 55
 - random operator, 55
 - supremum, 19, 82
 - Numerical dynamics, 49
 - Numerical scheme for affine RODEs, 149
 - affine-RODE-Taylor, 132
 - linear multi-step, 140
 - Runge–Kutta, 137
 - second order derivative-free, 136
 - Numerical scheme for general RODEs
 - essential RODE-Taylor, 146
 - local linearisation method, 6, 150
 - γ -pathwise RODE-Taylor, 144
 - RODE-Taylor, 143
 - the averaged Euler, 152

- Numerical scheme for Itô RODEs
 derivative-free explicit strong, 114
 derivative-free implicit, 117
 implicit RODE-Taylor, 116
 linear multi-step, 118
 B-stability, 158
 consistency conditions, 119
 higher order, 119
 RODE-Adams-Basforth-2, 117
 RODE-Adams-Moulton-2, 118
 strong RODE-Taylor, 112
- Numerical scheme for ODEs
 consistent, 102
 derivative-free, 102
 general one-step, 102
 Heun scheme, 102
 one-step explicit, 103
 p -Taylor scheme, 102
- Numerical scheme for SODEs
 Euler-Maruyama, 106
 Milstein, 106
 one-step, 104
 order γ strong Taylor, 106
- P**
 Path of nonsingularity, 218
 Positive, 24
 off-diagonal positive, 24
 quasipositive, 24
 strongly off-diagonal positive, 25
 strongly positive, 24
 strongly quasipositive, 25
- Positively invariant, 40
 Probability space, 15, 223
 Pullback
 absorbing set, 43
 absorption, 40
 attracting, 43
 attraction, 38, 53
 attractor, 43
 limit, 38
- R**
 Random attractor, 43
 partial, 53
 Random set, 43
 absorbing, 44
 closed, 43
 compact, 43
 tempered, 43
 Remainder set, 66
- Riemann sum, 5, 165
 RODE, 4
 “autonomous”, 74
 canonical form, 42
 coupled RODE-SODE, 117
 Taylor schemes: general case, 143
 with additive noise, 133
 with affine bounded noise, 77
 with affine noise, 73, 124, 129
 with affine structure, 5
 with commutative noise, 133
 with Itô noise, 109
- Rough paths, 36
- S**
 Singleton set, 46
 SODE, 29
 Itô, 31, 63
 solution, 32
 Stratonovich, 4, 11, 34
 Standard assumptions, 69, 106
 Stochastic calculus, 30
 Stochastic integral
 correlated, 10
 Itô, 30
 Itô RODE, 109
 simulation
 compound Poisson, 178
 fractional Brownian motion, 172
 Ornstein–Uhlenbeck process, 165
 Wiener process, 165
 Stratonovich, 34
 Stochastic integral equation, 63
 Stochastic process
 fractional Brownian motion, 167
 Cholesky method, 169
 covariance matrix, 231
 fast Fourier transformation, 169
 random midpoint displacement, 171
 Riemann integral, 172
 h-correlated stationary Gaussian, 4
 Itô process, 35
 Ornstein–Uhlenbeck, 11
 Poisson, 4, 177
 stationary Ornstein–Uhlenbeck, 35, 165
 two-sided Wiener, 37
 Wiener, 26, 30
 Sub-exponential growth, 49
- T**
 Taylor approximation, 7, 102

- affine-RODE, 129
- general affine-RODE, 77
- Itô SODEs, 62
- ODE, 61, 62
- pathwise γ -RODE, 88
- RODE-Taylor, 88
 - essential, 89
- strong Itô–Taylor, 69
- Taylor expansion, 6, 7, 61
 - affine-RODE, 76, 129
 - general Itô–Taylor, 66
 - integral form, 8
 - Itô–Taylor, 66
 - RODE-Taylor
 - truncation error, 88
 - stochastic, 63
 - strong Itô–Taylor, 69
- Taylor scheme
- affine-RODE-Taylor, 130, 132
- γ -pathwise RODE-Taylor, 144
- ODEs, 102
- RODE-Taylor, 143
 - essential, 146
 - for affine RODEs, 149
 - global discretisation error, 145
 - heuristic, 152
 - strong RODE-Taylor, 112
- Theorem
 - Arzelà Selection, 20
 - existence and uniqueness for SODE, 32
 - existence of random attractor, 44
 - Lebesgue dominated, 24
 - Peano’s Existence Theorem, 19
 - The Picard–Lindelöf Theorem, 17
- \mathbb{R}^d , 43, 47