Stochastic systems with time delay

Probabilistic and thermodynamic descriptions of non-Markovian processes far from equilibrium

vorgelegt von

M. Sc.
Sarah Anna Marie Loos
ORCID: 0000-0002-5946-5684

von der Fakultät II – Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
Dr.rer.nat.

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. Martin Schoen

Gutachterin: Prof. Dr. Sabine H. L. Klapp

Gutachter: Prof. Dr. Klaus Kroy

Tag der wissenschaftlichen Aussprache: 13. März 2020

Berlin 2020

Für meine Eltern,

die mir gezeigt haben, dass die
Dinge es wert sind, aus mehr als einem
Winkel betrachtet zu werden.

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Projektnummer 163436311 - SFB 910.

Acknowledgements

First of all, I would like to thank Sabine Klapp for giving me the opportunity to be part of your group, allowing me to work very independently and autonomously, and for supporting me with advice and guidance.
Warm thanks goes to Henning Reinken for supporting my work, and me personally, through all the years. Thank you for all the discussions and ideas concerning this work. I am very grateful for the incredibly nice experience of sharing the office with such a kind and clever person.
I would like to take the opportunity to thank Édgar Roldán, Roland Netz, Klaus Kroy and Victor Holubec, Christian Kuehn, Hannes Stuke, Thomas Ihle, Isabelle Schneider and Bernold Fiedler for inviting me to your seminars, and giving me the opportunity to discuss this work with you and your groups. Furthermore, I would like to express my gratitude to Rainer Klages, Alexander Carmele, and Gernot Schaller for continuous discussions, valuable advice and guidance.
Special thanks also goes to the students Jan Meyer, Simon Hermann, and Timo Doerries, who I had the opportunity to co-supervise and mentor. It was a great experience to be continuously challenged by your questions, and to understand myself better through teaching you. I would further like to thank my colleagues for creating a warm work environment. Special thanks goes to Nima Siboni, Sascha Gerloff, Guo-Jun Liao, Mohsen Khadem, Alexander Kraft, Christopher Waechtler, and Thomas Martynec for various interesting discussions. I thank Florian Buchholz, Felix Herrmann, Henning Reinken, Thomas Martynec, Vera and Peter Loos, Yousof Mardoukhi, Sascha Gerloff, Johanna Strunge, and Maximilian Seyrich for giving me valuable feedback on this thesis, and for your ideas and comments.
Lastly, I would like to express my gratitude for the constant support by my family and friends. I am especially thankful to you, Felix. You were always there when I needed you, and you never let me down.

Zusammenfassung

Neben zufälligen Fluktuationen aufgrund des allgegenwärtigen *Rauschens* in unserer Welt, spielen in vielen natürlichen und technologischen Prozessen *Gedächtnis*effekte, d. h. Geschichtsabhängigkeiten, eine Rolle. Gedächtnis kann beispielsweise in Form von Zeitverzögerungen auftreten, die wiederum von endlich schnellen Informationsübertragungen oder Signallaufzeiten hervorgerufen werden. Durch die Kombination von Rauschen und Zeitverzögerung, oder Gedächtnis im Allgemeinen, entsteht nicht-Markovsche Dynamik, welche mit Hilfe von zeitlich nichtlokalen Langevin-Gleichungen mit delta-verteilten Gedächtnisfunktionen beschrieben werden kann. Trotz ihrer Omnipräsenz sind stochastische Systeme mit Gedächtnis noch nicht gut verstanden. Das gilt vor allem, wenn zusätzlich nichtlineare Kräfte auf das System wirken, wie es in realistischen Situationen im Grunde immer der Fall ist.

In dieser Arbeit untersuchen wir ein Brownsches Teilchen, auf welches sowohl nichtlineare Kräfte als auch eine zeitverzögerte Rückkopplungskontrolle wirken. Damit betrachten wir ein experimentell zugängliches und analytisch behandelbares Beispiel, das als Prototyp für ein nicht-Markovsches, intrinsisches nicht-Gleichgewichtssystem angesehen werden kann. Wir konzentrieren uns auf zwei wichtige Eckpfeiler der Statistischen Physik: *Wahrscheinlichkeitsbeschreibungen* und *Thermodynamik*. Beide theoretische Konzepte haben sich als ausgesprochen wertvoll zur Beschreibung passiver Teilchen mit zeitlich lokalen Gleichungen erwiesen, und beide sind bislang für Systeme mit Gedächtnis unzureichend verstanden.

Es wurde bereits vor über 20 Jahren gezeigt, dass die Fokker-Planck-Beschreibung für Systeme mit Zeitverzögerung eine unendliche Hierarchie von gekoppelten Gleichungen liefert. Wir vergleichen verschiedene, in früherer Literatur vorgeschlage Herleitungen dieser Hierarchie. Darüber hinaus entwickeln wir einen neuen, konzeptionell einfacheren Ansatz, der auf einer sogenannten *Markovschen Einbettung* basiert. Außerdem betrachten wir das erste Glied der unendlichen Hierarchie genauer. Diese Gleichung wurde in der Vergangenheit benutzt um Approximationen für die einzeitige Wahrscheinlichkeitsdichte zu entwickeln. Ein konzeptioneller Nachteil bekannter Verfahren besteht darin, dass sie effektiv Markovsche Beschreibungen liefern, und somit die Essenz der Prozesse mit Gedächtnis nicht wirklich zu erfassen vermögen. Um dieses Problem anzugehen, schlagen wir ein neues Schema vor, welches wir "Force-linearization closure" nennen. Dieses Verfahren liefert eine echte nicht-Markovsche Beschreibung auf Kosten der Vernachlässigung von Nichtlinearitäten in den höheren Ordnungen. Durch die Anwendung auf ein bistabiles und auf ein periodisches Modell zeigen wir, dass die resultierenden einzeitigen Wahrscheinlichkeitsdichten genauer sind als die von früheren Approximationen, nämlich der Störungstheorie oder der sogenannten "Small delay expansion". Weiterhin zeigen wir, dass die Erkenntnisse, die eine isolierte Betrachtung der einzeitigen Wahrscheinlichkeitsdichte allein liefern, begrenzt sind. Dies liegt daran, dass die Dynamik entscheidend durch das Gedächtnis bestimmt wird, also von nichttrivialen zeitlichen Korrelationen. Aus diesem Grund untersuchen wir auch die höheren Glieder der Hierarchie. Insbesondere schlagen wir einen neuen Ansatz vor, der auf die zweizeitige Verbundswahrscheinlichkeitsdichte abzielt. Dieser neue Ansatz erlaubt das Beschreiben von dynamischen, nicht-Markovschen Effekten, z. B. von durch das Gedächtnis induzierten Oszillationen. Soweit uns bekannt, ist dies der erste Ansatz auf Fokker-Planck-Ebene, der solch komplexe nicht-Markovsche Dynamik erfassen kann.

Danach wenden wir uns der Betrachtung thermodynamischer Größen zu. Hierbei konzentrieren wir uns zunächst auf die Wärmeproduktion im stationären Zustand. Die passenden Formeln für diese Größe liefert die stochastische Thermodynamik. Wir zeigen, dass die Rückkopplungskontrolle zwangsläufig zu Wärmeproduktion führt, selbst wenn die Zeitverzögerung verschwindend klein ist. Wir führen diese Wärmeproduktion bei kleinen Verzögerungszeiten auf das Phänomen des "Entropiepumpens" zurück. Letzteres ist für Markovsche Systeme bereits ein bekanntes Phänomen, das durch geschwindigkeitsabhängige Kräfte hervorgerufen wird. Als konkretes Beispiel für ein nicht-Markovsches System betrachten wir wieder das bistabile Modell. Wir zeigen, dass die Rückkopplung sowohl wärmend, als auch abkühlend wirken kann. Des Weiteren weisen wir nach, dass die mittlere Mediumsentropieproduktion ein Maximum bei Kohärenzresonanzbedingungen hat.

Zum Schluss wollen wir die Irreversibilität des Prozesses quantifizieren. Hierbei stoßen wir auf

ein grundlegendes Problem: die Berechnung der totalen Entropieproduktion nicht-Markovscher Systeme. Als Versuch, dieses Problem für die hier betrachteten Beispiele anzugehen, nehmen wir die externe Kontroll-Apparatur explizit mit in das Modell auf. Zu diesem Zweck lösen wir uns von den Systemen mit diskreter Zeitverzögerung und betrachten nun Systeme mit verteiltem Gedächtnis und farbigem Rauschen. Diese können durch Markovsche Netzwerke (über die Markovsche Einbettung) dargestellt werden, was wiederrum die Anwendung des etablierten Standardformalismus ermöglicht. Wir zeigen, dass eine nicht-monoton abfallende Gedächtnisfunktion in der Langevin-Gleichung untrennbar mit nicht-reziproken Kopplungen im zugehörigen Netzwerk verbunden ist. Des Weiteren zeigen wir, dass bereits das alleinige Auftreten einer nicht-reziproken Kopplung eine endliche Entropieproduktion mit sich bringt. Wir leiten einen verallgemeinerten zweiten Hauptsatz der Thermodynamik her, der den auftretenden Informationsfluss enthält. Der ursprünglich betrachtete Fall der diskreten Zeitverzögerung stellt hier den Grenzfall eines unendlich großen Markovschen Netzwerks dar. In diesem Grenzfall finden wir eine divergierende Entropieproduktion. Dieser Fund impliziert unbegrenzte Kosten für das präzise Speichern einer Brownschen Trajektorie. Darüber hinaus zeigen unsere Überlegungen Zusammenhänge zwischen Gedächtnis, Rückkopplung und Aktivität auf. All diese Phänomene sind durch dasselbe vereinigende Konzept beschreibbar, nämlich durch nicht-reziproke Kopplung.

Abstract

In addition to the ever-present random fluctuations due to *noise*, many systems in biology, physics and technology involve discrete *time delay*, stemming, e.g., from finite information transmission times. The combination of noise and delay yields *non-Markovian* dynamics describable by time-nonlocal Langevin equations with a delta-peaked memory kernel. Despite their omnipresence, the understanding of stochastic systems with time delay is yet in its infancy. This is especially true in the presence of *nonlinear* forces, which are ubiquitous in realistic situations. In this thesis, we consider a Brownian particle subject to nonlinear forces and time-delayed feedback. This represents an experimentally accessible and analytically treatable prototype of a non-Markovian, intrinsically nonequilibrium system. We consider two important cornerstones of statistical physics: *probabilistic descriptions* and *thermodynamic laws*. Both have proven themselves to be valuable frameworks to describe passive particles with time-local equations. And both are, so far, not well-understood for time-delayed systems.

It is already known that the Fokker-Planck description for delayed systems is given by an *infinite hierarchy* of coupled equations. We compare different derivations of this hierarchy, and further present a new, simpler approach by means of a Markovian embedding technique. Further, we consider the first member of the hierarchy, which has been used in earlier literature as starting point for approximation schemes aiming for the one-time probability density of nonlinear systems. One conceptual disadvantage of the existing approximations is that they effectively render Markovian descriptions, thus, do not capture the essence of the process with memory. To tackle this problem, we propose a new scheme, called the force-linearization closure, which yields a non-Markovian description at the cost of negligence of nonlinear effects in higher-order terms. By application to a bistable and a periodic model, we demonstrate that the resulting one-time probability densities are more accurate than earlier approximations from the literature, in particular, the perturbation theory and the small delay expansion. We further show that the insights provided by the one-time probability alone are limited. This is because the dynamics is crucially determined by the memory, and, subsequently, the temporal correlations. In this spirit, we also shed light on the higher members of the hierarchy, and introduce a new approach aiming for the two-time joint probability density, allowing us to finally describe non-Markovian dynamical effects, e.g., delay-induced oscillations. To the best of our knowledge, this is the only approach on the Fokker-Planck level yet that captures this dynamical state.

Turning to thermodynamic quantities, we first focus on the net steady-state heat flow, which is accessible via the framework of stochastic thermodynamics. We show that the feedback inevitably leads to a finite heat flow, even for vanishingly small delay. We link the heat flow at small delay times to "entropy pumping", which is a phenomenon that is already established for Markovian systems subject to velocity-dependent driving. Application to our bistable model further reveals that the feedback induces heating as well as cooling regimes, and leads to a maximum of medium entropy production at coherence resonance conditions. We further discuss the possibility to quantify the irreversibility and the distance from equilibrium. Here we immediately encounter a fundamental problem, i.e., the notion of total entropy production of non-Markovian systems. As an attempt to address this problem, we additionally include the feedback controller state explicitly in the model. To this end, we step away from the discrete delay and consider systems subject to distributed delay and colored noise. These systems can be represented (via Markovian embedding) by *Markovian networks*, which can be treated by the standard formalism. We demonstrate that a *non-monotonic* memory kernel in the Langevin equation is inextricably linked to *non-reciprocal* coupling. Further, we show that the non-reciprocity alone implies a finite entropy production. We derive a generalized second law involving information flow. Discrete time delay is recovered as the infinite-dimensional limit. Here the entropy production diverges, implying unbounded cost for precisely storing a Brownian trajectory. Our considerations further reveal connections between memory, feedback, and activity, and show that these phenomena are all describable by the same unifying concept, i.e., non-reciprocal coupling.

Contents

Contents **I**

Publications **VII**

Lists of Abbreviations & Symbols **VIII**

Introduction **1**
Outline of the thesis 9

I Theoretical background & State of the art 13

1 The Langevin equation **14**
1.1 Stochastic modeling 14
1.1.1 Brownian motion 14
1.1.2 Colloidal suspensions 15
1.1.3 Side note: A more general view 16
1.2 The Markovian Langevin equation 16
1.2.1 Gaussian white noise 18
1.2.2 Ensemble averages & probability density 20
1.2.3 Solutions of the Langevin equation and the overdamped limit 21
1.2.4 Ornstein-Uhlenbeck process 24
1.2.5 White noise – Wiener process – Stochastic calculus 25
1.2.6 Path integral representation 27
1.3 How stochastic motion emerges from first principles 28
1.3.1 Infinite harmonic oscillators bath – An example of a Mori-Zwanzig projection 28
1.3.2 Coarse-graining – Forgetting some details 30
1.3.3 The Markov assumption 34
1.3.4 Real-world complications 35
1.3.5 Time-reversal symmetry & Causality 36
1.4 Introduction to the Langevin equation with time delay 38
1.4.1 Optical tweezers – An experimental tool to control 40
1.4.2 Time-delayed feedback 41
1.4.3 The Langevin equation with time delay 42

1.4.4 Side note: Delay differential equations . . . 43
1.4.5 Linear systems with time delay . . . 46
1.4.6 Autocorrelation function of linear delay equations . . . 47
1.5 Nonlinear example systems with time delay . . . 51
1.5.1 Bistable system: The doublewell potential . . . 51
1.5.2 Periodic system: The washboard potential . . . 52
1.5.3 Scaling . . . 54
1.6 Timescales . . . 55
1.6.1 Kramers escape times . . . 57
1.7 Delay-induced oscillations & Coherence resonance in nonlinear systems 58
1.7.1 Delay-induced oscillations . . . 58
1.7.2 Coherence resonance . . . 59
1.7.3 Bifurcation theoretical perspective on delay-induced oscillations 60

2 Fokker-Planck equations 62
2.1 Markovian case . . . 62
2.1.1 Natural boundary conditions . . . 63
2.1.2 Joint probability densities . . . 63
2.1.3 The probability current and steady states . . . 64
2.2 Introduction to Fokker-Planck descriptions of systems with time delay 66
2.2.1 Earlier approximation schemes . . . 67
2.2.2 Probability current and apparent equilibrium of time-delayed systems . . . 69
2.2.3 Side note: Delay in the probability density . . . 69

3 Stochastic thermodynamics 71
3.1 Side note: Some historical notes & Where is stochastic thermodynamics located in theoretical physics? . . . 72
3.2 Stochastic energetics . . . 73
3.2.1 Steady states . . . 75
3.3 Fluctuating entropy . . . 77
3.3.1 Thermal equilibrium & Nonequilibrium steady states . . . 79
3.4 Fluctuation theorems . . . 81
3.4.1 Route to first principles – Axiom of causality . . . 82
3.5 Information . . . 83
3.6 Previous results, expectations and apparent problems for systems with time delay . . . 85
3.6.1 Energetics in the presence of delay . . . 85
3.6.2 Entropic description . . . 88
3.6.3 The acausality issue . . . 88
3.6.4 Short review on Effective thermodynamics . . . 90
3.7 Side note: Active particles & Non-reciprocal interactions . . . 92
3.7.1 Active Ornstein-Uhlenbeck particles . . . 93
3.7.2 Connection between active matter and time-delayed systems . 94
3.7.3 Non-reciprocal interactions . . . 95

II Probabilistic descriptions for systems with time delay 97

4 Infinite Fokker-Planck hierarchy 98
- 4.1 Derivation of Fokker-Planck hierarchy from Novikov's theorem 98
  - 4.1.1 Alternative approach with two time arguments 101
- 4.2 Exact probabilistic solutions for linear systems with time delay 102
  - 4.2.1 Derivation of the second member of the Fokker-Planck hierarchy 103
  - 4.2.2 Steady-state solutions 106
  - 4.2.3 The notion of effective temperature 110
  - 4.2.4 Markovian vs. non-Markovian two-time probability density . . 111

5 A new derivation via Markovian embedding 113
- 5.1 Markovian Embedding 115
  - 5.1.1 Projection, Memory kernel & Colored noise 115
  - 5.1.2 Limit $n \to \infty$ 118
  - 5.1.3 Interpretation of the X_j variables 120
  - 5.1.4 Initial condition 120
  - 5.1.5 $(n+1)$-dimensional Markovian Fokker-Planck equation 121
- 5.2 Derivation of first member of Fokker-Planck hierarchy 121
- 5.3 Derivation of higher members via Markovian embedding 123
  - 5.3.1 Comparison to equation from Novikov's theorem 126
- 5.4 Side note: Discrete vs. distributed delay 127
  - 5.4.1 Trajectories of the delayed Ornstein-Uhlenbeck process 128
  - 5.4.2 Probability densities in the presence of discrete and distributed delay 129

6 Force-linearization closure 132
- 6.1 Details of the approximation 133
  - 6.1.1 Linearization of the deterministic forces 133
  - 6.1.2 Analytical probabilistic solution for linearized forces 134
  - 6.1.3 Vanishing steady-state probability current 135
  - 6.1.4 Specification to linear delay force 136
- 6.2 Comparison to earlier approaches 137
  - 6.2.1 Small delay expansion 137
  - 6.2.2 Perturbation theory 137
  - 6.2.3 Effective temperatures 139
- 6.3 Application to the periodic potential 139
  - 6.3.1 Discussion of results 141
- 6.4 Application to the bistable potential 142
  - 6.4.1 Discussion of results 143
- 6.5 Estimation of escape times 145
  - 6.5.1 Interwell dynamics 145
  - 6.5.2 The Kramers-FLC estimate 147
  - 6.5.3 Comparison with numerical results 148
  - 6.5.4 Delay-induced oscillations 149
  - 6.5.5 Side note: Normal diffusion despite non-Markovianity 150

7 Approximate two-time probability density 153
7.1 Application to the bistable delayed system . . . 155
7.1.1 Comparison with approaches for one-time probability density . 157
7.2 Concluding remarks . . . 159

III Thermodynamic notions for systems with time delay 163

8 The heat flow induced by a discrete delay 164
8.1 Main idea . . . 165
8.1.1 Polynomial energy landscapes . . . 165
8.2 Mean heat rate & Medium entropy production . . . 166
8.2.1 Linear systems . . . 168
8.2.2 Markovian limits in nonlinear systems . . . 170
8.2.3 Limit of vanishing delay time . . . 172
8.2.4 Discussion of the behavior for small delay times . . . 174
8.3 Application to the bistable potential . . . 176
8.3.1 Low thermal energy – intrawell dynamics . . . 177
8.3.2 High thermal energy – interwell dynamics . . . 178
8.4 Preliminary numerical results for fluctuation of heat, work and internal energy . . . 187
8.5 Concluding remarks . . . 189

9 Entropy, Information & Energy flows 190
9.1 Emergence of non-monotonic memory . . . 192
9.1.1 Interpretation of $X_{j>0}$ in the case of a feedback controller . . . 194
9.2 The role of non-reciprocal coupling – Connection to active matter . . . 196
9.2.1 A generic model with non-reciprocal coupling . . . 196
9.3 Non-reciprocal coupling & Non-equilibrium . . . 198
9.3.1 Fluctuation-Dissipation relation . . . 199
9.3.2 Total entropy production . . . 199
9.3.3 Analytical solutions . . . 203
9.4 Non-reciprocal coupling & Activity . . . 204
9.4.1 "Mild cases" . . . 205
9.4.2 Reversed heat flow . . . 206
9.5 Non-reciprocal coupling & Information . . . 208
9.5.1 Information flow and generalized second law . . . 208
9.5.2 Information-theoretic perspective on feedback control . . . 212
9.6 Total entropy production and heat flow in the presence of non-monotonic memory . . . 213
9.6.1 Limit of discrete delay . . . 215
9.6.2 Impact of measurement errors . . . 217
9.7 Irreversibility & Coarse-graining . . . 219

IV Concluding remarks **221**

10 Summary **222**

11 Outlook **228**

A Appendix **233**
A.1 Numerical methods 233
A.2 Derivation of Novikov's theorem 234
A.3 Green's function method 235
A.4 Connection to Fokker-Planck hierarchy from Novikov's theorem . . . 236
A.5 Fluctuation-dissipation relation for unidirectional ring of arbitrary length n 237

Bibliography **XI**

Publications

This thesis is based on three publications, and one preprint

[1] Sarah A. M. Loos and Sabine H. L. Klapp
Force-linearization closure for non-Markovian Langevin systems with time delay
Physical Review E 96, 012106 (2017)

[2] Sarah A. M. Loos and Sabine H. L. Klapp
Heat flow due to time-delayed feedback
Scientific Reports 9, 2491 (2019)

[3] Sarah A. M. Loos and Sabine H. L. Klapp
Fokker-Planck equations for time-delayed systems via Markovian Embedding
Journal of Statistical Physics, 177, 95-118 (2019)

[4] Sarah A. M. Loos, Simon M. Hermann, and Sabine H. L. Klapp
Non-reciprocal hidden degrees of freedom: A unifying perspective on memory, feedback, and activity,
ArXiv:1910.08372 (2019)

List of Abbreviations

Abbreviation	Description
LE	Langevin equation
FP, FPE	Fokker-Planck, Fokker-Planck equation
d.o.f.	degree of freedom
PDF	Probability density function
OU	Ornstein-Uhlenbeck process
BD	Brownian dynamics simulation
ss	steady state (used as subscript)
MSD	Mean square displacement
(S)DDE	(Stochastic) delay differential equation
FLC	Force-linearization closure
PT	Perturbation theory
small τ	Small delay expansion
FDR	Fluctuation-dissipation relation
EQ	Thermodynamic equilibrium
NESS	Nonequilibrium steady state
EP	Entropy production
IFT	Integral fluctuation theorem
CR	Coherence resonance
AOUP	Active Ornstein-Uhlenbeck particles
NR	Non-reciprocal
Fig., Eq., Sec., Ref.	Figure, Equation, Section, Reference
l.h.s., r.h.s.	left hand side, right hand side
w.r.t.	with respect to
i.e.	Latin *id est*, which means "that is"
e.g.	Latin *exempli gratia*, which means "for example"

List of Symbols

Symbol	Description	SI unit
t	Time	s
τ	Delay time	s
τ_{ir}	Intra-well relaxation time	s
τ_{K} , r_{K}	Kramers escape time, rate	s, 1/s
J	Probability current	1/s
τ_{B}	"Brownian" time $\tau_{\mathrm{B}} = \sigma^2/D_0$, *used as time scale*	s
x	Space coordinate	m
X	Stochastic variable, particles position	m
$\underline{X}$	Stochastic vector $\{X_0, ..., X_n\}$	m^{n+1}
$\mathbf{X}$, $\underline{\mathbf{X}}$	Stochastic path, i. e., random trajectory	m, m^{n+1}
ϕ	History function (initial condition)	m
Ω	Domain of x	m
$C(t')$	Position-autocorrelation function, $\langle X(t)X(t+t')\rangle$	m^2
σ	Particle diameter, characteristic length of potential, *used as length scale*	m
ρ	Probability density function (PDF)	1/m
ρ_n	n-point probability density	$1/\mathrm{m}^n$
ρ_{c}	Conditional probability density	1/m
Z	Normalization constant	m
P	Probability	1
$\mathcal{P}$	Path probability	1/m, $1/\mathrm{m}^n$
$\langle ... \rangle$	Ensemble averages	–
γ	Friction coefficient	kg/s
D_0	Diffusion coefficient	m^2/s
ξ	Gaussian white noise	$1/\sqrt{\mathrm{s}}$
$\mathcal{T}$	Temperature	K
k_{B}	Boltzmann constant $\sim 1.38 \times 10^{-23} \mathrm{kg\,m^2/(s^2K)}$	$\mathrm{kg\,m^2/(s^2K)}$
V	Potentials	$\mathrm{kg\,m^2/s^2}$
V_0	Barrier heights	$\mathrm{kg\,m^2/s^2}$
F	Forces	$\mathrm{kg\,m/s^2}$
$k_{\mathrm{B}}\mathcal{T}$	Thermal energy, *used as energy scale*	$\mathrm{kg\,m^2/s^2}$
u, q, w	Fluctuating inner energy, heat, work	$\mathrm{kg\,m^2/s^2}$
S_{Sh} , S_{tot}	Shannon and total entropy	$\mathrm{kg\,m^2/(s^2K)}$
$\mathcal{I}$	Mutual information	1
$\dot{I}$	Information flow	1/s
N	Number of realizations (simulations)	1
Δx, Δt	spatial step size, temporal step size (simulations) in units of $\sigma, \tau_{\mathrm{B}}$	–

Introduction

"[F]undamentally, feedback is one of the great ideas developed (mostly) in the last century, with particularly deep consequences for biological systems"

– John Bechhoefer, 2005 [8]

Many natural and artificial systems are governed by dynamical equations which involve both, noise and time delay [9–11]. Noise due to imperfections or random environmental influences is essentially omnipresent in real-world systems and experimental setups. From a mathematical point of view, it often results from the presence of hidden degrees of freedom that have disappeared in the course of a coarse-graining procedure. Stochastic models that account for the noise, are well-known to be a powerful tool to effectively describe complex systems by using just relatively few, *mesoscopic* (stochastic) degrees of freedom. Such mesoscale models, for example the *Brownian motion* studied in this thesis, also play a central role in statistical physics. However, many theoretical concepts rely on the Markov assumption, neglecting the fact that realistic systems often have *memory*, i.e., history-dependence, due to, e.g., viscoelasiticity [12–15], hydrodynamics [16–18], persistence [19–21], or due to *time delay*.

Time delay can have various sources. A prominent example from electrodynamics are the retarded potentials (e.g., Liénard-Wiechert potentials [22, 23]), where

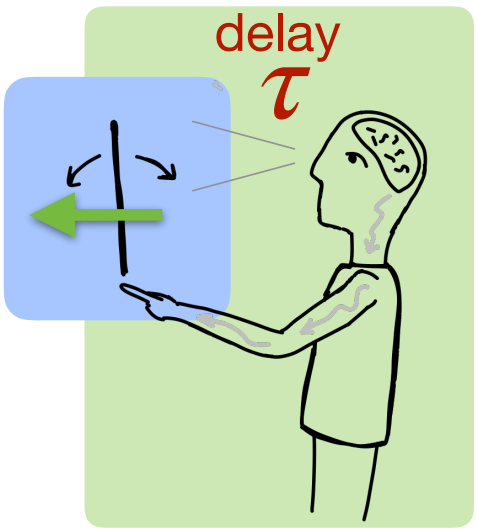

Figure 0.1: *A well-known control problem, where time delay matters: balancing. As an example, think of the attempt to balance a stick on your finger tip [5–7]. The delay τ is created within the control loop consisting of your eyes, the neural transmission to and back from your brain (or spine), plus the information processing and the muscle reaction time. As everybody knows, it is very difficult to balance the stick, but practice improves the performance, as the delay can be reduced.*

Figure 0.2: *The basic principle of feedback control: feeding back the output signal into a system, via a so-called feedback loop.*

delay arises due to the finite speed of light. In biological systems, maturation times can create delays (as in population growth [24] and prey-predator dynamics [25]). On the bio-molecular level, they arise in chemical reaction kinetics (e.g., cell metabolism [26–28]). Neural delays caused by finite refractory times [29–32] can even become apparent on a macroscopic level, as in the human pupil reflex [30], or in stick balancing experiments [5, 6]. The latter example, which is sketched in Fig. 0.1, illustrates that delay can have a destabilizing effect (the stick can be stabilized longer in the upright position, if the reaction time delay is shorter). In socio-economic systems, information processing and communication times create delays [33] (e.g., financial markets [34–36]).

Beyond these intrinsic delays, time delay is an important topic in control theory, particularly in the context of *feedback*, i.e., a driving that depends on an (earlier) system state [8, 10] (Fig. 0.2). Feedback protocols have an unavoidable time lag between signal detection and action of control [8, 10, 11, 37–40]. An everyday example is the adjustment of the shower temperature (see Fig. 0.3). This example further illustrates that delay may induce oscillatory behavior, which is quite characteristic for delay dynamics. Similarly, they arise in laser systems (due to the traveling time of light in the cavity [40–42]), and quantum-optical systems (coupled to a structured photonic reservoir [43–45]). Because delays induce complex behavior like oscillations, chaos, and can even have stabilizing effects (e.g., on unstable orbits [46–48]), they are also implemented intentionally (as in Pyragas control [10]). Besides numerous technical applications, feedback is a very important mechanism that is naturally encountered in biological systems [49–54], as in gene regulatory networks [55–58]. An important example is the chemotactal motion of bacteria which relies on feedback which involves delay [49, 59, 60]. In fact, delay often occurs because of feedback, and feedback quasi always includes delay[1].

In this thesis, we investigate a colloidal particle subject to nonlinear static forces and a position-dependent, time-continuously operating feedback force with a single, *discrete* delay $\tau \geq 0$. Experimentally, such a feedback can be realized with the help of a co-moving *optical trap* [61–66]. The control "target", i.e., the particle position, is accessible by an optical microscope due to the typical (micron-scale) size of a colloid [67–69]. The effect of the trap can be well approximated by a *linear* force towards the delayed position, with a tunable delay time[2] [61, 62]. We model the dynamics of the colloid with time- and space-continuous equations, and consider white noise. Due to the delay, the equation of motion, that is, the (overdamped) Langevin equation (LE), becomes time-nonlocal, or, in a more physical language,

[1] Sometimes its rather a matter of preference, which aspect is emphasized.

[2] A small time delay arises naturally, and it can further be artificially enlarged to a desired value within the resolution of the camera.

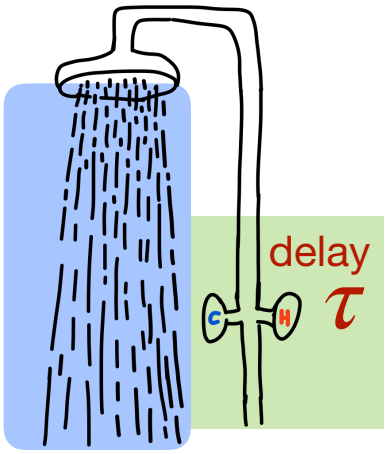

Figure 0.3: *An everyday control problem, where we experience the impact of delay: Adjusting the water temperature in your shower. Typically, there is a time lag between adjustment of temperature controller, and actual change of temperature. This can yield unpleasant oscillations between hot and cold (especially when the time lag is long).*

non-Markovian. This type of description is also appropriate to model various other systems, including some of the aforementioned examples; and this type of delay LE is subject of numerous recent theoretical works [40, 70–89]. The discrete delay can further be considered as a special case of distributed delay, where the memory kernel is a delta-distribution around τ. From a mathematical point of view, the delay LE is a stochastic *delay* differential equation (SDDE). It is well-known that SDDEs are *infinite-dimensional*[3] [10, 90]. The usual concepts and solutions for Markovian LEs then do not apply and the development of a general solution method for SDDEs is yet an unresolved problem. This is especially true in regard to the interplay of noise, delay, and *nonlinear* forces, which are omnipresent in realistic situations. Only for the special case of purely linear equations, closed-form solutions for probability density functions, position-correlations and stability boundaries were obtained in earlier literature [9, 77, 84, 89, 91–93].

As we have seen, despite the fact that stochastic (nonlinear) delay systems are naturally encountered in various contexts, there are still many open problems in their mathematical description. Moreover, the incorporation of time delays, or generally memory, in the theoretical frameworks of statistical physics is still in its infancy.

Statistical Physics Perspective

There is indeed a class of (mesoscale) systems with noise and memory that is already well-established in statistical physics, which for example comprises colloids in fluids with viscoelastic properties or hydrodynamic backflows. This class of systems is typically modeled by generalized Langevin equations [94], which also feature memory, i.e., time non-local terms. Specifically, they involve a memory kernel in the friction force and finite noise correlations (colored noise) resulting from conservative interactions between a colloidal particle (or, more generally, the "slow" variable) and the bath particles. As a consequence, the delayed friction is related to the noise via a fluctuation-dissipation relation [95], which is closely connected to the fact that the system may approach thermal equilibrium (see, e.g., [15, 96–98]).

[3] as reflected by an initial condition with infinitely many degrees of freedom

It is important to note that the type of memory treated in this thesis is crucially different. A fundamental difference is that there is no relation between the delay and noise correlations. From the simple comparison with generalized LEs alone, one may already suspect that a system subject to a time-delayed force is *intrinsically out of thermal equilibrium*, even in the absence of additional drivings. This guess turns out to be true, as was previously shown [99], and as we will also demonstrate by various means throughout this thesis. The system is pushed out of equilibrium by a *history-dependent* driving, yielding a crucially different state than time-dependent (or constant) drivings, which are much better understood.

The property of being intrinsically out of equilibrium already on the level of individual constituents is a feature that feedback-controlled systems have in common with *active systems*. This class of systems, which includes robots, active microswimmers, or actively moving motor proteins like kinesin [100, 101], is a prime focus of the latest research in nonequilibrium statistical physics [19, 102–110]. There is another –less obvious– shared feature, that is, the presence of memory. While the discussion of memory, or delay, in feedback loops has a long tradition, this viewpoint is less common in the context of active motion. However, the importance of history-dependence becomes apparent in the form of *persistence* of active swimmers [19–21], the viscoelastic properties of active gels [18], or the giant Kovacs-like memory effect in the Vicsek model [111]. Despite these apparent similarities, the theoretical description has essentially been worked out separately. Only recently, researchers have started to realize the delicate connections between both [38, 50, 66]. For example, recent experiments [66] and theoretical studies [89, 112] have demonstrated that time-delayed feedback can be used to *mimic active particles*. We believe that a further cross-fertilization between the fields would be very important not only for the understanding of individual elements, but also for the collective behavior on larger scales; a topic that is already of major interest for active systems [20, 38].

From the viewpoint of statistical physics, one may consider our time-delayed stochastic systems as well as active small-scale systems as "novel types of particles". Let us explain this in more detail. The theoretical description of the complex, living or artificial world surrounding us is the main objective of modern statistical physics. The building blocks of this world are subsystems which often operate far from thermal equilibrium themselves, already on the individual level[4]. Experiments and computer simulations show that such nonequilibrium systems display intriguing behavior from the individual to the collective level. For example, recent experiments [38, 66] and numerical studies [110, 113, 114] focus on the collective behavior of time-delayed systems, where delay often arises within the interactions between individual constituents. In [38], large swarms of robots with delayed communications were studied. Similarly, optical feedback was employed to stabilize "colloidal molecules" [66, 89]. These are molecule-like non-equilibrium structures consisting

[4]For such systems, the traditional picture of a (passive) system subject to external ("open-loop") forces is not appropriate, as the drive is *internal*, or in other words, correlated with the system state.

of a few Brownian particles, which exhibit bond length oscillations due to the delay. While these studies offer interesting insights into the statistical physics of nonequilibrium building blocks, the theoretical understanding is still in its infancy. This is because the common assumptions that are the basis of many established theoretical results are violated. Traditionally, the frameworks of statistical physics were designed for passive particles with time-local interactions, like hard disks, harmonic springs, billiard balls or Lennard-Jones particles [115], describable by Hamiltonian equations on the underlying, microscopic level. This implies *reciprocal* (conservative) interactions, which fulfill Newton's third law, *actio = reactio*. Such systems can reach thermal equilibrium, if no external driving prevents them.

In contrast to those passive particles, one may consider the aforementioned intrinsically nonequilibrium systems to represent a "novel type of particle". For such intrinsically nonequilibrium building blocks, the established frameworks of statistical mechanics need to be revisited. Theoretical challenges include the absence of an underlying Hamiltonian description, non-Markovianity, non-reciprocal interactions (which we will come back to at the end of this introduction), and the lack of a variational principle with suitable thermodynamic potentials. Colloids subject to (optical) feedback may be viewed as prototypes of a "novel type of particle", particularly useful to study the effect of non-Markovianity (the time delay represents a "tunable memory", which, by itself, already drives them out of equilibrium).

Two particularly relevant open problems in the context of modern statistical physics are the *probabilistic description* and the generalization of *thermodynamics*. Both, Fokker-Planck equations and thermodynamic laws have proven themselves of outstanding value for the development of traditional statistical physics of passive particles. Thus, it is natural to also reach out for Fokker-Planck descriptions and thermodynamics notions of time-delayed stochastic systems. This is the objective of the investigations presented in this thesis. In the following, we briefly lay out the main challenges and research questions.

Probabilistic treatment

It is already known for more than 20 years that for systems with time delay, the Fokker-Planck equation (FPE) of the one-time probability density function (PDF) does not provide a complete, self-contained probabilistic description, because it explicitly involves the two-time PDF. In fact, it represents only the first member of an infinite hierarchy, mirroring the infinite-dimensionality of the underlying LE. Contrary to Markovian systems, the route from the LE to a FPE is more involved [82, 86, 87, 92, 116]. Therefore, writing down explicit expressions for the various members of this hierarchy is already a nontrivial problem. In earlier literature, we could find a general expression solely for the first member. We will consider this mathematical problem from different angles and explore strategies to nevertheless obtain analytical expressions for the FPEs themselves and the PDFs, specifically for *nonlinear* overdamped systems.

As mentioned above, the only class of delayed noisy systems where closed-form

solutions are known, are those with linear forces. Here, all the n-time probability densities are given by multivariate Gaussian distributions. Exact solutions can be obtained on the LE [77] and the FPE level [92], or by utilizing both levels of description [84]. In the more general case of nonlinear systems, one has to rely on approximations [9]. The two most established strategies are the "small delay expansion", that is a Taylor expansion of the LE in τ [82, 83], and a perturbation theory on the level of the FPE [86, 87], where the entire delay force is treated as a small deviation from the Markovian dynamics. A conceptual disadvantage of these approaches is that they effectively render Markovian descriptions. This motivates us to ask: Is there a way to approximate the dynamics, but in a non-Markovian way? In this work, we develop such an approach. The trick is to exploit the closed-form PDF solutions from the linear delay case. By application to two generic nonlinear systems, a colloidal particle in a sinusoidal or in a bistable potential supplemented by a linear delay force, we will compare our new approximation (called "force-linearization closure") against the earlier approaches.

We will further show that some important features of the dynamics are caused by the interplay of non-Markovianity and nonlinearities. For example, the memory of an escape event over a potential barrier can cause a subsequent jump, inducing spontaneous (delay- and noise-induced) particle oscillations between neighboring potential wells. To capture such complex dynamics, one actually also needs higher-order correlations in time. In principle, these are described by the higher members of the Fokker-Planck hierarchy. The fact that explicit equations for the latter are, so far, not known hinders the development of further approximation schemes and solution methods. For this reason, we will also shed light on these equations and the higher-order PDFs. To this end, we will introduce a new derivation of the Fokker-Planck description by means of a *Markovian embedding technique* [97, 117–121] and a subsequent limiting procedure. The Markovian embedding is a generalization of the so-called linear chain trick [9, 122, 123], known in the context of deterministic delay equations. In this way, we find a closed, complete Fokker-Planck equation in an infinite-dimensional space, from which one can derive a hierarchy of equations. From a conceptual point of view, our approach yields an alternative perspective on the mathematical structure of delayed processes. Based on the second member, we will further propose an approximation for the two-time PDF, which is a central quantity in the description of delay systems, as it directly gives the correlation between the present and the delayed state. We will again test this approximation by application to the aforementioned bistable example model, this time focusing on the regime of delay-induced interwell-oscillations.

Thermodynamics

Another central framework in statistical physics is thermodynamics. In the last decades, special interest has been devoted to heat flow and other thermodynamic properties of small (mesoscale) systems, which are dominated by thermal fluctuations [124, 125]. *Stochastic thermodynamics* has emerged as an elegant and consistent framework to generalize thermodynamic notions to the level of noisy trajectories and to systems far from equilibrium [70, 126–132], with numerous applications to

soft matter [106], biological [131, 133], and quantum systems [134, 135]. For Markovian systems, concepts like entropy production, fluctuation theorems, stochastic energetics, and information exchange have been successfully formulated (and tested experimentally [63, 136, 137]). From a technical point of view, the FPE and path integrals (and their direct correspondence to the LE) turn out to be very useful. For example, the *total entropy production*, i.e., the *irreversibility* measure, can be directly calculated from path probabilities, or via the PDF and the associated probability current [125]. For non-Markovian systems, where probabilistic descriptions are much more involved, all these concepts have to be revisited. While extensions towards several non-Markovian systems have indeed been carried out in the past [138–147], the application to time-delayed systems with discrete time delay is still in its infancy. But before we dig further into technicalities, let us first think about relevant physical questions and expectations. As is commonly done in stochastic thermodynamics, we will focus again on the long-time behavior, where the system approaches a steady state.

While the incorporation of delay is a major challenge and it is not clear what we shall expect, the control aspect of our system is already well understood. Control is a deep-rooted topic in thermodynamics. In fact, major theoretical advances where initiated in the course of the long-lasting debate about the "Maxwell demon" [136, 148, 149]. This is an imaginary creature that can, by means of a sophisticated control, *extract energy from a single heat bath*, which means it "defeats the second law of thermodynamics". Later, simpler devices have been proposed, like the Szilárd engine [148, 150]. Interestingly, in the literature one even finds time-delayed, position-dependent feedback protocols suitable for this task, for example the feedback ratchet from Ref. [151]. However, there are two aspects that are fundamentally different from our feedback control. First, in these previous models the control is applied only at solitary instances in time. Between the control action, the system evolves "freely" and is on these time intervals, in fact, Markovian. Second, the control protocols typically *break the spatial symmetry* of the system externally. For example, in the Szilárd engine and the feedback ratchet, the control is "on" or "off" depending on whether the particle is left or right of a certain position. We wonder: Can the time-continuously operating feedback control considered in this thesis nevertheless also be used for the task of energy extraction from a single heat bath? In our setting, this would correspond to a reversed, *negative net heat flow*, which we shall thus look for!

Recent research has made clear that the search for a thermodynamic description of systems subject to continuous time-delayed feedback is indeed quite involved, even in the absence of nonlinearities [93, 99, 116, 152]. For the special case of a *linear* delay LE, it has been established that the delay force induces a net heat flow in the steady state [99] (implying nonequilibrium). However, there are three open problems concerning this earlier study. First, the derived heat flow is claimed to be positive definite, with an explanation that is based on an incomplete entropic treatment[5]. Second, the derived heat flow has an apparently unphysical behavior

[5] Information-theoretic terms in the entropy balance were forgotten and based thereon the heat flow is claimed to be positive definite due to the second law. We will come back to this later.

in the limit of vanishing delay time. Third, the approach taken in [99] is limited to linear systems only, which is a severe restriction. We want to clarify these issues, and extend the analysis towards nonlinear systems.

To this end, we will evaluate the heat rate in nonlinear (polynomial) potentials, and particularly focus on the behavior for small delay times. Our calculation explicitly yields a correction term to the results from [99], predicting that the heat rate is discontinuous at $\tau \to 0$, and drops to zero at $\tau = 0$ (where equilibrium is reached). We will discuss the origin of this behavior. We will further explicitly show that the direction of induced heat flow is indeed tunable upon adjusting the delay time and feedback strength (i.e., the laser intensity).

Thus, the "Maxwell demon type"-task can be performed by our feedback. Nowadays it is well-known that the underlying reason for the apparent "violation of the second law" in such controlled systems is that additional terms need to be added to the entropy balance, which account for the entropic cost associated with *information*. For example, deleting or storing information by the controller has entropic cost, as stated by Landauer's principle [148, 153]. Thus, in addition to the change of Shannon entropy (which automatically vanishes in steady states) and the medium entropy production (the heat flow divided by the temperature), information-theoretic terms add to the total entropy production. This is why one can have a negative steady-state heat flow, and at the same time a positive total entropy production. It would be very interesting to quantify these terms for the feedback control considered here. To this end, one should evaluate the total entropy production, which can be directly calculated by comparing the path probabilities of the process and its time-reversed image [125]. According to the framework and arguments by [93, 116], one here encounters a severe technical issue: the acausality of the backward process. Notably, the acausal nature of the backward process constitutes a crucial difference to the aforementioned protocols, where the control only acts at singular instances in time[6].

Considering all of this from a broader perspective, we actually run into fundamental problems of stochastic thermodynamics at this point: the notion of total entropy production of non-Markovian systems, the role of hidden degrees of freedom, the role of auxiliary variables, and coarse-graining. Recent research focuses on different strategies to approach these issues. One attempt is the formulation of "effective thermodynamic" descriptions[7] [93, 105, 107, 109, 116, 152, 154]. In this context, a generalized irreversibility measure for time-delayed systems was suggested in Ref. [93, 116], which fulfills nonequilibrium inequalities, generalizing the second law of thermodynamics and providing bounds to the extractable energy. However, this approach is not applicable for nonlinear systems, which is a major restriction. Furthermore, a general problem of effective thermodynamic descriptions is that the definition of such generalized irreversibility measures is a bit ambiguous.

Here, we will pursue a different strategy. In particular, we will include the controller in the model, via a specific representation of the feedback controller (via a ring network with unidirectional, linear couplings). This representation is closely

[6] Then one can chop the trajectory in a way that the control actions are at the boundaries only, creating a process whose backward image does not violate causality.

[7] We will later explain the idea in more detail, Chapter 3.

connected to the aforementioned Markovian embedding[8]. We will explore how much we can learn about the delayed process by studying the thermodynamic properties of the associated Markovian networks. In this way, we can for example access the additional information-theoretical entropy terms, and show that they are directly connected to information flows.

A characteristic feature of the Markovian networks is the involved unidirectional coupling, which is a specific type of a *non-reciprocal interaction*. As we have mentioned above, such non-reciprocal couplings are uncommon in statistical physics. Traditionally, one investigates systems consisting of elements with conservative interactions that are derivable from interaction Hamiltonians, and are thus reciprocal, like the atoms in a polymer, colloids in a dense suspension, or billiard balls. In the absence of driving forces or temperature gradients, such systems can equilibrate and are well described by traditional thermodynamics. However, it has already been realized in earlier literature that the idea of reciprocal couplings (and its thermodynamic implications) breaks down in many living and artificial active systems, where due to high complexity more general interactions between subsystems, or mesoscopic degrees of freedom, can emerge [110, 155]. This resembles the situation encountered here. Indeed, it appears somewhat natural that the feedback loop can be considered as a combination of non-reciprocally coupled entities, mirroring the intrinsic asymmetry of the elements' response to each other. Lately, researchers have started to investigate the consequences of non-reciprocal coupling on statistical properties of collective systems [110, 155]. However, the thermodynamic implications of the non-reciprocity have, to the best of our knowledge, not been considered in earlier literature. This is one focal point of this thesis.

Our investigation will reveal common features of delay dynamics and active matter. To rigorously analyze these connections, we will combine information-theoretical with energetic and entropic considerations. This yields an interesting perspective on the role of memory, activity and feedback for thermodynamic quantities.

Outline of the thesis

Here we lay out the structure of this thesis, which is visualized in Fig. 0.4.

Part I is an extended introduction to stochastic systems and provides the methodological background by reviewing important theoretical concepts, specifically, the Langevin equation (Chapter 1), the Fokker-Planck equation (Chapter 2), and stochastic thermodynamics (Chapter 3). At the end of each Chapter, we reconsider the framework with regard to time-delayed feedback, give an overview of the state of the art, and point out apparent open problems and ongoing discussions. These considerations provide the basis and motivate the detailed analysis presented in Parts II and III, which contain the main new results of this thesis.

In Part II, we thoroughly discuss the probabilistic description of systems with time delay. To begin with, we extend two earlier derivations towards higher orders (Chapter 4). We discuss linear delay equations separately, for which we explicitly derive higher-order probabilistic solutions (Sec. 4.2). Furthermore, we present

[8] which automatically relates time delay to a unidirectionally coupled Markovian network

an alternative derivation of the Fokker-Planck hierarchy (Chapter 5), based on a Markovian embedding (Sec. 5.1). The embedding makes explicit that the discrete delay can be regarded as the limiting case of distributed delay and further reveals connections between delay and non-reciprocal coupling. We introduce a new approximation scheme for the one-time probability density in Chapter 6 (called force-linearization closure), which exploits the solutions derived before for the linear case. We compare it to earlier approaches and discuss applications to two nonlinear example systems, a periodic and a bistable one. In Chapter 7, we present an approximation scheme aiming at the two-time joint probability density. Again, we apply this approach to the bistable example system, and compare it to the earlier approaches from Chapter 6.

Part III is dedicated to the thermodynamic description of stochastic processes with delay. First, we discuss the mean heat flow induced by the time delay (Chapter 8). Again using the bistable system as an example, we consider on both, the intrawell dynamics and the delay-induced interwell oscillations, and test how well the approximations from Part II can estimate the heat. We reveal connections to "entropy pumping", and investigate the implications of the overdamped limit. In Chapter 9, we turn to the concept of entropy, now considering a specific representation of the feedback controller, which involves non-reciprocal coupling (and is closely related to the Markovian embedding technique introduced in Part II). We investigate the (information-)thermodynamic implications of the encountered non-reciprocity, and discuss common features of delay dynamics and active matter.

Lastly, in Part IV we summarize our main results and provide an outlook for future research on stochastic systems with time delay.

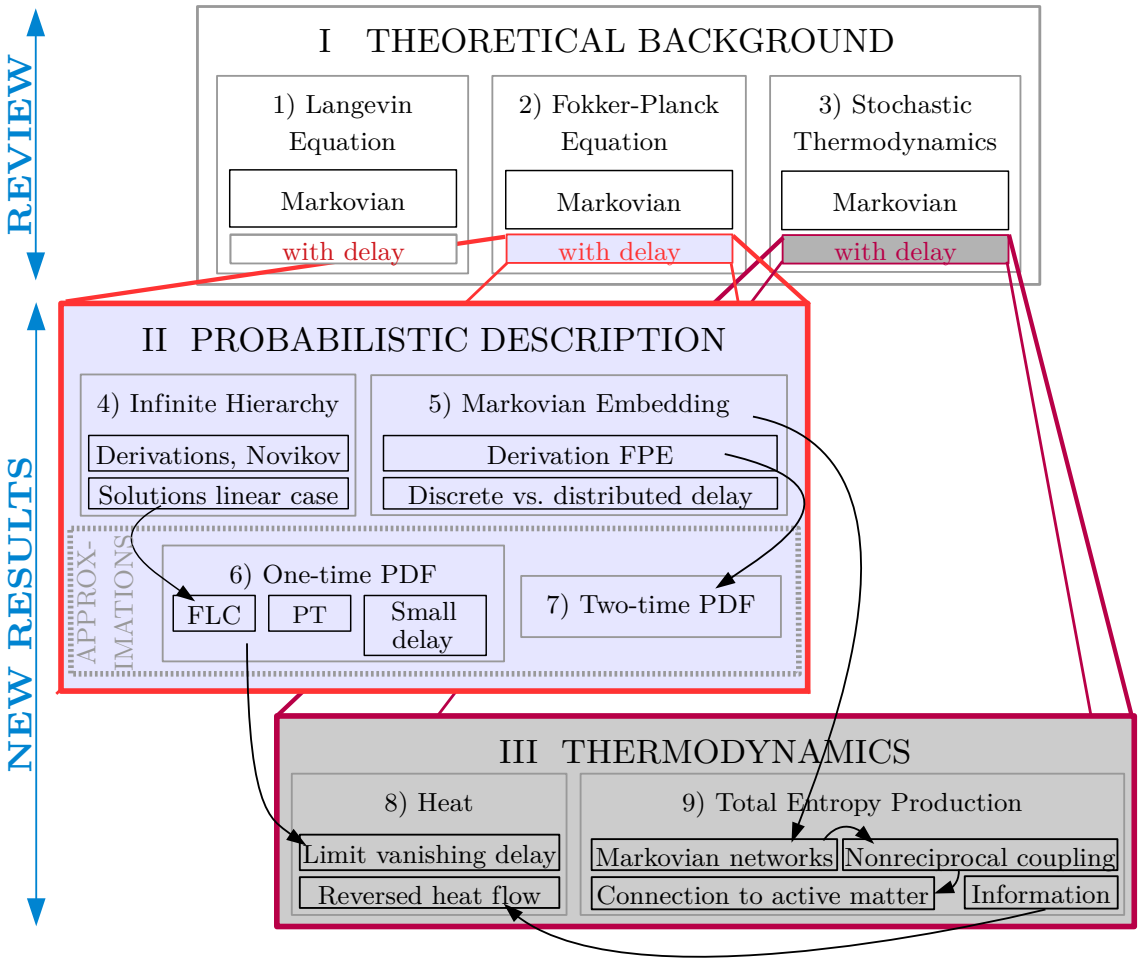

Figure 0.4: *Visualization of the structure of this thesis. The three large boxes represent the superordinate structure: Part I is a long introduction with review character, Parts II and III provide new results concerning probabilistic (II) and thermodynamic (III) descriptions of systems with time delay. The Arabic numbers in the inner boxes represent the respective Chapter numbers. A list of the meaning of the abbreviations can be found on page VIII.*

Part I

Theoretical background & State of the art

"FIRST REMARK. Non-Markov is the rule, Markov is the exception."

– Nicolaas G. van Kampen, 1998 [156]

In this Part I, we review some important theoretical concepts of the description of stochastic systems, in particular, the Langevin equation, the Fokker-Planck equation, and stochastic thermodynamics. We will always start with introducing established frameworks that are already subject of standard literature. As we will see, these frameworks are often limited to Markovian cases, while non-Markovian systems are by far less understood. At the end of each Chapter, we will reconsider the frameworks with regard to *time-delayed* dynamics. We will give an overview of the state-of-the-art research on stochastic delay equations and hint to ongoing debates.

Besides providing the methodological background, the purpose of these considerations is to make the Reader aware of apparent technical and conceptual problems in the description of non-Markovian, nonequilibrium or active systems, and to formulate relevant research questions. Another goal is to differentiate our findings from previous research. In this way, Part I builds up the basis for our new results presented mainly in Parts II and III of this thesis.

1 | The Langevin equation

"Mathematicians may flatter themselves that they possess new ideas which mere human language is as yet unable to express."

– James Clerk Maxwell, 1865 [157]

1.1 The stochastic way of describing things

1.1.1 Brownian motion

While you are reading this thesis, at every instance in time countless particles of the surrounding air hit your skin due to their irregular thermal motion. There are of the order of $\sim 10^{23}$ molecules in every liter of air [115, 158].

We cannot feel this bombardment, since the particles are tiny. But now imagine you would be much smaller, lets say, of the size of a single cell[1], yet larger than the molecules of your surrounding. Then, you would feel the bombardment very dramatically, in fact, these collisions would make you move randomly around yourself, even when you would try to stand still. You would perform a *Brownian motion*, like every small particle, called *Brownian particle*, suspended in a fluid at a finite temperature does [159]. Mathematically, this type of erratic, random motion is described by a *Langevin equation*.

Langevin equations describe systems subject to random perturbation. It is indeed typical that the system itself is "intermediate-sized", and studied on a *mesoscopic* length and timescale. This means, it is small enough to expose the collisions in the form of thermal fluctuations (by performing the random Brownian motion). This randomness would average out, if the system was much larger, or if it was considered on a *macroscopic* scale, i.e., a much bigger time- and length scale. Yet, the system is larger than the subsystems responsible for the perturbations, whose impact is described only implicitly by a single random term and a friction force. Typically, this is appropriate if multiple collisions take place at the same time, see

[1] like in the famous TV Series "Il était une fois...la vie" [*English title*: "Once upon a time...life"], where this effect is completely neglected, drawing a physically incorrect picture of the processes in our body.

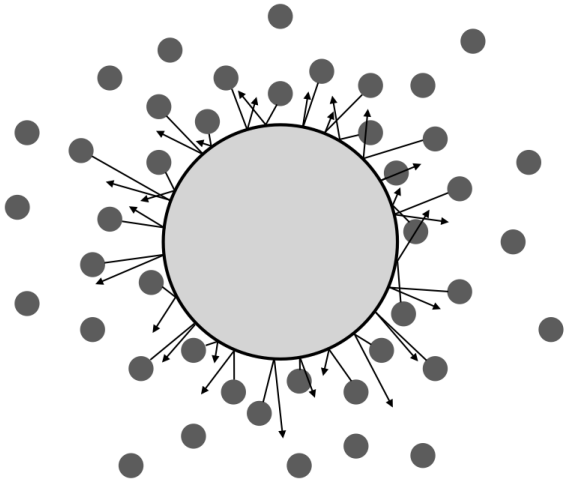

Figure 1.1: *Illustration of a colloid (large, light gray sphere) immersed in a fluid of much smaller particles (dark spheres). Multiple collisions take place at the same time, creating a Brownian motion of the colloidal particle. In the mesoscopic, i.e., coarse-grained ,description, this bombardment is modeled by a random force and a friction force.*

Fig. 1.1 for an illustration. In contrast, a *microscopic* description would mean that the positions and velocities of all particles are precisely described. To this end, one would need to solve a vast amount of equations[2] at every instance in time (about $\sim 10^{23}$, in the case of a colloidal particle immersed in one mole of fluid).

The stochastic way of modeling (for, e.g., colloidal suspensions) is the best description we have to date to describe complex phenomena on the nano- to micrometer scale. In the modern framework of statistical physics, stochastic modeling by the Langevin equation, and in particular Brownian motion, plays a central role. It provides a key step on the route from microscopic first principles to macroscopic laws, and enables a fully probabilistic formulation of statistical mechanics. Furthermore, it forms the backbone of a dynamical interpretation of statistical effects [159].

A key feature of Langevin equations is that they are *not* limited to systems in thermal equilibrium[3], which is of crucial importance in many applications. They can thus also be used to model various *active* particles, i.e., living or artificial matter with an internal energy source [102], like bacteria or robots (see Sec. 3.7). This will be important for our considerations, specifically the second Part of this thesis.

1.1.2 Colloidal suspensions

Colloids are a common example of Brownian particles, useful to experimentally investigate Brownian motion. The term colloidal particle denotes a few nanometer- to micrometer-sized, i.e., $\sim 10^{-9}$m -10^{-6}m, objects immersed in a fluid of atomically scaled particles, i.e., several Ångström, i.e., $\sim 10^{-10}$m. Colloids can be single macromolecules, lipid drops in water, artificial spheres, or pollen grains, as in the case of Robert Brown's name-giving, original observation of this motion [161]. Various further examples of colloidal suspensions can be found in [162]. From an experimental perspective, a pleasant property of colloidal particles is that they are comparably easy to access in experiments, as they are usually visible in optical microscopes. They can further be manufactured in a controlled manner, making it possible to

[2] An example for a numerical approach is multi-particle collision dynamics [160]. Analytically, however, there is no hope.

[3] Except for an equilibrium assumption for the bath, which we will explain in Sec. 1.3.1.

design colloids with desired properties (size, shape, charge) within a quite wide range [162]. Also the interactions between colloidal particles are often controllable, e.g., by external electrical fields [163, 164] or optical feedback [66]. Furthermore, experimental techniques enable precise dynamical control of the position of a colloidal particle, for example by optical traps (this tool is particularly relevant in the context of this thesis, and will be reviewed in Sec. 1.4.1). On the collective level, colloids are used to study friction, phase transitions, glassy dynamics, or critical phenomena [163]. Further, by tracking the movement of individual colloidal particles, one can study hydrodynamic effects, viscoelastic properties of the surrounding fluid, or diffusion properties. Colloidal suspensions can also be found in nature. For instance, macro-molecules, viruses, individual and small clusters of cells are on this very (colloidal) length scale, which is why the stochastic way of modeling introduced in the following, is particularly useful in (cell-)biology.

1.1.3 Side note: A more general view

Besides the example of Brownian motion, there are endlessly more processes in physics and other disciplines, where a detailed description at all length scales is impossible, because there are too many components, due to uncertainties in the model resulting from the fact that not all relevant processes are known precisely, or because of external random influences. Thus, it is no surprise that stochastic modeling is very widely spread, and Langevin, or Langevin-like, stochastic differential equations have proven themselves extremely useful to understand complex dynamical systems. Examples include gene expression [165], population growth dynamics [166], climate prediction [167], the stochastic motion of stock market prizes [168], the human pupil reflex [30], neural networks [169], just to name a few[4]. Other common examples are the fluctuating electric voltage in RC circuits [170, 171], and the concentration of a chemical component in a mixture [12, 172]. We will mostly consider colloidal systems, but many results presented in this thesis also apply to other stochastic processes.

For pedagogical and comprehensive introductions to stochastic systems in physics, we refer the interested Reader to the books by Gardiner [158] and Risken [173], and to [174] by Karatzsas and Shreve for a more mathematical perspective. A great book on stochastic thermodynamics, which also contains an educational introduction to stochastic modeling in physics, is the book by Sekimoto [124].

1.2 The Markovian Langevin equation

The motion of a Brownian particle is mathematically described by the Markovian *Langevin equation* (LE)

$$m\frac{\mathrm{d}^2X(t)}{\mathrm{d}t^2} = -\gamma\frac{\mathrm{d}X(t)}{\mathrm{d}t} + F + \sqrt{2k_\mathrm{B}\mathcal{T}\gamma}\,\xi(t), \tag{1.1}$$

[4]By the way, in all these examples, time delays are in fact also known to play a role, but let us first focus only on the stochastic part here and take it step by step.

which is a stochastic differential equation with the time $t \in [0, \infty[$, and $X \in \Omega$. $\Omega \subseteq \mathbb{R}$ denotes the spatial domain. X is a stochastic variable and represents the dynamical quantity of interest. Depending on the system at hand, this can be the position of a Brownian particle, the angle of a stochastically rotating system, the concentration of a chemical component, or any other *degree of freedom* (d.o.f.) whose stochastic dynamics is describable by (1.1). We will nevertheless call X *position* throughout this thesis for the sake of simplicity. The term *Markovian* refers to the fact that the LE (1.1) is *local* in time, which means that one only needs to know the system state at time t to infer the probabilities of its future evolution, while additional knowledge of past states does not provide further useful information for this task (we will give a more rigorous definition based on transition probabilities in Sec. 2.1.2).

The LE is a force balance, with friction and noise.Let us comment on all terms of (1.1) from left to right:

- $m\ddot{X}(t)$ is the inertia term with the mass of the particle m and the second derivative of the position, i.e., the particle acceleration. We will later consider the overdamped limit which amounts to neglecting this term. Note that we use both notations for time derivatives: $\dot{X}(t') = \frac{\mathrm{d}X(t)}{\mathrm{d}t}\big|_{t=t'}$.

- $-\gamma\dot{X}(t)$ is a Stokes friction force[5] that is proportional to the velocity with the friction coefficient γ. In Sec. 1.3.1, we will sketch a possible derivation of this term for a toy model with harmonic oscillators.

- F denotes the sum of all *external forces* acting on the particle. They may originate from a static potential $V(x)$, then $F(x) = -V'(x) = -\partial V(x)/\partial x$ is conservative, but they may also stem from other external sources. In particular, F can also be time-dependent and nonconservative, or be a combination of both, conservative and nonconservative parts. We will later introduce time-delayed (feedback) forces, making it *history*-dependent. We will denote the forces in F *deterministic*, to distinguish them from the random Langevin force (next bullet point).

- The last term is the Langevin force, with the noise strength $\sqrt{2k_\mathrm{B}\mathcal{T}\gamma}$, and the *noise* ξ. It involves the temperature $\mathcal{T}$ of the surrounding fluid, the Boltzmann constant k_B, and the friction constant γ. As the Langevin force considered here is not X-dependent, this noise is referred to as *additive noise* (as opposed to multiplicative noise "$g[X(t)]\xi(t)$" with some function g).

An overview over all coefficients and variables, including the respective SI units, is given at the beginning of this thesis on page IX. Further, a list of all abbreviations is given on page VIII.

In the following, we will further specify and discuss the individual components of the Langevin equation, introducing the suitable mathematical framework and

[5]The simple form of the friction is because we consider a fluid at low Reynolds numbers with laminar flow. According to Stokes law, the friction coefficient is related to the diameter of the colloid σ, and the viscosity of the fluid ν, via $\gamma = 3\pi\sigma\nu$.

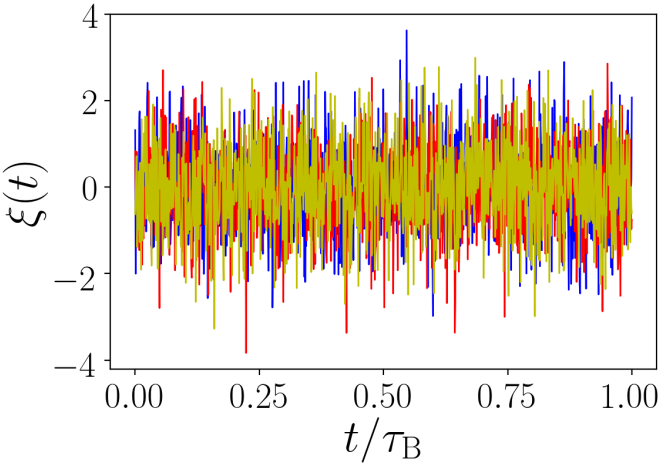

Figure 1.2: *Numerically generated white noise with a temporal discretization of* $\mathrm{d}t = 10^{-3}\tau_{\mathrm{B}}$*, with the (here) arbitrary timescale* τ_{B} *(we will later give physical meaning to* τ_{B}*). Three different random number seeds where used (indicated by different colors).*

provide physical explanations. Special attention needs to be drawn to the noise term. In particular, we need to define the noise and the stochastic calculus (i.e., Itô, Stratonovich) below, to give (1.1) mathematical meaning. While we will use a notation indicating one-dimensionality of the considered system, many of the definitions and considerations can indeed readily be applied to multidimensional systems as well, for example by simply replacing X by a vector $\underline{X} = \{X_0, X_1, ..., X_n\}$, and the noise ξ by a vector $\underline{\xi} = \{\xi_0, \xi_1, ..., \xi_n\}$. We will come back to this in later parts of this thesis (e.g., Chapters 5 and 9).

1.2.1 Gaussian white noise

The noise ξ is assumed to be Gaussian distributed, hence, it is fully determined by the first two moments. Further, we will assume *white noise*[6], that means

$$\langle \xi(t) \rangle = 0 \tag{1.2}$$

$$\langle \xi(t)\xi(t') \rangle = \delta(t - t'), \tag{1.3}$$

involving the Dirac delta distribution, and the *ensemble average* $\langle ... \rangle$ (which will be defined in Sec. 1.2.2). The temporal correlations $\langle \xi(t)\xi(t') \rangle$ decay infinitely fast. This amounts to the assumption that the random collisions of the colloid and bath particles are so numerous and fast that they are independent from each other at different instances in time. In contrast, noise terms with account for *finite* correlations in time (not a delta-distributed correlation) are denoted *colored noise*. We will consider a particular case of colored noise in Chapters 5 and 9. The white noise (1.2) has infinitely many "jumps" on each time interval, no matter how short it is. Indeed, this is physically not reasonable, but yields a good approximation for many situations, and is much easier to deal with from a technical point of view, as opposed to correlated, i.e., colored noise. Figure 1.2 shows exemplary, numerical realizations of the white noise process.

[6] "White" refers to the power spectrum of the noise, which is constant, resembling the spectrum of white light. The power spectrum is given by the Fourier-transformation of the correlation function according to the Wiener-Khintchine theorem [173].

The *path probability* to observe a specific realization $\boldsymbol{\xi} = \{\xi(t)\}_{t_1}^{t_2}$, of the white noise on a time interval $t \in [t_s, t_f]$ is given by the path integral [175] (also see Sec. 1.2.6 below)

$$\mathcal{P}[\boldsymbol{\xi}] \propto e^{-\int_{t_s}^{t_f} \frac{\xi(t')^2}{2} \mathrm{d}t'}. \tag{1.4}$$

A mathematically rigorous treatment can be found in [175].

We will come back to the mathematical description of white noise below, when we introduce the Wiener process and the stochastic calculus.

Gaussianity – Central limit theorem Besides making the mathematical treatment much easier, the assumption of *Gaussianity* of the noise term is, in various cases, indeed very reasonable from a physical point of view. The underlying reason is the *central limit theorem*. It states, loosely speaking, that the sum of a great number of independent random events[7] will asymptotically approach a Gaussian distribution [115, 158, 176]. Thus, Gaussian noise in a sense "automatically comes up" whenever random processes are modeled which are a result of many (approximately) uncorrelated, random processes. This is for example the case for all the collisions between the Brownian particle and the surrounding bath particles, and among the latter.

The heat bath and the temperature The friction force and the Langevin force both stem from the interaction between the Brownian particle and the surrounding fluid particles. The collection of these fluid particles is called the *reservoir*, the *heat bath*, or simply the *bath*. Friction slows the particle down, while the Langevin force prevents the particle from standing still, creating the erratic motion.

As the surrounding particles are much smaller and thus lighter than the Brownian particle, the timescale of their motion is also much shorter (lighter particles typically move much faster[8]). Thus, both timescales are well-separated, making it plausible to assume that the bath particles can exchange their momenta and distribute their positions "infinitely fast" from the perspective of the "slow" Brownian motion one is mainly interested in. Further, the bath is assumed to be *infinitely large* and at a *constant temperature* $\mathcal{T}$. This amounts to the important assumption that *the bath is in thermal equilibrium at any time*. The temperature is connected to the mean quadratic velocity (in fact, their kinetic energy) of the bath particles via the *equipartition theorem*. This theorem states that in equilibrium, the energy is (on average) *equally* distributed among the degrees of freedom [158], with

$$\frac{1}{2} m_{\mathrm{bath}} \langle v^2 \rangle = \frac{d}{2} k_{\mathrm{B}} \mathcal{T}, \tag{1.5}$$

for d-dimensional systems, with m_{bath} being the mass of the individual bath particles. The associated *thermal energy* of a particle in the bath is $k_{\mathrm{B}}\mathcal{T}$. One should stress

[7] under fairly general conditions on their individual statistics, as long as they have a finite variance

[8] The reason is the equipartition theorem, see below.

that the temperature is defined on the basis of the bath only, not taking into account the state of the Brownian particle (which does not have to be in equilibrium, and can thus also move faster than the equipartition theorem would state). In Sec. 1.3.1, we will sketch how friction and noise emerge from interactions with the bath particles, and how the temperature is introduced in the stochastic description. The *infinite size of the bath*, which corresponds to a thermodynamic limit, makes it plausible that the temperature can stay constant, even if external driving induces a steady net heat flow from the particle to the bath, which we will discuss in Chapter 8.

1.2.2 Ensemble averages & probability density

An ensemble is an imaginary (infinite) set of identically prepared systems (with equal temperature, same initial conditions,...), each yielding one possible realization of the random process. The ensemble average [e.g., in Eq. (1.2)] is the average over multiple, in principle all possible, realizations of the random process, i.e.,

$$\langle A(X)\rangle = \frac{1}{N}\sum_{i=1}^{N} A(X_i) \tag{1.6}$$

for an arbitrary function A of the random system state X. Here X_i denotes an individual realization, and N is the total number of all realizations ($N \to \infty$). For *ergodic* processes, the ensemble average is equivalent to the time-average of the process (over infinitely long trajectories) [158]. All processes considered in this thesis are assumed to be ergodic. As opposed to the random process X, ensemble averages are *deterministic* quantities (i.e., noise-free), and hence, contain "reliable" information about the process. In some contexts, they are hence of main interest.

Probability density A particularly important ensemble-averaged quantity is the *probability density function* (PDF)

$$\rho_1(x,t) = \langle \delta[X(t) - x]\rangle, \tag{1.7}$$

with $x \in \Omega$ (note the distinction between X which is the stochastic system state, i.e., the colloidal position, and x which is the spatial variable). The probability of finding the particle at time t within a given interval $[x, x+\Delta x]$ is $\int_x^{x+\Delta x} \rho_1(x',t)\mathrm{d}x'$. As the overall probability to find the particle somewhere in the spatial domain Ω has to be one, the normalization is $\int_\Omega \rho_1(x',t)\mathrm{d}x' = 1$. The subscript emphasizes that this is a PDF for a single stochastic event $\{x,t\}$, as opposed to joint PDFs which will be consider in Sec. 2.2. Finding an analytical expression for the PDFs of systems with time delay will be the main focus of Part II of this thesis. Like (1.6), the PDF can as well be used to calculate ensemble averages of a function $A(X)$, through

$$\langle A(X)\rangle = \int_\Omega A(x)\rho_1(x,t)\mathrm{d}x. \tag{1.8}$$

Steady states In the *long time limit* many systems approach a stationary PDF

$$\text{Steady states:} \quad \partial_t \rho_{1,\text{ss}}(x,t) = 0, \tag{1.9}$$

which is the defining criterion of *steady states*. The latter can be distinguished as thermal *equilibria* (EQ), or *non-equilibrium steady states* (NESS), which we will precisely define in Chapter 3. Due to their time-independence, steady states generally allow for more analytical treatment than transients. They play a central role in statistical physics, in general, and stochastic thermodynamics, in particular (see Chapter 3), comparable with the central role of EQ in traditional macroscopic thermodynamics. Hence, they are a major focus in our investigations.

Throughout the thesis, we will occasionally use the subscript "ss" to emphasize that an ensemble average or PDF refers to the steady state (however, we will not always indicate steady states in that way).

Correlation functions Another common type of ensemble-averaged quantities are *autocorrelation functions*, for example the *(temporal) position-autocorrelation function*

$$C(\Delta t) = \langle X(t)X(t+\Delta t)\rangle_{\text{ss}}. \tag{1.10}$$

As already indicated by the argument here, in steady states the correlation functions only depend on the time-difference Δt, which is the typical scenario that we will consider in this thesis. When considering correlation functions of other quantities than X, we will denote them with corresponding subscript, for instance, $C_\nu = \langle \nu(t)\nu(t+\Delta t)\rangle$ is the autocorrelation of ν. Correlation functions contain a lot of information about a (random) process, and, in particular, tell how different instances in time are connected with each other. Thus, they are on the one hand important to describe dynamical behavior, and on the other hand, give us insight into the memory of a process.

In fact, for processes with Gaussian-distributed probability densities, the complete probabilistic information is encoded in the position-autocorrelation function, because Gaussian-distributed densities are fully determined by their first two moments. As we will later see, the steady-state solutions of *linear* LEs are typically Gaussian PDFs.

1.2.3 Solutions of the Langevin equation and the overdamped limit

Ballistic timescale Let us consider (1.1) in the absence of external forces, i.e.,

$$m\frac{\mathrm{d}^2X(t)}{\mathrm{d}t^2} = -\gamma\frac{\mathrm{d}X(t)}{\mathrm{d}t} + \sqrt{2k_\mathrm{B}\mathcal{T}\gamma}\,\xi(t), \tag{1.11}$$

describing a "free" particle. Setting the velocity v_0 as initial condition, i.e., $\frac{\mathrm{d}X(t)}{\mathrm{d}t}\big|_{t=0} = v_0$, this equation has the formal solution

$$\dot{X}(t) = v_0 e^{-(\gamma/m)t} + \int_0^t e^{-(\gamma/m)|t-t'|}\sqrt{2k_\mathrm{B}\mathcal{T}\gamma/m}\,\xi(t')\mathrm{d}t'. \tag{1.12}$$

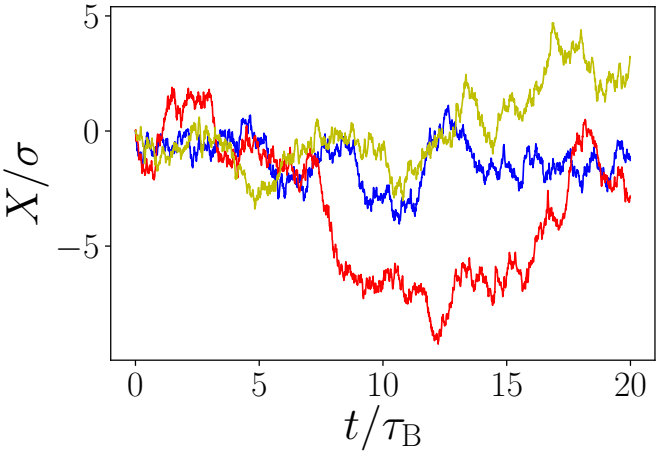

Figure 1.3: *Three example trajectories of a free Brownian particle in the overdamped limit with $X(0) = 0$. The data stem from Brownian dynamics simulations (see Appendix A.1 for details). At the same time, these trajectories represent a Wiener process (1.25) with $W(t)\hat{=}X(t)$, see below.*

The term "formal" refers to the fact that the solution still contains the random force, owing to the stochastic character of the Langevin equation. Such formal solutions are very valuable for analytical calculations, and are hence important for the remainder of this thesis. We stress that for *nonlinear* LEs, formal solutions are not always available (see, e.g., Ref. [177] for a discussion of exact solutions of nonlinear differential equations), making their treatment much more involved.

Equation (1.12) shows that the initially set velocity v_0 exponentially decays, as the velocity relaxes on a timescale m/γ, called the *ballistic* or *velocity-relaxation timescale*. This is typically the shortest characteristic timescale of the process.

Using the properties of white noise (1.2), one finds from (1.12) the velocity-auto correlation function

$$\left\langle \dot{X}(t_1)\dot{X}(t_2)\right\rangle = v_0^2\, e^{-(\gamma/m)(t_1+t_2)} + \frac{k_{\mathrm{B}}\mathcal{T}}{m}\left[e^{-(\gamma/m)|t_1-t_2|} - e^{-(\gamma/m)|t_1+t_2|}\right], \tag{1.13}$$

which in the long time limit $t_1 \gg 1$, $t_2 \gg 1$ relaxes to the stationary value

$$\left\langle \dot{X}(t_1)\dot{X}(t_2)\right\rangle_{\mathrm{ss}} = \frac{k_{\mathrm{B}}\mathcal{T}}{m}\, e^{-(\gamma/m)|t_1-t_2|}, \tag{1.14}$$

(see [173] for further details). Thus, the correlations decay exponentially on the ballistic timescale.

The value of the stationary velocity-autocorrelation at $t_1 = t_2$ further implies that the mean kinetic energy of the free Brownian particle in the long-time limit is

$$\langle E_{\mathrm{kin}}\rangle = \frac{m}{2}\left\langle \dot{X}^2\right\rangle_{\mathrm{ss}} = \frac{k_{\mathrm{B}}\mathcal{T}}{2}. \tag{1.15}$$

Thus, as one should physically expect, the free Brownian particle approaches thermal equilibrium, where its mean velocity is adjusted by the bath's temperature (and the equipartition theorem holds).

Overdamped Langevin equation For many situations, the velocity relaxation (which takes place on the ballistic timescale m/γ), is very fast as compared to

timescales one is mainly interested in, such as the diffusive timescale (which we will discuss below). Thus, one can often assume the velocity relaxation to be infinitely fast. This gives rise to the *overdamped* LE

$$\gamma \frac{\mathrm{d}X(t)}{\mathrm{d}t} = F + \sqrt{2k_\mathrm{B}\mathcal{T}\gamma}\,\xi(t), \tag{1.16}$$

representing the overdamped limit, i.e., $m/\gamma \to 0$, of (1.1). We will mainly consider the overdamped LE in this thesis. We will refer to (1.1) as the *underdamped* LE.

Diffusive timescale Now, we again consider a free Brownian particle, i.e., at $F = 0$, this time described by the overdamped equation (1.16). Figure 1.3 displays some example trajectories. Starting at position $X(0)$, the formal solution reads

$$X(t) = X(0) + \int_0^t \sqrt{2k_\mathrm{B}\mathcal{T}/\gamma}\,\xi(t')\mathrm{d}t'. \tag{1.17}$$

It implies a linear increase in time of the so-called *mean square displacement* (MSD) $\langle [X(t) - X(0)]^2 \rangle$, specifically,

$$\langle [X(t) - X(0)]^2 \rangle = \frac{2k_\mathrm{B}\mathcal{T}}{\gamma} \int_0^t \int_0^t \delta(t' - t'')\mathrm{d}t'\mathrm{d}t'' = 2\frac{k_\mathrm{B}\mathcal{T}}{\gamma}\, t = 2D_0 t. \tag{1.18}$$

In the last step, we have brought (1.18) in a more common form by introducing the *diffusion coefficient* D_0 and by employing the *Einstein relation* [178]

$$k_\mathrm{B}\mathcal{T} = \gamma D_0. \tag{1.19}$$

This renowned relation connects the diffusion coefficient with the temperature and represents a *fluctuation-dissipation relation*. It further implies another way of expressing the thermal energy $\gamma D_0 = k_\mathrm{B}\mathcal{T}$. The MSD is a measure for the diffusive spreading of the Brownian particle due to the noise. Its linear increase is characteristic for Brownian motion. Importantly, the underdamped equation has the *same* behavior for times larger than the ballistic timescale[9].

Considering a typical length scale σ of the system, the average time a free particle would need to explore, i.e., travel, this distance by diffusion is according to (1.18) given by[10]

$$\tau_\mathrm{B} = \sigma^2/D_0, \tag{1.20}$$

which is called the *Brownian time*. This is a *diffusive timescale*. From now on, we will measure time in units of τ_B throughout this thesis. For a free particle σ could be, for example, the diameter of the colloid. Later, σ will be a typical length scale of an external potential V_s in our example systems.

In the last decades, there was a lot of interest in the search for models that produce (for long times) nonlinearly growing mean square displacements, in particular,

[9] If one does not neglect inertia, the MSD increases approximately quadratically $\sim t^2$ on very short times up to m/γ, and then passes over to the diffusive regime with $2D_0 t$.

[10] In other words, according to (1.18), the MSD equals $2\sigma^2$ at time τ_B.

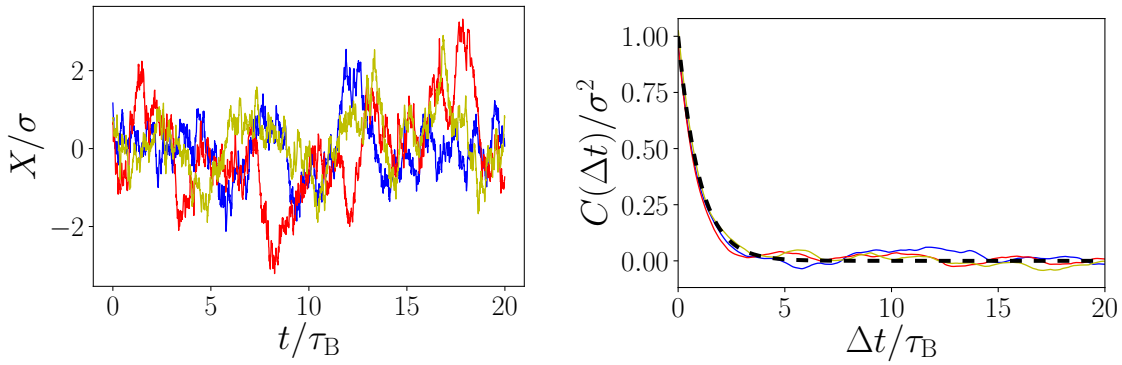

Figure 1.4: *Left panel: Three example trajectories of an overdamped Brownian particle in a harmonic trap, whose position X is performing an Ornstein-Uhlenbeck process. Right panel: Corresponding position-autocorrelation function, calculated from the three single trajectories (in the respective colors), and from the analytical results (1.22) shown by a dashed black line. We have cut off about 10^5 time steps to ensure steady state conditions. The only free parameter a is set to $a = \gamma/\tau_B$. It may seem strange to calculate the correlation function from a single trajectory. However, the system is ergodic, i.e., in the steady state the ensemble average is equivalent to the time-average of the process.*

$\langle [X(t) - X(0)]^2 \rangle = t^\alpha$ with $\alpha \neq 1$, called *anomalous diffusion* processes [13, 179]. (Specifically, $\alpha < 1$ denotes sub-diffusion, $\alpha > 1$ super-diffusion.) To find examples, we refer the interested Reader to [180] and references therein. Later, we will also briefly consider the long-time diffusion behavior of a system with delay from this perspective, see Sec. 6.5.5.

1.2.4 Ornstein-Uhlenbeck process

Another prominent example system is an overdamped particle in a quadratic or "harmonic" potential $(a/2)x^2$. This potential confines the particle's motion around its minimum leaving no way to escape, and is hence sometimes denoted a "*trap*". The motion of the particle is described by a fully linear LE like the one of a free particle, but this time with an additional linear restoring force $F = -aX$ in (1.16). Then, the LE generates a so called *Ornstein-Uhlenbeck process* (OU) in the position. We have in fact already discussed another OU above, that is Eq. (1.11), defining an OU in the velocity. Accordingly, the solution of the overdamped particle in a trap is analogously to (1.12) given by

$$X(t) = x_0 e^{-(a/\gamma)t} + \sqrt{2k_B\mathcal{T}/\gamma} \int_0^t e^{-(a/\gamma)|t-t'|} \xi(t') \mathrm{d}t'. \tag{1.21}$$

In the long-time limit, this system approaches a stable steady state (if $a > 0$), with a Gaussian-distributed PDF with zero mean. The Gaussianity, which is "preserved" by the *linear* LE, stems from the noise. The position-autocorrelation function reads

$$C(|t_2 - t_1|) = \langle X(t_1)X(t_2) \rangle_{ss} = \frac{k_B\mathcal{T}}{a\gamma} \left[e^{-(a/\gamma)|t_1 - t_2|} \right] = \frac{D_0}{a} \left[e^{-(a/\gamma)|t_1 - t_2|} \right]. \tag{1.22}$$

In contrast to the free particle, the MSD reaches a constant value D_0/a, due to the confinement by the quadratic potential. For example, when starting at $X(0) = 0$, the MSD reads

$$\begin{aligned}\langle [X(t) - X(0)]^2 \rangle &= \langle X(t)^2 \rangle - 2\langle X(t)X(0) \rangle - \langle X(0)^2 \rangle = C(0) - 2C(t) \\ &\overset{(1.22)}{=} \frac{D_0}{a} \left[e^{-(a/\gamma)} - 2e^{-(a/\gamma)t} \right] \end{aligned} \tag{1.23}$$

Thus, the MSD saturates at the timescale where the autocorrelation function decays. As can be seen from Eq. (1.22, 1.23), the corresponding relaxation time is of the order of γ/a.

Figure 1.4 displays some exemplary trajectories, as well as the corresponding autocorrelation function (1.22) of a Brownian particle in an harmonic trap in the steady state.

1.2.5 White noise – Wiener process – Stochastic calculus

Although very common in physics literature, the treatment of stochastic equations with white noise (1.2) is less popular among mathematicians, who typically write down the LE in the form

$$\gamma \mathrm{d}X(t) = F\mathrm{d}t + \sqrt{2k_\mathrm{B}\mathcal{T}\gamma}\,\mathrm{d}W(t), \tag{1.24}$$

with the *Wiener process* $W(t)$. The latter is a random process defined by independent, Gaussian-distributed increments $\mathrm{d}W(t) = W(t + \mathrm{d}t) - W(t)$, with

$$\begin{aligned} W(0) &= 0, \\ \langle W(t + \mathrm{d}t) - W(t) \rangle &= 0, \\ \langle W(t)W(t + \mathrm{d}t) \rangle &= t, \ \forall\, \mathrm{d}t > 0. \end{aligned} \tag{1.25}$$

As opposed to ξ, the Wiener process $W(t)$ is *continuous* in time. It is not differentiable.

The reason why mathematicians rather favor the description with the Wiener process is self-evident. While $\mathrm{d}X$ in (1.24) is well-defined, the velocity $\dot{X}$ in the LE (1.16) is not well-defined and its absolute value is almost surely infinity at each instance in time [158]. We will nevertheless consider white noise equations in this thesis, consistent with physics literature. After all, both representations have to be treated with equal care, and require the definition of the stochastic calculus.

Itô vs. Stratonovich The LEs in the form (1.24), (1.16) [or (1.1) for underdamped motion] are in fact not sufficient to fully characterize the noisy dynamics, but one additionally has to specify the *stochastic calculus*. This is not only true for multiplicative noise, but also for additive noise (despite occasional misconception in the literature), as we explain in the following. The stochastic calculus decides on the question of how to interpret integrals of the type $\int_0^t X(t)\mathrm{d}W(t)$, and roots in the issue of the zero–infinite nature of the *white noise*, i.e., its delta-autocorrelation.

Two common choices are Itô, i.e., the starting-point rule, and Stratonovich, i.e., the mid-point rule, which amount to the integration of some arbitrary functional $\mathcal{A}$

$$\int_0^T \mathcal{A}[X(t)]\ \mathrm{d}W(t) = \begin{cases} \sum_{i=0}^{k-1} \mathcal{A}\left[\frac{X(t_{i+1})+X(t_i)}{2}, \frac{t_{i+1}+t_i}{2}\right]\ \left[W(t_{i+1}) - W(t_i)\right], & \text{Stratonovich} \\ \sum_{i=0}^{k-1} A\left[X(t_i), t_i\right]\ \left[W(t_{i+1}) - W(t_i)\right], & \text{Itô} \end{cases} . \quad (1.26)$$

See [173, 181] for further details. More generally, both are special types of α-type calculus[11], where Itôs corresponds to $\alpha = 0$ and Stratonovich to $\alpha = 0.5$. Since they are very uncommon in physics literature, we do not discuss other α-type calculuses here.

As one can see from (1.26), only in the Itô calculus X is *non-anticipating* (or an adapted process), as it does not depend on future X. An important difference is that Stratonovich preserves the normal rules of calculus known from deterministic equations, like the chain rule, while Itô yields different rules. Another distinction is given by the connection between LE and PDF. When Itô calculus is applied, the associated Fokker-Planck equation (see Sec. 2.2) has an additional term, the *spurious* or *noise-induced drift*, as opposed to the corresponding Stratonovich FPE, see e.g., [173, 181]. This additional drift term, which is only present in Itô, yields a different solution, i.e, a *different PDF*.

However, the noise-induced drift vanishes for additive noise (as it depends on the spatial derivative of the noise strength in the Langevin-force, which is zero for multiplicative noise), thus, both calculus yield the exact same PDF, making the distinction on that level impossible. For this reason, some authors argue that the distinction of the calculus is not necessary at all in this case. This is indeed not true. For additive noise, *other* quantities may still differ, for example the cross-correlations $\langle X(t)W(t')\rangle$. For example, Itô calculus implies $\langle X(t)W(t)\rangle = 0$, whereas in Stratonovich $\langle X(t)W(t)\rangle = t/2$, [for initial condition $X(0) = 0$]. This underlines that the one-time PDF, after all, does not contain the full information about the noisy process.

However, one should emphasize, that *both calculuses can be used in any case, and the corresponding equations can be transformed into each other in a unique way* by means of Itô's formula [158]. This is for example very useful for path integrals, which are easier to handle in Itô calculus (where the Jacobian of the transformation from noise to space coordinates, i.e, $\mathcal{J}$ in (1.27), is unity, see [182]). Consistency is important, whenever using a formula one needs to make sure which calculus it was defined in (an example being the stochastic heat, see Chapter 3).

In the 1970's and 1980's there was a debate about the interpretation (and "correct" calculus in physical systems), see, e.g., [183] for a historical review. The main outcomes of this debate may roughly be summarized as follows: Stratonovich typically shall be used to model physical system which are time-continuous. On the contrary, intrinsically discontinuous systems shall be treated with Itô's rule (e.g., financial markets, population growth dynamics). While there seems to be consent

[11] Note that this α has nothing to do with the exponent defining anomalous diffusion.

on this point, there is in fact still an ongoing controversy about the more appropriate calculus in physical systems [183]. We aim to add to this debate the following argument: when one thinks of a physical system on the molecular scale, the question of continuity vs. discreteness may actually be fuzzy (the number of Adenosine triphosphate, short ATP molecules, fueling a process is discrete, while the fueled molecular motor [184] itself is moving through continuous space). Due to improved experimental techniques there is even renewed interest in the question of what the "correct" calculus is and a recent debate see, e.g., [185–188], (also arguing that other calculus than the α-type might be more appropriate [189]). However, as pointed out in [183], some of the arguments and questions should be viewed critically.

We will employ *Stratonovich calculus* throughout the thesis, because this is, in general, the more common choice in physics literature, and, in particular, consistent with the formulation of the stochastic energetics formalism employed in Part III, see Chapter 3.

1.2.6 Path integral representation

Above, we have shown how the probability density function $\rho_1(x,t)$ can be defined on the basis of the stochastic process X. It gives the probability to observe a certain value at a given time t, see Eq. (1.7). Besides the PDF, another way of expressing the process's probabilities are *path probabilities*, which tell how likely it is to observe an entire trajectory. In particular, one defines *paths* as $\mathbf{X} = \{X(t)\}_{t_s}^{t_f}$, i.e., $X(t)$ for all $t \in [t_s, t_f]$ giving one specific realization of the random process. For a Markovian process with white noise (1.4), the path probability $\mathcal{P}$ conditioned on the starting point $X(t_s) = x_s$, is given by the exponential of the *Onsager-Machlup action* [117, 125]. The latter readily follows from the Markovian LE, yielding (for the overdamped case)

$$\mathcal{P}[\mathbf{X}|x_s] = \quad \mathcal{J}\, e^{-\int_{t_s}^{t_f} \left(\dot{X}(t') - \gamma^{-1} F[X(t')]\right)^2 / (4k_\mathrm{B}\mathcal{T}\gamma)\mathrm{d}t'}. \tag{1.27}$$

The $\mathcal{J}$ factor is a Jacobi determinant stemming from a coordinate transformation from noise to spatial coordinates, $\xi(t_j) \to X(t_j)$, see Ref. [182] for details. In Itô calculus $\mathcal{J}$ is unity, in Stratonovich not. As we will mostly be interested in path probability ratios in this thesis, these factors anyway cancel out. We will thus not further discuss them.

In the framework of stochastic thermodynamics, which we will introduce in Chapter 3, the path probabilities play a central role, in particular to measure irreversibility. More generally, they can be used to calculate the ensemble average of a given functional $\mathcal{A}$ along an entire trajectory, via

$$\langle \mathcal{A}[\mathbf{X}] \rangle = \int_{t_\mathrm{s}}^{t_\mathrm{f}} \mathcal{A}[\mathbf{X}] \mathcal{P}[\mathbf{X}]\, \mathfrak{D}\mathbf{X} \tag{1.28}$$

with $\mathcal{P}[\mathbf{X}] = \mathcal{P}[\mathbf{X}|x_s]\rho_1(x_s, t_s)$.

The concept of path probabilities can be generalized to non-Markovian processes. For example, Ref. [190] discusses path integrals for processes with colored

noise, and Ref. [93, 116] employs path probabilities for system with time delay. A crucial issue in stochastic thermodynamics of non-Markovian processes is the path probability of the *backward* process. We will come back to this issue in Chapter 3.

In the following, we will make a little detour, exploring the aforementioned connection between mechanical laws and stochastic dynamics in more detail. We will proceed with the properties of the process generated by the LE in Sec. 1.4, where we will finally introduce the time delay.

1.3 How stochastic motion emerges from first principles

The introduction of time-reversal asymmetry

1.3.1 Infinite harmonic oscillators bath – An example of a Mori-Zwanzig projection

There is an established and well-understood link between the Langevin equation (1.1, 1.16) and *first principles*, via Mori-Zwanzig projection method. To be more specific, in the picture of the Mori-Zwanzig derivation, the "microscopic world" consists of point particles governed by Newton's laws (i.e., fulfill *actio = reactio*), obeying momentum and energy conservation. These laws are considered the first principles here[12]. We are not going into details but just give a rough overview over this derivation. Details can be found, e.g, in Ref. [12, 94, 124, 193].

One starts with Hamiltonian dynamics for *all* particles with various, presumably identical, "bath particles", say $Y_{j\in\mathbb{N}}$, and one special particle, say X. In our consideration this represents the colloid. One particularly simple possibility is to consider for the Y_j *harmonic oscillators* with frequencies ω_i, which are all *linearly* coupled to X, but not among each other (see, e.g., [12] p. 22). X is reciprocally (i.e, the interaction fulfills *actio = reactio*) coupled back to each harmonic oscillator, and may further "feel" external forces F. This situation is described by the following coupled equations

$$\begin{aligned} m\ddot{X} &= F + \sum_j (b_j Y_j - (b_j/\omega_j)^2 X) \\ \ddot{Y}_j &= -\omega_j^2 Y_j + b_j X. \end{aligned} \tag{1.29}$$

These are the *microscopic* equations for the present case. Due to the *linearity* of the Y_j-equations, this set of equations can be handled quite well analytically.

In particular, one can *project* all equations onto the dynamical equation of the colloid, by formally solving and substituting them into each other. To this end, we perform the Laplace-transformation $\hat{Y}(s) = \int_0^\infty Y(t)e^{-st}\mathrm{d}t$ of the harmonic oscillator

[12] as well as in kinetic gas theory [191, 192]

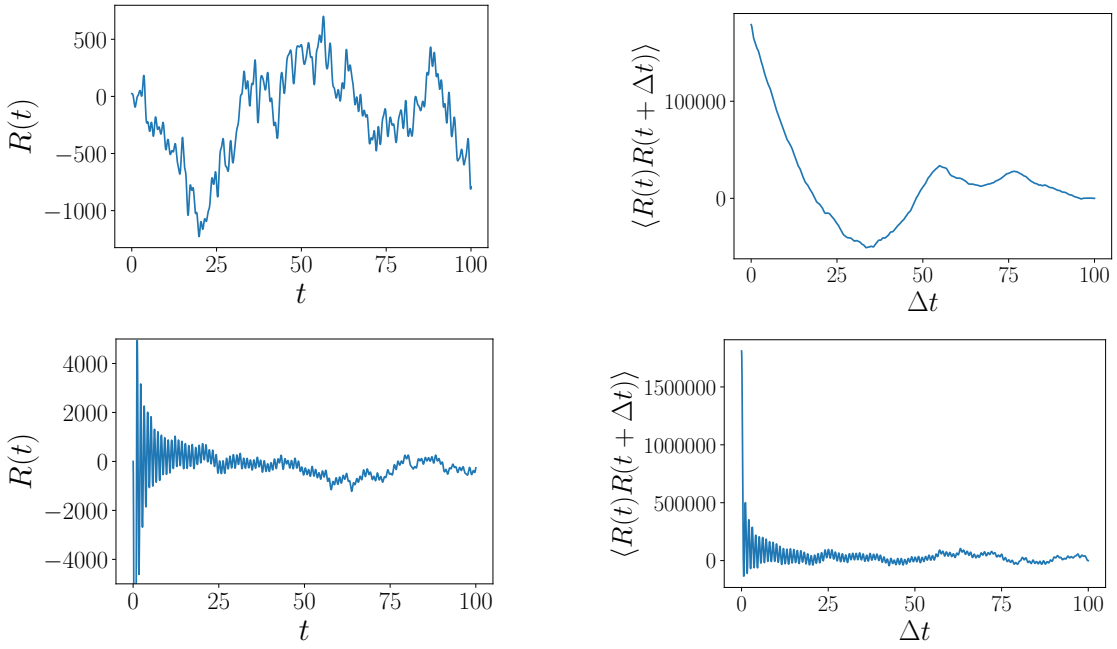

Figure 1.5: *Left panel: R from Eq. (1.34) and right panel: its autocorrelation function, for a system with $N = 10^5$ harmonic bath oscillators with equidistantly distributed frequencies ω_j among the oscillators within $[0, 2\pi]$, i. e,. $\omega_j = j\,2\pi/N$ with $j = \{1, ..., N\}$. The initial velocities $\dot{Y}_j(0)$ are randomly drawn from a Gaussian distribution (with variance 1), while the initial positions differ between upper and lower plots. Upper panels: Gaussian-distributed random initial positions $Y_j(0)$, with variance 0.27. Lower panels: Equidistantly distributed positions $Y_j(0)$ within $[-0.2, 0.2]$. The parameters are in both cases chosen such that the equipartition theorem is fulfilled (see text). Other parameters: $X(0) = 0$, $b_j = 1\ \forall j$.*

equations, i.e., the second line of (1.29). This yields for each oscillator j

$$\begin{aligned} s^2\hat{Y}_j - sY_j(0) - \dot{Y}_j(0) &= -\omega_j^2\hat{Y}_j + b_j\hat{X} \\ \Rightarrow \hat{Y}_j &= (s^2 + \omega_j^2)^{-1}[b_j\hat{X} + sY_j(0) + \dot{Y}_j(0)]. \end{aligned} \tag{1.30}$$

Transforming back to real space, this yields

$$\dot{Y}_j(t) = b_j \int_0^t \frac{\sin(\omega_j t')}{\omega_j} X(t - t')\,\mathrm{d}t' + \cos(\omega_j t)Y_j(0) + \frac{\sin(\omega_j t)}{\omega_j}\dot{Y}_j(0). \tag{1.31}$$

In the last step, we have utilizes the convolution theorem to transform the first term. Using partial integration, one can further rewrite this as

$$\begin{aligned} \dot{Y}_j(t) = &- \int_0^t \frac{b_j}{\omega_j^2}\cos(\omega_j t')\dot{X}(t - t')\,\mathrm{d}t' - \frac{b_j}{\omega_j^2}\cos(\omega_j t)X(0) + \frac{b_j}{\omega_j^2}X(t) \\ &+ \cos(\omega_j t)Y_j(0) + {\omega_j}^{-1}\sin(\omega_j t)\dot{Y}_j(0). \end{aligned} \tag{1.32}$$

Now, we consider again the first line of (1.29), and substitute the $\dot{Y}_j(t)$-terms by the

corresponding expressions (1.32), yielding the following equation for X [12]

$$m\ddot{X}(t) = F - \int_0^t \underbrace{\sum_j (b_j/\omega_j)^2 \cos(\omega_j t')}_{=\gamma(t')} \dot{X}(t-t')\mathrm{d}t' + R(t), \tag{1.33}$$

which describes the exact same X dynamics as Eq. (1.29). In Eq. (1.33), we have already grouped the interaction with the bath into two terms, that is the term $R(t)$ given in (1.34) below, and the integral with the memory, which is the "friction" kernel $\gamma(t')$. Interestingly, by choosing the b_j and ω_j accordingly, one can in this way generate arbitrary kernels, as in the limit of infinitely many oscillators the sum converges to a Fourier-integral . The integral only depends on the velocity $\dot{X}$ (and the constants b_j, ω_j) and gives the "friction". This term does not depend on the initial conditions of the bath oscillators.

Further, the R term reads

$$R(t) = \sum_j \dot{Y}_j(0)\frac{b_j}{\omega_j}\sin(\omega_j t) + \sum_j b_j \left[b_j Y_j(0) - (b_j/\omega_j)^2 X(0)\right]\cos(\omega_j t). \tag{1.34}$$

It involves the initial position $X(0)$ [but not the initial velocity $\dot{X}(0)$], and all initial positions and velocities of the bath degrees of freedom, $Y_j(0)$ and $\dot{Y}_j(0)$, respectively. For a given set of initial conditions, this a *deterministic* term, describing the impact of Y_j on the dynamics of X. In contrast to the friction, it does not depend on $\dot{X}$ at any time t [nor on $X(t>0)$].

In the here demonstrated simple form, the projection method is only applicable for heat baths with linear microscopic equations. However, there is also a way to project out more general microscopic equations, e.g., nonlinear baths, which involves transformation to the Hilbert-space and application of projection operators [12]. However, the main results [a generalized LE (1.36) with FDR (1.35), as given below] and the key arguments are, in principle, always the same. We further note that we will later in Sec. 5.1 perform a projection very similar to the one shown here (however, there we start from coupled stochastic equations).

Let us return to Eq. (1.33). Yet, we have not obtained a mesoscopic description. Now comes an important, final step, which transforms equation (1.33) into a stochastic one.

1.3.2 Coarse-graining – Forgetting some details

It is unpractical and indeed not necessary to carry around the cumbersome terms $R(t)$ and $\gamma(t)$. Further, for more realistic systems, which have typically much more complicated underlying equations (where some details could be unknown), it may even be impossible to work with deterministic, i.e., precise equations such as (1.33). To overcome these complications, one typically performs a *coarse-graining* procedure by throwing away some information about the process. In this way, one can create a simpler, stochastic (approximate) description for X , by a mesoscopic equation.

In the present context, one *assumes* to this end that

- there are infinitely many bath d.o.f..
- the bath d.o.f. initial conditions $Y(0)$ are uncorrelated among each other [which they are not in naturally evolving dynamics from (1.29) at $t > 0$]

These assumptions bring one in the position to apply the *central limit theorem*. Further, one assumes that

- the equipartition theorem holds, i.e., the mean kinetic energy is $\langle \dot{Y}_j(0)^2 \rangle = k_\mathrm{B}\mathcal{T}$ and the potential energy in the springs is $\langle \omega_j^2 [Y_j(0) - (b_j/\omega_j^2)X(0)]^2 \rangle = k_\mathrm{B}\mathcal{T}$.
- alternatively to the previous items, one can assume that the initial conditions are independently drawn from a Boltzmann distribution $\sim e^{-H_\mathrm{B}/(k_\mathrm{B}\mathcal{T})}$, with the total bath energy H_B. The result is equivalent.

In other words, the bath is assumed to be infinitely large and in thermal equilibrium. This is a conceptually very important step, as it introduces stochasticity and, at the same time, an element of time-asymmetry (which we elucidate further below). By application of the central limit theorem, these assumptions allow to replace the entire R term by a Gaussian, zero-mean *noise* ν, whose correlations can be proven [12] to obey the famous *second fluctuation-dissipation relation* (FDR)

$$\text{FDR:} \quad \langle \nu(t)\nu(t') \rangle = k_\mathrm{B}\mathcal{T}\, \gamma(|t-s|), \tag{1.35}$$

also called Kubo relation of second kind [95]. Due to the construction of the LE, in particular, the assumptions on the bath which yield the FDR, the LE describes systems that relax to thermal equilibrium in the absence of external non-conservative forces F. More precisely, the FDR is connected to a vanishing entropy production. While this cannot immediately be seen from Eq. (1.35) alone, it will become clear when we will later in Part III (see Chapter 9). Thus, for a wide class of systems, the existence of an FDR implies the possibility to reach thermal equilibrium [15, 95].

However, one should be aware that corresponding statements are not trivial in the presence of different driving mechanism. For example, in the presence of velocity-dependent forces, the system can reach a state with vanishing heat dissipation that fulfills detailed balance[13] [125], while the FDR might be broken. The fundamental connections between FDR, heat dissipation and EQ, are particularly well-understood for Markovian systems, see, e.g., [194]. For systems subject to feedback, the connections seem to be less understood. We could not find rigorous statements in the literature.

In total, starting from mechanical laws and plugging in, by hand, uncorrelated initial conditions with equally distributed energy among them, one obtains a *non-Markovian* LE, called *generalized Langevin equation*

$$m\ddot{X}(t) = F - \int_0^t \gamma(t'-t)\dot{X}(t')\mathrm{d}t' + \nu(t) \tag{1.36}$$

with *friction kernel* $\gamma(t-t')$ and colored noise ν, which fulfills the FDR (1.35). The term *non-Markovian* refers to the fact that the future of the process X generated by

[13] which is normally as well an indicator for thermal equilibrium.

this LE, at each instance of time t, cannot only be predicted based on $X(t)$ alone, but one further needs information over its *past* $X(t' < t)$. Both friction and noise have history-dependence, i.e., *memory*. Despite its non-Markovianity, there are a lot of results known for generalized LE. The FDR is an important theoretical result in regard to the physical understanding. Furthermore, it also represents a practical tool in the mathematical treatment of GLEs. An example is the unique mapping from the correlation functions to a generalized LE provided by Volterra-like *memory equations*, for which we give a concrete example below.

Vanishing correlations Another implication of the construction of the bath, particularly of the assumed independence of the initial velocities, is that

$$\begin{aligned} \langle R(t)\dot{X}(0)\rangle &= 0 \\ \Rightarrow \langle \nu(t)\dot{X}(0)\rangle &= 0. \end{aligned} \tag{1.37}$$

This can readily be seen from Eq. (1.34). Thus, the future noise is not correlated with the colloid's velocity at time $t = 0$. We will come back to this point in Sec. 1.3.5. One should stress that (1.37) is not true for other times $t' > 0$, i.e., $\langle \nu(t)\dot{X}(t')\rangle$ not necessarily vanishes.

Memory equation Here we briefly derive a Volterra-like *memory equation*, which uniquely maps the velocity correlation function $C_v(t) = \langle \dot{X}(0)\dot{X}(t)\rangle$ to a specific generalized LE. To this end, we will make three important assumptions: First, the force F shall be linear in X, i.e., $F = aX$ with some constant a. Second, (1.37) holds. We further assume that the ensemble shall be in equilibrium already at the beginning $t = 0$, such that $C_v(t) = \langle \dot{X}(t')\dot{X}(t'+t)\rangle_{\mathrm{ss}}$, although the initial point was indeed a special instance in time [e.g., fulfilling (1.37), which would normally not hold in a naturally evolving system (1.29)]. Formally, the last assumption is not strictly necessary, but makes the formula much more general and applicable, and is hence usually made. This is a delicate aspect, as this assumption contradicts (1.37).

Now we start with multiplying both sides of (1.36) with $\dot{X}(0)$ and taking the ensemble average of the entire expression. This yields

$$m\langle \dot{X}(0)\ddot{X}(t)\rangle = \langle \dot{X}(0)\left[-aX(t)\right]\rangle - \int_0^t \gamma(t'-t)\langle \dot{X}(0)\dot{X}(t')\rangle \mathrm{d}t' + \langle \dot{X}(0)\nu(t)\rangle. \tag{1.38}$$

Utilizing (1.37) and the identities

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}C_v(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\langle \dot{X}(0)\ddot{X}(t)\rangle, \\ \int_0^t C_v(t')\mathrm{d}t' &= \langle \dot{X}(0)X(t)\rangle, \end{aligned} \tag{1.39}$$

one readily finds

$$\frac{\mathrm{d}}{\mathrm{d}t}C_v(t) = -\int_0^t m^{-1}[a + \gamma(t'-t)]C_v(t')\mathrm{d}t'. \tag{1.40}$$

This gives a unique mapping from C_v to γ, which, in combination with (1.35) gives a unique generalized LE. However, mind the necessary assumptions, e.g., linearity of the LE, that the system is in thermal equilibrium already at the initial time, and (1.37).

Side node: Taking this simplified model serious Although the picture of infinitely many harmonic oscillators is clearly rather a metaphor for a real fluid, than an actual, realistic description, it is still a very wide spread model for a heat bath of undeniable importance from the viewpoint of theory development. Its obvious advantage is the simple, linear form. For this reason it is worth to take a closer look onto the "noise" R.

To get an impression of how well the approximation by a Gaussian noise is, for this particular construction of the heat bath, we consider possible realizations of R from (1.34). Figure 1.5 shows the term $R(t)$ itself supplemented with its autocorrelation function. We have chosen a uniform (equidistant) distribution of ω_j and Gaussian-distributed random velocities $\dot{Y}_j(0)$, and we have depicted two different cases for the positions: once with randomly distributed initial positions $Y_j(0)$, and once with equidistantly distributed, that means *correlated*, $Y_j(0)$. Only in the former case, a truly irregularly looking temporal evolution emerges, while the latter case is rather oscillatory, with irregular amplitudes. Indeed, a random positional distribution, but equidistant velocity distribution has a similar effect (not shown here). Hence, fully uncorrelated initial conditions are essential. We conclude that the initial disorder is enough to create an irregularly looking $R(t)$ (despite the linearity of the model), as is confirmed by testing different variations[14] among the three distributions of ω_j, $Y_j(0)$, $\dot{Y}_j(0)$, and other system sizes. Remarkably, the system size seems to play a minor role. We have checked different numbers of harmonic oscillators N between $N = 10^4$ and $N = 10^6$, and found surprisingly similar results (almost not distinguishable)[15]. In no case, we found an evolution comparable to the white noise (plotted in Fig. 1.2). At most, R realizations resemble a Wiener process (compare with Fig. 1.3).

It is noteworthy that in our tests, a system with randomly distributed ω_j seems to produce an irregular $R(t)$, even if $Y_j(0)$, $\dot{Y}_j(0)$ are correlated (results not shown here). This suggests that a (frozen) disorder in the system parameters has a somewhat stronger effect than "initial disorder". However, we have not analyzed this in detail.

These considerations show that the noise approximation is justified, but only for uncorrelated initial conditions. In a sense, the randomness must hence be introduced in this model by hand. This is indeed not surprising given the linearity of all forces. A nonlinear model (which would anyway be more realistic), is of course expected to yield more disorder, and can even be truly *chaotic* for finite system sizes [196]. It is interesting to note that there is some debate about under which conditions a linear differential equation can create chaotic trajectories. This appears to be only

[14] We have not studied the effect of random b_j distribution (but simply set them all to unity here).

[15] Except for the recurrence time, e.g., how long it takes until the initial configuration is recovered, which we do not discuss here. This time strongly depends on N [195], as can indeed easily be checked from (1.34).

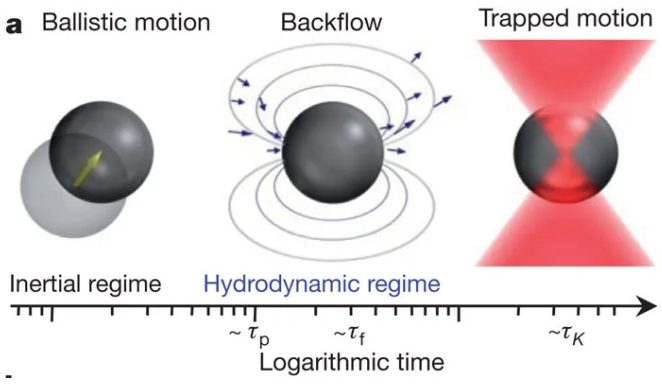

Figure 1.6: *This Figure stems from Ref. [16]. On very short timescales ($t \ll \tau_{\mathrm{p}} \approx m/\gamma$), the particle undergoes ballistic motion (left). On the timescale τ_{f}, hydrodynamic backflow develops (center); the gray lines indicate the emerging fluid velocity field; the arrows stem from computer simulations, see [16]. For $t > \tau_{\mathrm{K}}$, the harmonic potential from the optical trap (see Sec. 1.4.1) confines particle diffusion (right). The relation between τ_{f} vs. the diffusive timescale τ_{B} decides whether the Markov assumption is appropriate.*

possible for infinitely large systems, see, e.g., [197].

At this point one may further hint to the long-standing debate in statistical physics concerning the true *origin* of the *randomness* of Brownian motion and the diffusive behavior from microscopic, deterministic equations. In particular, is it a sufficient, or is it a necessary requirement that the micropscopic dynamics is chaotic, or is the extreme high dimensionality of the phase space the true reason that by reduction causes the randomly appearing motion? We refer the interested Reader to [159, 198] and references therein. Famous theoretical models in this context are the (chaotic) Lorentz-Gas [199, 200] an the (non-chaotic) "wind-tree" model [201].

1.3.3 The Markov assumption

In many situations, the colloid is much heavier, and thus *slower* than the bath particles. Then the memory kernel and colored noise both typically decay fast compared to the characteristic timescales of the colloidal motion.

When the colloidal timescales can be well-separated from the others, the memory kernel and noise-correlations are often assumed to decay *infinitely fast*, i.e.,

$$\langle \nu(t)\nu(t')\rangle \to 2\,k_{\mathrm{B}}\mathcal{T}\;\delta(t-t'),$$

and at the same time, due to FDR,

$$\gamma(t-t') \to 2\;\delta(t-t').$$

This is the *Markov assumption*, yielding a Markovian Langevin equation with *instantaneous* friction and *white noise* (1.1), free of any memory effects.

Returning to the "actual" system, i.e., a colloid in a surrounding fluid, the Markov assumption amounts to assuming the collisions between the bath particles and the colloid to be infinitely fast, and, in fact, neglecting any temporal correlations among the bath particles.

1.3.4 Real-world complications

In the real world, especially when one considers complex systems which contain various different length and timescales, the idealized, strict separation of timescales is not always possible. This can be due to the fact that the heat bath particles themselves have internal degrees of freedom. For example, a heat bath containing polymers or worm-like micelles [179], which have internal relaxation times on timescales comparable to the colloidal motion, results in viscoelastic memory [12–15, 202]. Another example of colloidal systems, where memory effects do play a non-negligible role, are suspensions with significant hydrodynamic flows [16–18, 203]. The hydrodynamic timescale of a colloidal sphere in acetone was measured in Ref. [16], see Fig. 1.6 for a schematic overview. Further, for an even much more complicated example system, think of diffusion in a cell, where the different cellular components, liquids and lipids form a highly nontrivial environment. Here, the picture of one big colloidal particle in a bath of identical, tiny particles, might loose meaning. In all these counter example, the Markov assumption is not appropriate.

Moreover, the equilibrium assumptions on the bath particles, which yield the FDR between friction and noise, are as well not always justified. An example are active systems, which we briefly discuss in Sec. 3.7. Again, the aforementioned living cell may serve as an example. Due to the fact that various components, like molecular motors, transform chemical energy into motion [184], this system is intrinsically out of equilibrium, such that FDR is not expected to hold.

Recent research aiming for generalizations An interesting research direction are the various recently derived generalization of the fluctuation-dissipation relation to non-equilibrium systems, see, e.g., [15, 98, 147, 204, 205]. Immediate applications are systems in active baths, where the bath particles themselves are self-propelled. We will later introduce active systems and explain the terminology (see Sec. 3.7). We are not aware of related research in regard to time-delayed processes, e.g., a bath or colloid under time-delayed feedback. In this thesis, we have not considered such non-equilibrium fluctuation-dissipation relations. However, it would be very interesting to see whether rigorous results can be found, which might be possible for purely linear systems.

Time delay In most parts of this thesis, we will consider another kind of non-Markovian LE than the generalized Langevin equations discussed here, the delay LE. It describes a very different type of system, where the memory comes from "outside" (not from interactions and correlations within the bath). Consequentially, there is no relation between the delay and noise correlations. Hence, we can already expect the treatment to be even more involved. Compared to the situation discussed here, the delay LE is so to say the "Markovian limit" of the generalized LE (i.e., the memory effects due to the noise are neglected), with a time delay in the feedback force F, which again renders the motion non-Markovian. Thus, for delay LEs, the non-Markovianity is "tunable", and the assumption that the system can find an equilibrium is not *per se* justified and needs to be revisited (in fact, the system is intrinsically out of equilibrium, as we will see later in Part III).

Before we finally discuss the delay LE, let us make another brief excursion discussing the notions of causality and time-reversal symmetry.

1.3.5 Time-reversal symmetry & Causality

The mechanical time-reversibility The "microscopic level" in the derivation sketched above is given by point particles governed by mechanical laws. Such mechanical, Hamiltonian systems are *time-reversal symmetric*. For the sake of illustration, imagine a set of billiard balls bumping into each other and then spreading apart, conserving total momentum and energy (we neglect friction, which slows them down). Now imagine two observers, who see a short sequence, e.g., around the moment where the billiards hit, lets say of length $\{t_0, t_1\}$ with $t_0 < t_1$. One observer should make a prediction of where each of the balls will be at a short time *later* (at $t_1 + \Delta t$), while the other observer shall make a guess where all the balls have been an (equally long) short time *before* the observed sequence (at $t_0 - \Delta t$). *Who has the better chances to be correct?* Naively one might think that the first observer may have the easier task, as he has observed the "cause" and should guess the "effect". However, in fact, both observers have the exact same chances (which are, for observers who can measure and compute perfectly, 100%).

The "microscopic", mechanical world (recall that we neglect friction) is completely reversible in time. Each mechanical trajectory seems equally realistic as its time-reversed picture[16], and one can just as well predict the "cause" by observing the "effect", as the other way round. This property of mechanical laws, the *mechanical time-reversal symmetry* [124, 206].

The stochastic (ir)reversibility and causality Things change on the "mesoscopic", stochastic level, where a *causality* principle becomes important that distinguishes between future and past.

For example, the LE with white noise describes a process where the correlations between the system state X and the past noise, $\langle X(t_2)\xi(t_1)\rangle$ and $\langle \dot{X}(t_2)\xi(t_1)\rangle$, with $t_1 < t_2$, can be *non*-zero, but where in sharp contrast [85]

$$\langle X(t_1)\xi(t_2)\rangle \equiv 0 \tag{1.41a}$$

$$\langle \dot{X}(t_1)\xi(t_2)\rangle \equiv 0 \tag{1.41b}$$

strictly hold. The stochastic variables cannot be correlated with the future noise, but with its past noise, which is often denoted causality principle. Remarkably, this does indeed not necessarily disagree with our physical intuition, and almost seems self-evident. Yet, it implies a sharp distinction between future and past, contradicting the above described mechanical time-reversal symmetry [124]. For the physical understanding of this symmetry breaking it is important that it concerns ensemble average, i.e, averages over all possible realizations. One should further emphasize that (1.41) even holds in thermal equilibrium, when the system obeys *thermodynamic reversibility*. The latter is another concept, which we will explain below.

[16] This concepts will be further elucidated in Chapter 3 in the context of entropy production.

Further, the non-Markovian generalized LE with colored noise considered above, yields the asymmetric correlations

$$\langle \dot{X}(t_1)\nu(t_2)\rangle \neq \langle \dot{X}(t_2)\nu(t_1)\rangle. \tag{1.41c}$$

(The strict statement $\langle \dot{X}(t_1)\nu(t_2)\rangle = 0$ is only true for white noise). Again, this implies breaking time-reversal symmetry, even in equilibrium.

The Eqs. (1.41) demonstrate that the LE does not fulfill time-reversibility in the way the mechanical laws do, i.e., the LE already distinguishes between future and past. The underlying reason is an element of asymmetry introduced within the derivation, which is often done by a specific choice of initial conditions. In the aforementioned derivation by Mori and Zwanzig that is the assumption of *un*correlated initial conditions of the bath degrees of freedom. In particular, at the end of the projection methods, one must assume that all bath particles have independent initial positions and momenta. Although not directly formally related, this assumption resembles the Stosszahlansatz or molecular chaos hypothesis[17] from kinetic gas theory [191, 192]. This step in the derivation is essential to obtain a LE with Gaussian white or colored noise, which fulfills the second FDR. Indeed, such a step is also crucial in the famous derivation of the fluctuation-dissipation relation by Kubo [95], as is explicitly discussed in Ref. [209]. This derivation is somewhat more general (not relying on a specific model for the bath), but in fact is build upon the assumptions $\langle \nu(t)\dot{X}(0)\rangle = 0$, and, more trivially, $\langle \nu\rangle = 0$ (which is consistent with the expected translational symmetry of the overall system[18] [210]). This first assumption, which is identical to (1.37), appears to be essential. All these considerations show that a *causality* argument, or other time-reversal symmetry breaking, which may be introduced by the initial conditions, is crucial to find a stochastic description. Such an assumption is also denoted *axiom of causality* [211].

Considering the following course of this thesis, we discuss this topic here for three reasons. First, we will employ (1.41) multiple times in Parts II and III. Second, we will encounter a mathematical and physical problem in the formulation of stochastic thermodynamics for non-Markovian processes, which is related to the acausality of the backward process. This will be discussed in Chapter 3. As a primal step towards this problem, it is necessary to be aware of the role of causality in the considered stochastic models. Third, it is interesting and important to understand the *different notions of reversibility*.

We have already explained mechanical reversibility, and the stochastic irreversibility above. Further, in thermodynamics, (ir)reversibility is a central concept, which we will introduce in Chapter 3. *Thermodynamic reversibility* signals equilibrium situations, where the system is stationary and energy (and probability) flows vanish on average (we will elucidate this concepts later in Chapter 3).Under suitable conditions, Langevin equations can indeed describe such thermodynamical reversible

[17] Here one should also mention Lochschmidt's paradox [207, 208], which concerns the impossibility to derive time-asymmetric laws (referring to the second law of thermodynamics, see Chapter 3) from time-symmetric dynamics (mechanical microdynamics).

[18] or equivalently the overall momentum conservation of all particles including bath's, as given by Noether's theorem

equilibrium trajectories, despite the inherent mechanical time-reversal symmetry loss. Furthermore, Langevin equations are also capable of describing nonequilibrium processes, which attempt to maximize their entropy (more precisely, their free energy) possibly hindered by external driving mechanisms. However, Langevin equations *cannot* generate "unphysical" trajectories, which violate the second law of thermodynamics (i.e., reduce the total entropy, see Chapter 3), although purely mechanical systems *can*[19] We will come back to this point and extend the discussion about causality in LEs in Sec. 3.4.

Macroscopic reversibility We briefly aim to mention that there is indeed yet another notion of reversibility relevant in this context, describing macroscopic processes. It is an important concept in traditional macroscopic thermodynamics, for example for the realization of a Carnot cycle [115]. It refers to the possibility of retracing the history of macrostates in the opposite order in time, which often requires *external driving* of the system in a *quasi-static manner*, i.e., so slow that the system has the chance to "reach equilibrium" at each step. To this end, the macroscopic states shall be infinitesimally close to each other [212]. In this thesis, we are not concerned with such (often cyclic) processes, and will thus not consider this type of macroscopic reversibility. However, it is indeed build upon the aforementioned thermodynamic reversibility of equilibrium systems, which is, in contrast, concerned with individual trajectories.

Experimental confirmation of the time-reversal asymmetry In real systems, the white noise assumption never strictly holds, such that one shall expect $\langle \dot{X}(t_1)\nu(t_2)\rangle \neq 0$ in the corresponding stochastic model. As we have mentioned above, the causality principle then amounts to the asymmetry of $\langle \dot{X}(t_1)\nu(t_2)\rangle \neq \langle \dot{X}(t_2)\nu(t_1)\rangle$. In Ref. [17], this asymmetry was experimentally confirmed for a colloid in a trap (realized by "optical tweezers", see Sec. 1.4.1). Due to hydrodynamic coupling between colloid and bath, the colloid's velocity is also correlated with the future noise, as one indeed would expect considering the underlying mechanical motion of all individual particles. The measured asymmetry of the correlations underlines the time reversal symmetry-breaking (in this equilibrium system).

1.4 Introduction to the Langevin equation with time delay

Now we turn to the main topic of this thesis, that is, *stochastic systems with time delay*. In the last Section, we have seen that stochastic motion emerges by coarse-graining out of a microscopic world which is governed by Newton's laws. We have further seen that this stochastic motion is describable by a (generalized) Langevin

[19]There are, in fact, equally many possible trajectories of a mechanical system that minimize and maximize entropy. For every trajectory which increases entropy, there is one which decreases entropy, obtained by reversing all velocities of the end point configuration and use this as new initial condition.

equation, which approaches thermal equilibrium in the absence of external driving. In the Markovian limit, this gives rise to a Langevin equation with white noise and instantaneous friction. On the other hand, if the timescales between coarse-grained environment and particle dynamics are not well-separated, one encounters memory in the form of a friction kernel and noise correlations, which are connected by the FDR, as automatically follows from the underlying first principles.

In realistic systems, however, memory may also arise from other sources (not only from the bath itself), such that this type of description is not always the end of the story. For example, in various biological, physical or socio-economic systems, memory in the form of *time delay* naturally arises. Such a delay can have various origins, including finite information processing or decision making times [33–36], signal transmission can create delays in the communication between different agents [213], or finite refractory times [29–32]. Another very important origin of delay are feedback loops [8, 10].

In the following, we make a little detour and briefly discuss an important experimental technique that can be used to realize time-delayed feedback on colloidal particles, creating a process describable by a Langevin equation with time delay. After this, we will return to the theoretical considerations and review some important features of the LE with delay.

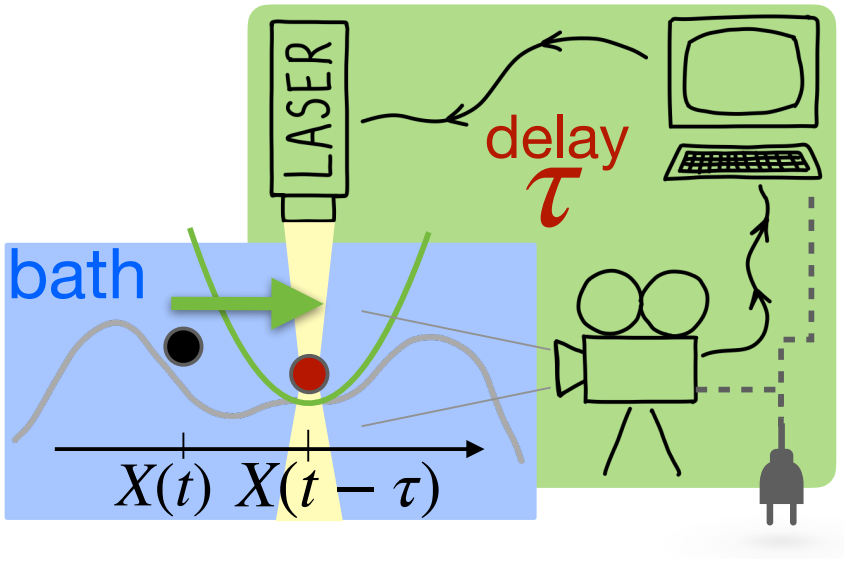

Figure 1.7: *Sketch of a possible experimental realization of a feedback control loop acting on a colloidal particle (black disk). The particle at position $X(t)$ feels a linear force (green arrow and green parabola) towards the delayed particle position $X(t-\tau)$ (red disk). This force stems from the feedback loop (green area) consisting of a camera, a computer and a laser (optical tweezers) and involves a time delay τ. The colloidal particle is immersed in a fluid which represents a heat bath (blue area), and is further subject to an external potential (gray line).*

1.4.1 Optical tweezers – An experimental tool to control

"*The possibility of using light [...] to move physical objects may make one think of Star Trek and tractor beams, and sound like pure science fiction.*"

– Anders Irbäck, from the Nobel prize ceremony speech, 2018 [214]

In this thesis, we are mainly concerned with stochastic systems with time delay. Such a feedback can be implemented in experiments via optical tweezers [215, 216]. To this end, the particle position is measured via a camera, and the tweezers are then focused onto the previously measured position. Figure 1.7 shows a sketch of such a setup. As measurement and information processing take time, this feedback loop inevitably involves a time delay τ, such that the optical tweezers are actually focused onto the *delayed* position. Of course, the delay can in principle also be intentionally enlarged to an arbitrary value, by storing the past measured positions, and focusing the laser onto the position a specific time τ ago. It should be noted that there are further, similar techniques to realize feedback on colloidal systems or single molecules, for example by temperature gradients, see, e.g., [66–69, 217].

Arthur Ashkin received in 2018 the Nobel prize in physics for his development of a device which enables trapping without damage of various small sized objects, called *optical tweezers* [218], "and for their applications to biology" (official prize motivation [219]). Enabling targeted trapping and moving of individual colloids, molecules, or atoms, i.e., objects from Ånström to micrometer scale, this device introduced a whole new way of investigating small scale systems and of probing theoretical ideas. Today, this tool is very widely used [61–63, 65, 150, 220–222], and experiments based on it represent the backbone of single-molecule biophysics, as well as stochastic thermodynamics.

The optical tweezers or *optical traps* consist of a Gaussian-profiled laser beam. An example for a colloid which can be controlled in this way is a transparent latex sphere, whose refraction index is higher than the surrounding fluid. Due to refraction of the laser beam at the surfaces of the latex sphere, its non-uniform light field exerts a force onto the sphere (stemming from radiation pressure). This force is directed towards the center of the beam, where the light intensity is highest. Thus, the sphere gets trapped in the center, and by moving the laser beam one can move the colloid around in a very precise manner. This technique can also be used to control other objects, like (bio)molecules, or single atoms [223]. By clever choices of laser frequencies, beam geometry and light intensity, it is even possible to trap individual cells or organisms, like blood cells [224], algae, or bacteria [225], as well as individual components within a living cell [226] while keeping the heating effects from the laser moderate, such that living organisms can survive this procedure. (See [227] for an estimation of the heating effect due to the laser light).

Experimental applications of optical tweezers Here we briefly mention some experimental techniques based on optical traps, which have proven themselves particularly important in the field of stochastic thermodynamics in the last years.

First, an optical trap can be used like a spring scale[20]. This enables measurement of forces, e.g., involved in biomolecular processes. For instance, by trapping the cargo of a Kinesin enzyme walking along a microtubule, its strength can be measured [101]. Kinesin is a type of molecular motor, fueled by ATP, which is essential for transport of, e.g., protein and membrane components within cells [100]. Similarly, when attaching a protein, e.g., DNA, to two trap- and movable latex spheres, the forces and the (fluctuating) work necessary to fold and unfold the protein can be tested [228]. (From such nonequilibrium experiments, one can even infer the free energy by means of the Jarzynsky relation. The latter relation will be given later in Chapter 3.)

Second, the precisely tunable optical trap can be used for "compression" or "expansion", similar to a movable piston on a macroscopic gas. In particular, a stiffer or looser trap yields more or less confinement of the thermal fluctuations, which is essential for the mesoscopic realization of a Carnot engine [229].

Third, optical traps can be used to build specific potential landscapes [150, 230], or generate desired interactions between colloidal particle [66]. In this way, one can experimentally realize desired forces in the LE of collective or single particle systems.

Finally, another relevant experimental technique based on optical traps, which is particularly important for the present thesis, is *time-delayed feedback* of colloidal particles [215, 216]. The resulting motion of the colloid can then be described with a delay Langevin equation. In the following, we will briefly outline such an experiment.

1.4.2 Time-delayed feedback

In a real experiment, the measurement is typically not continuous, but happens at discrete time, due to a finite temporal resolution of the camera. However, in this thesis, we will assume that the camera precision is high compared to the other timescales of interest, giving rise to consider the limit of infinitely fast measurements, i.e., a time-continuously operating feedback.

While recent experiments even realize delayed feedback control on multiple colloids at the same time [66], we will here only consider *single* Brownian particles, as there are still a lot of open questions for this case. The interplay of delay and noise makes the analytical investigation of such systems particularly challenging, but is, at the same time, found to play a crucial role for its behavior. Colloidal particles under optical tweezers feedback represent an ideal system to study this interplay.

The theoretical study of the effects of such a *time delay* on the particle position fluctuations is the main focus of this thesis. While the corresponding Fokker-Planck equation is not trivially given and requires special attention (this is the goal of Part II), the equation of motion of the colloid, that is, the Langevin equation can readily be written down.

[20]The force can be measured through measurement of the distance to the trap center.

1.4.3 The Langevin equation with time delay

Returning to our theoretical considerations, we now investigate the overdamped delay Langevin equation

$$\dot{X}(t) = \gamma^{-1} F[X(t), X(t-\tau)] + \sqrt{2D_0}\,\xi(t), \tag{1.42}$$

where F denotes the deterministic force which is now *nonlocal in time* and is given by some (generally nonlinear) function depending on the instantaneous and on the delayed system state, $X(t)$ and $X(t-\tau)$, with $\tau \geq 0$ being the single *discrete time delay*. From a mathematical point of view, (1.42) represents a *stochastic delay differential equation* (SDDE).

In the main examples considered in this thesis, we typically separate the force into a conservative part $F_\mathrm{s}[X(t)] = -\partial_x V_\mathrm{s}(x)|_{x=X(t)}$ stemming from a static potential $V_\mathrm{s}(x)$, and a "delay force" F_d, which explicitly depends on the delayed particles position $X(t-\tau)$. This force introduces *feedback* into the system, which is *a drive that depending on an (earlier) system state.* (An overview over different aspects of feedback control can be found in [8].) This feedback may stem from an external controller as mentioned in the last Section. The total deterministic force can hence typically be represented as

$$F[X(t), X(t-\tau)] = F_\mathrm{s}[X(t)] + F_\mathrm{d}[X(t), X(t-\tau)]. \tag{1.43}$$

Due to its simple functional form, a common choice for the delay force, or "feedback force", is a linear force of type

$$F_\mathrm{d}[X(t), X(t-\tau)] = -k[X(t) - X(t-\tau)]. \tag{1.44}$$

For instance, Eq. (1.44) is a common choice in control theory, defining a Pyragas control (see Sec. 1.4.4). Further, at $k > 0$, such a force can be used to approximate the force of optical tweezers. Particularly, one can consider (1.44) to be stemming from a *history-dependent* quadratic potential

$$V_\mathrm{d}(x, x_\tau) = (k/2)[x - x_\tau]^2, \tag{1.45}$$

which can be used to model an optical trap centered around $X(t-\tau)$ [62], with $F_\mathrm{d}[X(t), X(t-\tau)] = -\partial_x V_\mathrm{d}(x, x_\tau)|_{x=X(t), x_\tau = X(t-\tau)}$.

Positive/negative feedback The parameter k in (1.44) denotes the *feedback strength*, or delay force strength. In the case of the optical tweezers it corresponds to the laser intensity (which can only be positive). More generally, its sign decides whether the particle is pulled *towards* the delayed position for $k > 0$, which is usually called *negative feedback*, or if its pushed *away* from the delayed position for $k < 0$, denoted *positive feedback*. Note that negative k corresponds to positive feedback, and vice versa, which can be confusing! For the sake of illustration, one may think of the former type as attractive feedback, and to the later as repulsive feedback (in a metaphorical sense). In a steady state, where $X(t) \approx X(t-\tau)$, negative (i.e., attractive) feedback has the tendency to counteract perturbations, while positive

(repulsive) feedback rather has a destabilizing effect. These tendencies are also observed in more general control schemes [8]. We will see, however, that both types can also have quite different effects than that. While we consider linear delay forces in most parts of this thesis, the other deterministic forces will, in general, be *nonlinear*.

Please note that while we will consider position-dependent feedback in this thesis, feedback forces may, in general, also involve other dynamical quantities, like velocities. Examples can be found in [63, 231, 232].

As can immediately be seen from the definition (1.44), the force loses its history-dependence in the trivial cases $k = 0$ or $\tau = 0$. Thus, the dynamics becomes Markovian in these limits. If we do not want to distinguish between both cases, we will sometimes refer to the (general) Markovian limit with $k\tau = 0$ for the sake of a simpler notation.

Initial conditions Due to the non-locality in time, the initial condition is a *history function* $\phi(-\tau \leq t \leq 0)$, i.e., $X(t) = \phi(t)$, for $t \in [-\tau, 0]$. It can be fixed for the entire ensemble or be drawn from a distribution $\mathcal{P}(\phi)$. As setting $\phi(t)$ amounts to specifying infinitely many values, i.e., infinitely many degrees of freedom, the initial condition indicates that the delayed process is *infinite dimensional*. We will encounter mathematical and physical properties connected to this infinite-dimensionality at various times throughout the thesis.

Memory kernels and distributed delay In general, the control could also involve *distributed delay* [32, 91, 92, 122, 233–240], where the delay force $F_\mathrm{d}\left[\int_{-\infty}^{t} K(t-t')X(t')\mathrm{d}t'\right]$ involves a convolution with a *memory kernel* K specifying the distribution of delay, and may depend on the entire past trajectory. The discrete delay introduced above corresponds to the special case of a delta-distributed memory kernel $K(t-t') = \delta(|t-t'|-\tau)$. In the case of distributed delay, the initial conditions in general require to specify a function on $]-\infty, 0]$.

We will in this following mostly focus on the discrete delay case, and later give more insight into the distributed delay cases in Parts II and III, (e.g., see Sec. 5.1, and Chapter 9). In Sec. 5.4 we will compare trajectories of systems with discrete and distributed delay.

1.4.4 Side note: Delay differential equations

Before we begin with discussing dynamical properties of Langevin equations with time delay, it is instructive to first consider the deterministic case, given by a *delay differential equations* (DDE). To this end, let us for a moment forget about the noise term in the Langevin equation (1.42), i.e., $\mathcal{T} \to 0$. Alternatively, one can consider the ensemble average of (1.42)

$$\langle \dot{X}(t) \rangle = \gamma^{-1} \langle F[X(t), X(t-\tau)] \rangle + \underbrace{\langle \sqrt{2D_0}\,\xi(t) \rangle}_{=0}. \tag{1.46}$$

Both strategies render DDEs.

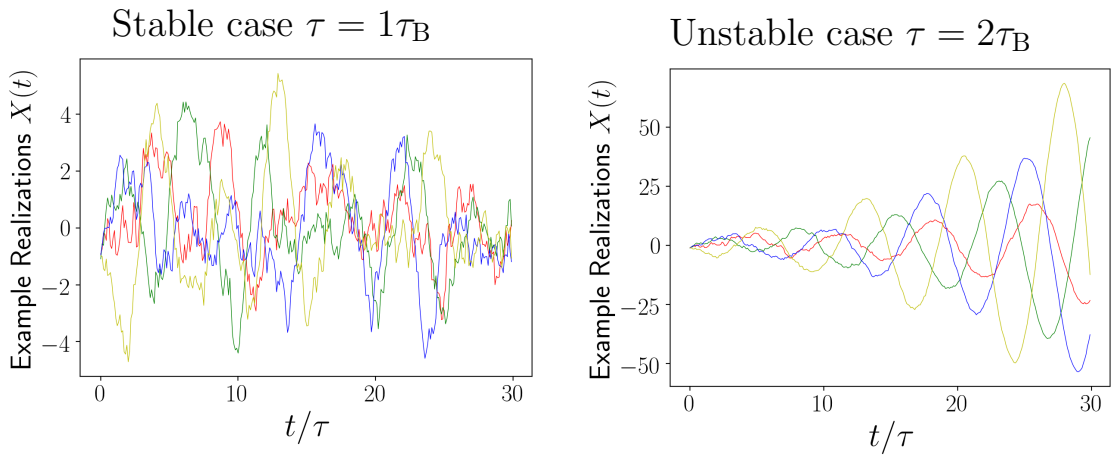

Figure 1.8: *Change of stability upon the increase of delay time. Both panels depict transient exemplary trajectories (in units of σ) of the system with LE (5.46), with history function $X(t') = -\sigma$, $\forall\ t' \in [-\tau, 0]$. This system is stable at short delay times, e.g., at $\tau = 1\tau_B$ (left panel), while it is unstable if $\tau > (\pi/2)\tau_B$, e.g., at $\tau = 2\tau_B$ (right panel). The stability boundaries have been calculated using the deterministic limit $\mathcal{T} \to 0$ (see text). Figure 5.3 shows corresponding steady-state trajectories and the autocorrelation function of this system in the stable regime.*

The reason to start with the noise-free case is that DDEs have already been quite well studied in mathematics, see [122, 241–245] for comprehensive introductions, and Ref. [122] for applications to life sciences. The discussion of DDEs (which are a special type of functional differential equations) is far beyond the scope of this thesis, and we will here only discuss a few selected aspects relevant for our work.

First, we mention that delay is known to have a smoothing effect on the solutions. For example, when the initial condition has a discontinuity, this is typically washed out in the course of time rendering smooth solutions for $t \gg \tau$.

Furthermore, it is generally known that delay can induce *oscillatory behavior*. For example, a first-order, scalar (one-variable) differential equations with time delay, i.e., a scalar DDE, can have oscillatory solutions, while an ordinary scalar differential equation cannot. We illustrate this using the simplest possible example

$$\dot{X}(t) = -X(t - \tau). \tag{1.47}$$

It is quite easy to see that at $\tau = \pi/2$, this equation has an oscillatory solution $X = \sin(\omega t + \psi)$, with ψ depending on the initial condition.

A very interesting, and often overlooked property of DDEs is that their *backward solutions are not unique*. When instead of the initial condition, one fixes the last interval (of minimal length τ), and tries to infer the corresponding solution on the preceding time intervals, which means solving the equation backwards in time, this problem has no unique solution. In fact, even when fixing the solution on $[\tau, \infty)$, it is not guaranteed that the solution on $[0, \tau]$ is uniquely defined. See Ref. [242] p. 248 for a concrete example. This is in sharp contrast to ordinary differential equations (ODEs), which can be solved equally in both directions in time.

Now we turn to the stability properties. Delay can have both, destabilizing and

stabilizing effects. Indeed, the linear delay force in Eq. (1.44), which is in mathematical literature denoted Pyragas control [246–248], can be used for stabilization. It has been extensively studied in the context of chaos control [10], and is still subject of ongoing research, for example, aiming for generalizations and proofs of controllability [46, 48]. For deterministic systems, Pyragas is a *non-invasive* control, that means, it vanishes when the system is in a steady state, or on a *periodic orbit* with period τ. This is because in both of these cases $X(t) \equiv X(t-\tau)$. Although being zero on a targeted periodic orbit or fixed point, the control term can nevertheless change its stability and, thus, provides a practical tool for stabilization.

On the contrary, delay can also destabilize a system. Again, we look at our simple exemplary DDE (1.47), which is in the absence of delay, $\tau = 0$, clearly a stable system which will always eventually approach $X = 0$, no matter how the initial condition is chosen. It turns out that it is also stable, as long as $0 \leq \tau < \pi/2$, but becomes unstable at $\tau \geq \pi/2$. In particular, the "fixed point" at $X = 0$ changes its stability. This can be shown by a linear stability analysis. To this end, one makes an exponential ansatz $X = Ce^{\lambda t}$ (with some constants C, λ). If one finds $\lambda < 0$, this indicates that the system is stable, while $\lambda > 0$ signals instability. In the present example, the exponential ansatz readily yields $\lambda X = -Xe^{-\lambda\tau}$, thus, implying the so called characteristic polynomial

$$\Delta(\lambda) = \lambda + e^{-\lambda\tau} \stackrel{!}{=} 0, \tag{1.48}$$

whose roots mark the point where the stability changes. Due to the delay, the characteristic polynomial is a "transcendental" algebraic equation with infinitely many complex[21] solutions λ, given by the "branches" of the Lambert W function $\lambda = W_k(-\tau)$ (k marking the branch). The solution λ with the highest real part changes its sign at $\tau = \pi/2$, where the system gets unstable. For $\tau < \pi/2$, (1.47) may exhibit decaying oscillations, e.g., when $X(t < 0) = \text{const} \neq 0$, while the oscillations increase for $\tau > \pi/2$. Thus, stable oscillations only occur right at very bifurcation point, i.e., the parameter where the system's stability changes. Figure 1.8 shows exemplary realizations of the transient behavior of (1.47), at a delay time before and after the stability boundary $\tau = \pi/2$. Here, we have already added noise to the system [i.e., the DDE is actually (5.46)].

A quite generally applicable strategy to find explicit solutions of DDEs is by solving them in an iterative manner, using the so-called *method of steps* [9]. This amounts to solving the equation from the Nth interval $[(N-1)\tau, N\tau]$, to the next, $N+1$st one, each time utilizing the previous solution to replace the delay term, and starting with the history function for $N = 0$. By construction, when applied directly to the equation of motion, this technique is not suitable to analytically access the steady-state dynamics, as one would have to repeat the procedure in principle infinitely often (while numerically, this can be done). In Sec. 4.2.1, we will explicitly demonstrate how this technique can be used to access steady-state properties, by applying it to the autocorrelation function.

[21] The imaginary part signals oscillatory solutions.

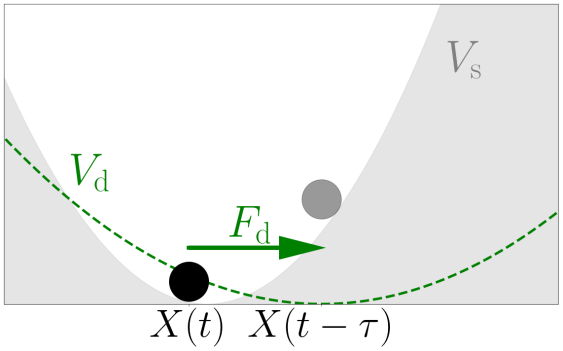

Figure 1.9: *Visualization of the linear system with time delay from Eq. (1.49). The particle moves in a static quadratic potential V_s and is subject to a linear feedback force $F_\mathrm{d} = -k[X(t) - X(t-\tau)]$, which stems from a quadratic trap around the delayed position.*

Now we return to the Langevin equation with noise and delay. This is a stochastic delay differential equations, which is a specific the type of DDE. Some results for the corresponding DDEs carry over, an example being the stability boundaries (discussed above). However, due to the combination of noise and delay, the analytical treatment of SDDEs is very involved, and, in general, much fewer rigorous results are known. For example, no general solution method has been found till date. An exception is the case of linear SDDEs, which we will discuss next (in Sec. 1.4.5). In this special case, formal solutions can be found for example by employing the aforementioned method of steps.

SDDEs are subject of ongoing research. A good introduction is provided in [9], numerical aspects are discussed in [249], and examples for further recent studies in mathematical literature include [250–252], in physics [84], and nonlinear dynamics [253]. Stability of systems with distributed delay was recently considered in [233, 236]. However, for many nonlinear cases, which we will extensively consider in this thesis, the method of steps and many other techniques are not applicable. We remind the Reader that for nonlinear differential equations often no explicit solutions are known, even in the absence of noise and delay (see, e.g., Ref. [177]).

1.4.5 Linear systems with time delay

To begin with, we consider a very important example system, that is the purely *linear system with time delay*

$$\gamma\dot{X}(t) = -\alpha X(t) - \beta X(t-\tau) + \sqrt{2D_0}\,\xi(t), \tag{1.49}$$

which was already discussed by Kuechler and Mensch in the important 1992 paper [77]. The LE (1.49) can be used to describe a Brownian particle in a static harmonic trap $V_\mathrm{s} = \frac{a}{2}x^2$ subject to linear feedback $F_\mathrm{d} = -k[X(t) - X(t-\tau)]$, from optical tweezers (see Sec. 1.4.1), then $\alpha = a + k$, $\beta = -k$. Figure 1.9 shows an illustration of this system.

We will consider (1.49) at multiple times throughout this thesis. We will see that for these purely linear models, many mathematical problems due to the delay are, in fact, not that dramatic after all. For example, the steady-state position-autocorrelation function can be computed analytically [77].

1.4.6 Autocorrelation function of linear delay equations

The temporal position-autocorrelation function (1.10) between the systems state at time t and time $t+z$, i.e., $C(z) := \langle X(t)X(t+z)\rangle_{\mathrm{ss}}$ on $z \in [0,\tau]$, is given by

$$C(0) = D_0 \frac{1 + (\beta/\omega)\sinh(\omega\tau)}{\alpha + \beta\cosh(\omega\tau)}, \tag{1.50}$$

$$C(z) = C(0)\cosh(\omega z) - (D_0/\omega)\sinh(\omega|z|), \tag{1.51}$$

with $\omega = \sqrt{\alpha^2 - \beta^2} \in \mathbb{C}$. Note that ω becomes imaginary if $|\alpha| < \beta$, such that the hyperbolic functions in Eqs. (1.50) convert to trigonometric ones. This explicit closed-form expression for the correlation function on the time difference interval $[0,\tau]$ was derived in earlier literature [82, 85]. The first derivation (based on the LE) stems from Kuechler and Mensch [77]. Later in Sec. 4.2.2, we will present a derivation for $C(z)$ up to 2τ.

In Sec. 1.4.4 we have briefly discussed a way of finding stability boundaries based on a linear stability analysis. For the system with delay and noise, another indicator for stability is the finiteness of the correlation function: A diverging correlation function implies that there exists *no stable* steady state. Based on the explicit expressions for $C(0 \le z \le \tau)$, stability conditions for the linear delay LE were rigorously proven in [77]. In particular, if $|\alpha| < \beta$, there are critical values of the delay time $\tau_c\omega = \arccos(-\alpha/\beta) + 2\pi\kappa,\ \forall\kappa \in \mathbb{Z}$, for which C diverges. Furthermore, no stable state exists, if $-\beta \le \alpha$, or $\tau\alpha \le -1$. Only if $-\alpha < \beta < \alpha$, the system is stable for all delay times τ. In Fig. 1.8 we have shown examples for stable and unstable transient dynamics of the system (1.49) with $\alpha < \beta$ (in particular, $\alpha = 0$, $\beta = \gamma/\tau_{\mathrm{B}}$).

Now we focus on regions where the system is stable, such that we can consider steady states of (1.49). Let us first have a look at (numerically generated) trajectories and see what we can learn from them. Figure 1.10 (left panels) displays exemplary steady-state trajectories for different signs of β. First we notice that the irregular trajectories look, in fact, rather similar for positive and negative β, although these two cases are quite different ($\beta > 0$ corresponds to positive, repulsive feedback, i.e., a force *away* from the delayed particle position, while the feedback pushes the particle away if $\beta < 0$).

To further get an impression for the impact of delay, we compare them to the trajectories of the (non-delayed) Ornstein-Uhlenbeck process depicted in Fig. 1.4 (see Sec. 1.2.4), which corresponds to the linear delay model (1.49) at $\tau = 0$ (with $a = \alpha+\beta$). We observe that the trajectories[22] of (1.49) are not clearly distinguishable from those of a colloid subject to a static restoring force. Thus, visually inspecting the erratic trajectories alone, does not provide interesting insights.

Delay-induced oscillations in the linear system While the erratic trajectories did not yield an interesting view on the impact of delay in general, or positive vs. negative feedback, the position-autocorrelation functions, which are depicted in the

[22] Please note that we purposefully used the same random number seeds in all example trajectory plots shown in this Chapter, in order to make the comparison even easier.

right panels of Figs. 1.4 and 1.10, are found to have very distinct characteristics. For example, only the non-delayed case has a monotonically decaying correlation function (Fig. 1.4, right panel). As a quite generic feature of the cases with delayed feedback, we here detect a pronounced peak around (multiples of) τ. These peaks indicate oscillatory behavior of the particle position, which occurs for both, positive and negative feedback. These position oscillations have a stochastic character, as can be seen from the fact that they are not really visible in the trajectories. This is in sharp contrast to oscillations induced by external periodic drivings (e.g., in systems without no delay), where all trajectories would be periodic and in phase.

Closely inspecting the autocorrelation functions, we find that the *mean oscillation period* is about 2τ for the case of positive (i.e., repulsive) feedback, i.e., a force away from the delayed particle position, and roughly τ for negative (attractive) feedback, which pulls the particle back. Furthermore, the positive feedback is found to yield negative values of C, i.e., *anti*-correlations, corresponding to a high probability to have different signs at times $X(t)$ and $X(t+t')$.

As we have already mentioned in Sec. 1.4.4, the occurrence of oscillations is a quite ubiquitous phenomena encountered in various dynamical systems with time delay. In the absence of noise, they typically have a transient character, while they here appear in the steady state. We will also consider delay-induced oscillations in nonlinear systems later.

Physically, one might expect that the delay-induced oscillations are only observable when τ is large compared to the other timescales. In the present case, the relevant timescale is the diffusive relaxation within the static quadratic potential. Assuming that this timescale was independent of the time-delayed force, the relaxation timescale can be estimated to be about $\gamma/a = \gamma/\alpha = \tau_{\mathrm{B}}$ here (see Sec. 1.2.3). Indeed, the oscillations are very pronounced if $\tau = 10\tau_{\mathrm{B}}$ (Fig. 1.10), but they are not visible for $\tau = \tau_{\mathrm{B}}$. The latter case is depicted in Fig. 1.11, where again positive and negative feedback are shown.

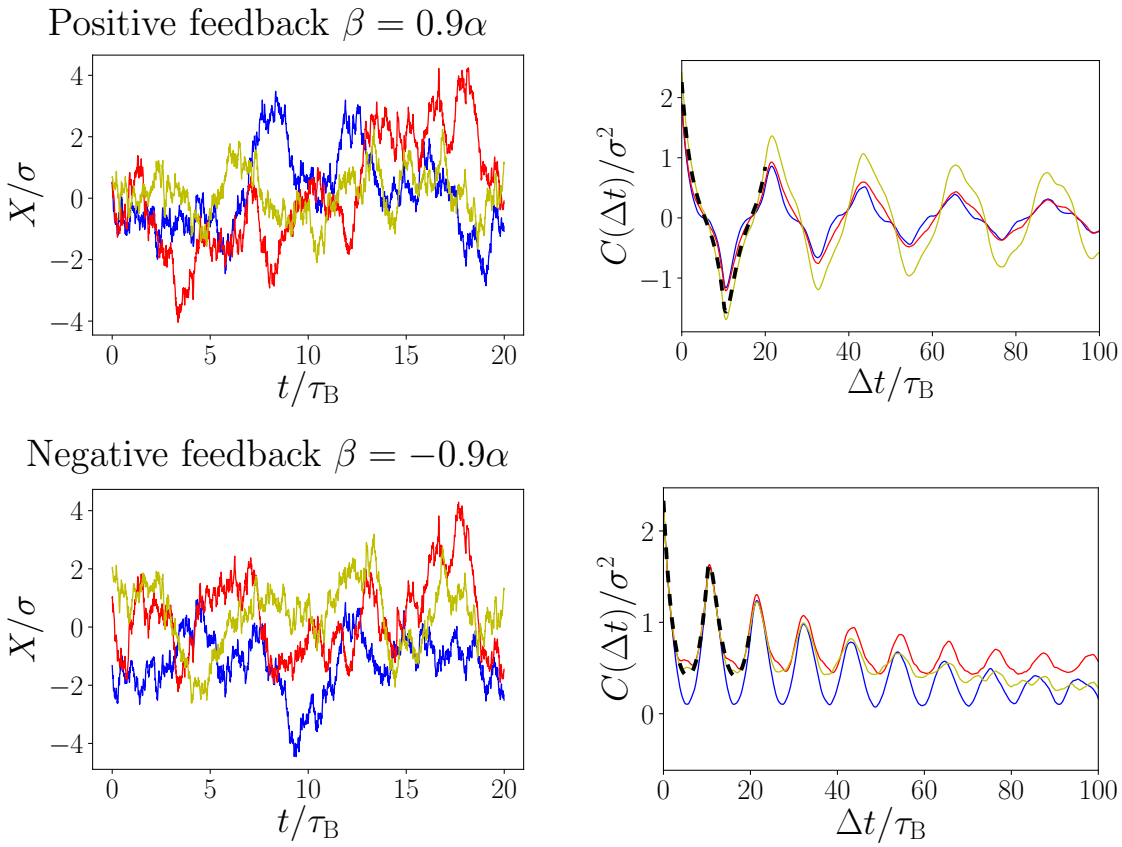

Figure 1.10: *Left panel: Three numerically generated example steady-state trajectories of an Brownian particle in a harmonic trap $V_s = [(a-k)/2]x^2$, subject to a linear time-delayed feedback $-k[X(t) - X(t-\tau)]$. The LE reads (1.49), with $\alpha = a$, $\beta = -k$. The delay time is rather large, $\tau = 10\tau_B$. Right panel: Corresponding position-autocorrelation function from the three single numerical trajectories (recall that the process is ergodic), and as shown by a dashed black line from the analytical results (1.51, 4.23) for $t \in [0, 2\tau]$. The derivation of the explicit expression (4.23) for $t \in [\tau, 2\tau]$ will be presented later in Chapter 4. To ensure steady state conditions, we have cut off about 10^5 time steps in the numerical simulations (see the Appendix A.1 for further details). Here, $\alpha = 1(\gamma/\tau_B)$, $\beta = \pm 0.9(\gamma/\tau_B)$. The upper panel case is positive ("repulsive") feedback, where the particle is pushed away from the delayed position, while it is pulled towards it in the lower panels, corresponding to so-called negative ("attractive") feedback.*

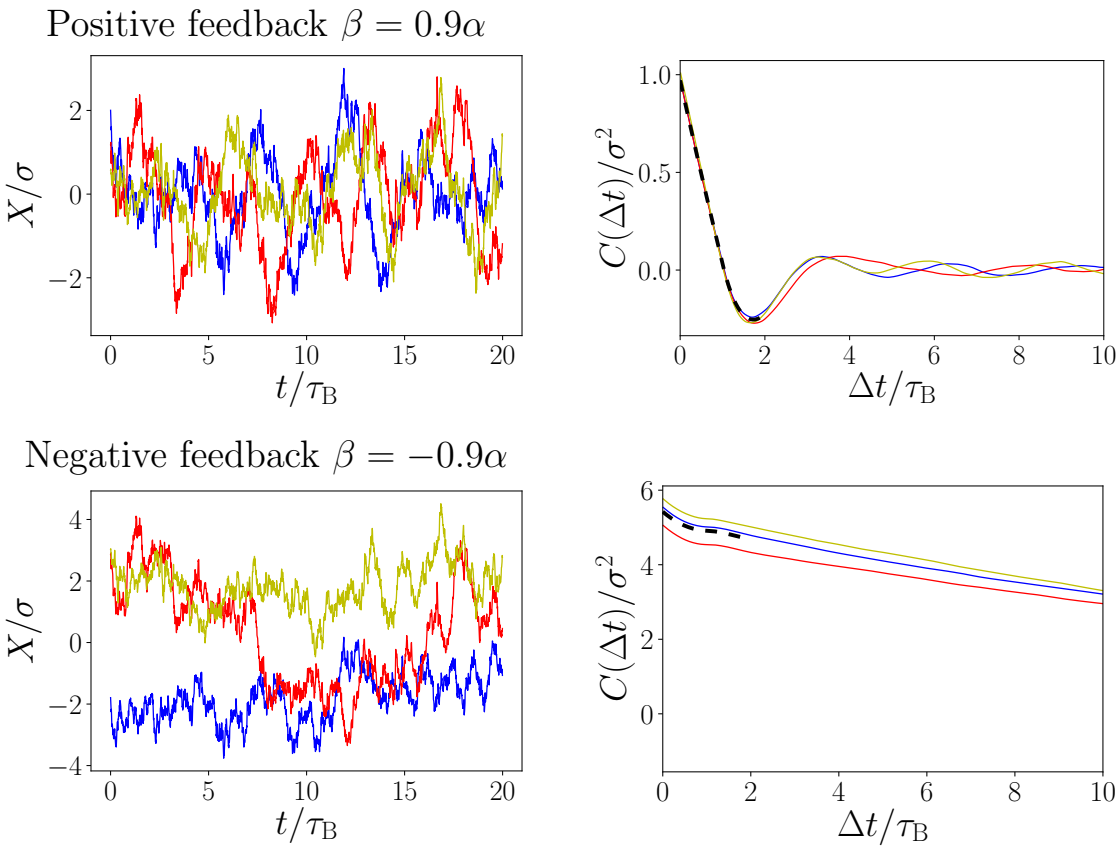

Figure 1.11: *Example steady-state trajectories and corresponding correlation function as in Fig. 1.10, but here in the case of a rather short delay time, $\tau = \tau_B$. In both Figures, only the positive (i.e., "repulsive") feedback, i.e., a force away from the delayed particle position is found to yield negative values of C, i.e., anticorrelations.*

1.5 Nonlinear example systems with time delay

In more realistic scenarios, the purely linear case is often of limited use. An important step towards applications is thus the incorporation of *nonlinearities* in the description, which are essentially omnipresent in real-world systems, and are, at the same time, very poorly understood in the context of time-delayed, stochastic systems. For this purpose, we will consider two nonlinear toy models, which are already well-established in the context of diffusion and transport of Markovian, Brownian particles. In this Section, we will introduce these models; specifically, the involved bistable and periodic static potentials.

Due to their prototypical character, the here introduced exemplary systems are suitable to investigate the effect on time delay in quite generic nonlinear equations. In Parts II and III, they will serve as toy models for the theoretical considerations, and be used as applications for the proposed approaches.

1.5.1 Bistable system: The doublewell potential

As our prime example system, we will study a particle in a *doublewell potential*

$$V_{\mathrm{s}}(x) = V_0\,[(x/\sigma)^4 - 2\,(x/\sigma)^2]. \tag{1.52}$$

This potential has two minima at $x = \pm\sigma$, separated by a barrier of height V_0. By the quartic potential walls, the particle is confined, which means the chances of finding the particle decay to zero for $x \to \pm\infty$. Typically, the probability will be highest in the vicinity of one of the two minima, i.e., within the center of one *well* or *valley*. However, due to the thermal motion, the particle can also leave its valley and jump to the other one. The distance between barrier and minima introduces a length scale, σ. We will thus measure lengths in units of σ.

A Brownian particle in such a doublewell potential is indeed a very well-known, generic model for bistable, noisy dynamics. For example, it is commonly used to study escape problems [158, 173, 254, 255]. This is further used to model reaction kinetics, the particle representing a reaction coordinate [255]. Also individual memory cells of memory devices may be modeled in this way [153] (e.g., the particle being on the left or right side may correspond to 0 and 1, respectively). More generally, this model may describe any continuous system, that randomly switches between two characteristic states. Experimentally, such a bistable potential landscape can be generated by optical traps [150].

We will consider the combined impact of this nonlinear potential and (linear) delay force stemming from a quadratic trap (1.44), giving rise to the LE

$$\gamma\dot{X}(t) = (4V_0/\sigma)[-(X(t)/\sigma)^3 + (X(t)/\sigma)] - k[X(t) - X(t-\tau)] + \sqrt{2D_0}\xi(t). \tag{1.53}$$

One can equivalently consider the forces to stem from a *total potential*

$$V_{\mathrm{s}}(x) + V_{\mathrm{d}}(x, x_\tau), \tag{1.54}$$

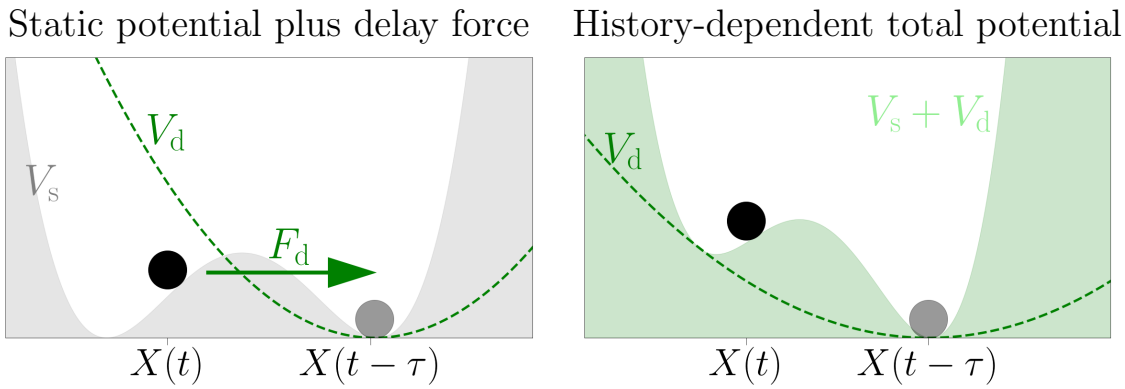

Figure 1.12: *Two ways to visualize the bistable system with time delay from Eq. (1.53). Left panel: Static doublewell potential V_s (1.52) supplemented by a linear delay force F_d defined in (1.44). Right panel: Total potential $V_\mathrm{s}(x)+V_\mathrm{d}(x,x_\tau)$ (1.54). Due to the dependency on the delayed position, the total potential is history-dependent.*

with $V_\mathrm{d} = (k/2)[X(t)-X(t-\tau)]^2$ from Eq. (1.45). Due to the delay, the total potential is a *history-dependent* energy landscape. Figure 1.12 shows a schematic illustration of the system (1.53) (from both viewpoints).

In recent years, bistable, noisy systems have indeed already been considered under the impact of a linear delay force by different authors [83, 256–260]. Experimental realizations have been suggested in Refs. [259, 260], where X represents the polarization dynamics of lasers with optical feedback (and τ is associated with the external cavity length). However, these studies are rather concerned with dynamical phenomena, and thus have a different focus then the work presented in this thesis, which mainly aims at a probabilistic description and thermodynamic notions of delayed systems. We will refer to earlier results in the respective sections at various occasions in Parts II and III.

1.5.2 Periodic system: The washboard potential

Second, we will consider the periodic *washboard potential*

$$V_\mathrm{s}(x) = -(V_0/2)\cos(x/\sigma) \tag{1.55}$$

with infinitely many barriers of height V_0 and minima at $x = 2\pi\sigma i$ and $i \in \mathbb{Z}$. Again, σ represents a typical length scale and is used as a unit for distances.

Particles in sinusoidal potentials (in the absence of delay forces) represent a well-studied paradigm to model transport on rough surfaces [230, 261, 262], and can be realized experimentally, e.g., by optical traps [230]. More generally, this gives a simple model for spatially extended (not confined) systems with a structured environment. Analogously to the bistable example, we will consider the motion of a particle subject to delayed linear feedback in the static potential, described by

$$\gamma\dot{X}(t) = -(V_0/2)\sin[X(t)/\sigma] - k[X(t)-X(t-\tau)] + \sqrt{2D_0}\,\xi(t), \tag{1.56}$$

A schematic plot of the system is shown in Fig. 1.13.

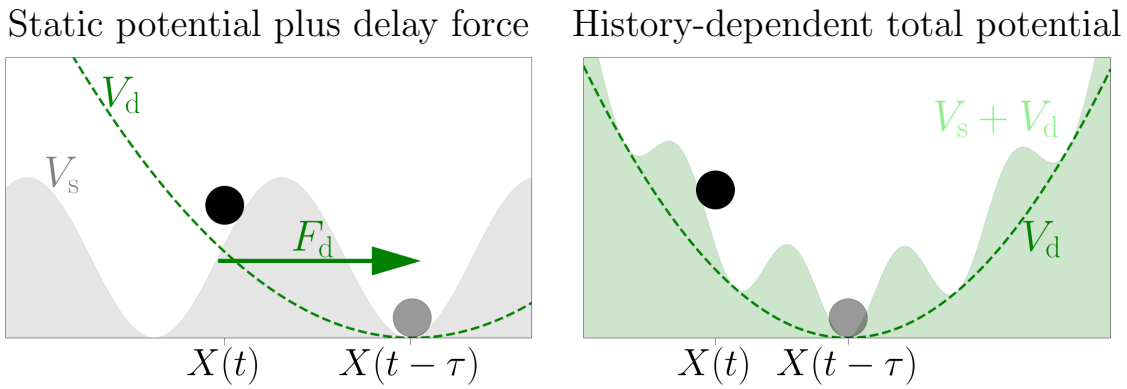

Figure 1.13: *Two ways to visualize the periodic system with time delay from Eq. (1.56). Left panel: Static doublewell potential V_s (1.55) supplemented by a linear delay force F_d defined in (1.44). Right panel: Total potential $V_\mathrm{s}(x) + V_\mathrm{d}(x, x_\tau)$ (1.54). Due to the dependency on the delayed position, the total potential is history-dependent.*

Exact steady-state solution, normalization of periodic system In contrast to the linear trap and the doublewell, this potential is not confining, thus, the probability to find the particle does not quickly vanish for $|x| \gg 0$. Therefore, we have to be a bit more careful in regard to the normalization of the probability densities, on the one hand, and the numerical simulations, on the other hand.

In the absence of external driving (which implies $k\tau = 0$), this Markovian system naturally approaches a steady state in the long time limit, which, in this case, is the thermal equilibrium [125], (see Chapter 3). The PDF is then simply the Boltzmann-distribution

$$\rho_\mathrm{ss}(x) = \frac{1}{Z}\exp\left[-\frac{V_\mathrm{s}(x)}{\gamma D_0}\right] = \frac{1}{Z}\exp\left[\frac{V_0\cos(x/\sigma)}{\gamma D_0}\right]. \tag{1.57}$$

However, due to the system's infinite extent, the normalization constant $Z = \int_\Omega \exp\left[-V_\mathrm{s}(x)/(\gamma D_0)\right]$ is somewhat problematic. Contrary to the other examples, we will hence consider a PDF[23] on a reduced interval $[-\pi, \pi]$ with normalization $\int_{-\pi}^{\pi} \hat{\rho}(x)\mathrm{d}x = 1$. Please note that this does not imply periodic boundary conditions at $\pm\pi$, which would not be suitable in the case where the delay force is present. While this reduced viewpoint on the system is appropriate for the steady state and one-time PDFs, it is less useful to consider transient dynamics, or higher order correlations.

The PDF (1.57) is shown in Fig. 1.14, which also displays the results of a Brownian dynamics simulation (see Appendix A.1). In the simulations, uniformly distributed initial conditions were used. The agreement between simulation results and exact solution confirms our numerical method.

To simulate the system with washboard potential, we have implemented a comoving simulation box, and whenever moving the box, we have made sure that the delayed particle position is transformed accordingly. This is particularly important to simulate the long-time diffusive behavior, which we later briefly discuss in Sec. 6.5.5.

[23] This reduced PDF already contains the full information due to the spatial 2π-periodicity of the system.

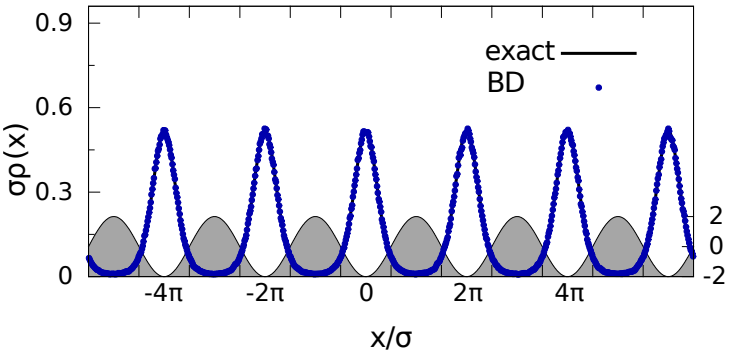

Figure 1.14: *The steady-state PDF in the washboard potential (1.57) with barrier height $V_0 = 2k_{\mathrm{B}}\mathcal{T}/\sigma^2$. Black line: exact results (1.57), blue disks: numerical results from BD simulations. Parameters: $N = 9\times10^5$, $t = 150\tau_{\mathrm{B}}$.*

1.5.3 Scaling

In principle, the LE could be reduced by one parameter, without loss of information, by rescaling the entire equation. For example, the particular value of each energy-term (like barrier height) does not matter in detail for the dynamics, but only the relation between them (like barrier height vs. thermal energy). However, by scaling one parameter out (like the barrier height), this parameter can then afterwards not be changed anymore (without changing all other parameters). As we will be interested in the varying different parameters in different Chapters of this thesis, we would have to use alternative rescalings. For the sake of more consistency, we therefore choose not to rescale the equation, and instead consider the equation with all physical coefficients, like friction γ or diffusion D_0, and length scale σ. This might as well make comparison with experimental data easier.

We are not going to choose specific values for the parameters in SI units, like $V_0 = \ldots\ \mathrm{kg\,m^2/s^2}$, but rather express the values by their ratios (for example, we will say the barrier height is three times larger then the thermal energy, i.e., $V_0 = 3k_{\mathrm{B}}\mathcal{T}$), making use of this scalability. We nevertheless give the corresponding SI units of all coefficients on page IX. To find suitable real values, we refer the interested Reader to the cited experimental works, like [16, 263], and to the book [162].

We emphasize that the theoretical frameworks discussed in this thesis do not only apply to Brownian particles, but also to other natural and artificial systems. Having this in mind, the friction and diffusion coefficient, the Brownian timescale and the particle diameter introduced before just provide appropriate energy, time- and length scales, which one can adjust to the specific system under consideration.

Further, if one has a more mathematical viewpoint on the problem and is less interested in the physical context, it might be more convenient to consider corresponding dimensionless LEs, which can be obtained by rescaling[24] position $x^* = x/\sigma$ and time $t^* = t/\tau_{\mathrm{B}}$, $\tau^* = \tau/\tau_{\mathrm{B}}$, and by rescaling $V_0^* = V_0/(\gamma D_0)$ and $k^* = k(\tau_{\mathrm{B}}/\gamma)$.

[24] A quite popular alternative is to simply give all parameters dimensionless numbers. This could be done at any time for the following results.

1.6 Timescales

The stochastic motion is characterized by various characteristic timescales. We have already introduced some of them in earlier Sections of this Chapter 1. Let us still, at the cost of redundancy, give here a brief overview over all main timescales, from the shortest to the longest. This provides an interesting view on the investigated motion. First, we will consider timescales known for Markovian systems. We recall that lists of all abbreviations and symbols can be found on pages VIII and IX, respectively.

Ballistic motion On very short times $0 \leq t \leq m/\gamma$, the motion is *ballistic*, i.e., dominated by inertia, resulting in a quadratic increase of the MSD $\sim t^2$. However, this aspect of the dynamics on this timescale is not resolved by the overdamped equation considered in this thesis (which neglects the second temporal derivative in the LE, see Sec. 1.2.3). The overdamped limit is appropriate when the particle mass m is small compared to the friction γ, such that the timescale m/γ is negligible (as compared to, e.g., τ_B).

Hydrodynamic memory Further, in our colloidal example system, hydrodynamic effects may be important above the ballistic timescale. In particular, the motion of the colloid affects the flows in the surrounding fluid, which, at a later time then, in turn, impacts the particle position. This *hydrodynamic backflow* gives rise to negative position-autocorrelation functions at intermediate times. However, this effect is typically neglected in the description of single colloidal particles, giving rise to a Markovian LE with white noise and instantaneous friction (also see Sec. 1.3.1). We do not further discuss the impact of hydrodynamics within the scope of the thesis.

Diffusion At yet larger times, the dynamics is dominated by *diffusion*. Here, a free particle has a linear MSD $\sim t$. A characteristic scale is the *Brownian time* $\tau_\mathrm{B} = \sigma^2/D_0$ describing the typical time it takes for a free (in the absence of deterministic forces) particle to diffuse over a distance σ. Except for the purely linear system, which has no typical length scale (which is the reason why it is convenient to have the somewhat abstract length scale σ), the minima and maxima of the considered potentials have a distance of order σ.

Furthermore, any external potential will influence the dynamics, yielding additional timescales.

Intrawell relaxation For example, a (quadratic) trapping potential $V_\mathrm{s} = (a/2)x^2$ confines the motion. This yields a saturation of the mean square displacement with a characteristic saturation time, called *intrawell relaxation time* τ_ir, which can be estimated by γ/a (see the description of the Ornstein-Uhlenbeck process in Sec. 1.2.3, and see Sec. 6.5.5 later). More generally, a potential well does not necessarily fully confine the particle, but can also allow for escapes out of the

valley (as we discuss next). Still, the MSD will show a saturation on the intrawell relaxation time (before the escapes occur). During this time, the particle position has "relaxed" within the well, meaning, initial condition effects have decayed (and the PDF is given by the Boltzmann distribution). Employing the estimation explained and derived in Ref. [158], p. 348, one can use $\tau_{\mathrm{ir}}/\tau_{\mathrm{B}} \approx \gamma/V_{\mathrm{s}}''(x_{\mathrm{min}})$, yielding $\tau_{\mathrm{ir}}/\tau_{\mathrm{B}} \approx 1/8\,(k_{\mathrm{B}}\mathcal{T}/V_0)$ for doublewell potential (1.52) at $k = 0$.

When the considered static potential has multiple (local) minima, *escape events* can occur, i.e., the particle spontaneously leaves its current valley (due to random kicks from the heat baths). These escapes are characterized by further typical timescales.

Jump duration time First, the time that the particle actually on average needs to cross the separating barrier, that is, after it passed one minimum for the last time, until it reaches the next minimum for the first time. The average *jump duration time* depends on the shape of the potential, and, of course, on the barrier height as compared to the thermal energy of the particle. In this thesis, we will not study this timescale in detail, but refer the interested Reader to Ref. [71], where it was thoroughly discussed for Markovian systems. Instead, we will here only use numerical estimations[25]. This timescale is found to be important for a delay-induced oscillatory state, as we will describe later.

Kramers escape time Furthermore, the typical time that passes between two jump events occur is the famous *Kramers escape time* τ_{K}. For Markovian processes, this timescale is described by the Kramers theory and the Arrhenius formula. Due to the importance and prominence of the Kramers escape time, we will dedicate a little section to it in the following. (An explicit example for the impact of the escape time on the MSD will be given later in Sec. 6.5.5.)

Time delay Of course, the *time delay* itself introduces an additional timescale into the system, namely τ. Only by thinking about at the well-known timescales mentioned above, it is already clear that we shall expect different delay-induced dynamical phenomena, depending on the value of τ as compared to all these timescales. We will come back to this thought multiple times in the remainder of this thesis. A prominent example are the delay-induced oscillations in the linear system, see above Sec. 1.4.5.

Moreover, the presence of a delay force may also shift the other timescales. However, we will often still use "Markovian estimates" for the timescales, which are, indeed, nevertheless found to be appropriate in many cases.

[25] We will employ the minimum-based technique of Ref. [71]: Basically, one numerically detects the times it takes for the particle to sequentially pass the three positions: one minimum – the maximum – the adjacent minimum, and then average over the results.

1.6.1 Kramers escape times

For the Markovian case ($k\tau = 0$), the total waiting time between two jumps Δt_{jump} is described by the Kramers theory [173, 254, 255]. This framework predicts exponentially decaying waiting time probabilities, with an average value, called *mean escape time* given by Arrhenius formula [given by Eq. (1.58)], and relies on several assumptions and simplifications, which we briefly lay out in the following.

The Kramers theory, which was developed for *Markovian* systems characterized by a *static* potential landscape U. It gives an approximation for the mean escape time τ_{K} over a barrier of height ΔU induced, or as one sometimes says "activated", by white noise. Importantly, the potential barriers ΔU shall be large compared to the thermal energy, i.e., $\gamma D_0 \ll \Delta U$. This enables a quasistatic assumption of the dynamics within the well. Further, around both extrema (the well and the barrier), the potential shall be well-approximated by a quadratic potential[26]. The Kramers theory provides an estimation for the escape rate by the Arrhenius formula, which reads [1, 71, 158, 254–256]

$$r_{\text{K}} = \sqrt{U''(x_{\min})|U''(x_{\max})|}/(2\pi\gamma)\, e^{-\Delta U/\gamma D_0}, \tag{1.58}$$

with the corresponding potential minimum $x_{\min}$ and maximum, $x_{\max}$. There are many further results known for this and similar problems, see, e.g., [255] for a detailed overview.

The estimated escape time from the rate, depends on the number of ways to exit the potential well. In our one-dimensional models, this can either be one, as in the case of the bistable potential, or two, as for the period system, yielding $\tau_{\text{K}} = 1/r_{\text{K}}$ or $\tau_{\text{K}} = 1/(2r_{\text{K}})$, respectively.

What happens in the presence of time delay? We stress that the estimate (1.58) relies on a quasistatic approximation for ρ_1 and does not involve ρ_2, or any higher order correlation between different times and spaces. Applying it to the *non-Markovian* system, as we will do in Parts II and III, must be considered a crude approximation. There are actually generalizations towards non-Markovian systems, for example [255, 264]. An interesting comparison between the Kramers predictions and alternative, more sophisticated approximations to the actual escape times for a system with time delay were reported in [89]. Indeed, this is a research topic on its own and not the focus of the present thesis. We will here instead only employ Eq. (1.58). The main goal of the crude approximation presented in Part II, is rather to check how far we can get with the approach for the PDF presented there.

[26] Generalizations of the Kramers estimate have been worked out, which take higher order terms of a expansion around the extrema into account and can be used if this assumption is not appropriate for the potential at hand, see e.g. [173].

1.7 Delay-induced oscillations & Coherence resonance in nonlinear systems

In this Section, we will discuss an important dynamical state present in our nonlinear example applications, that are spontaneous, delay-induced oscillations between different potential wells. As it has already been realized in earlier literature (for similar models) [40, 256, 257, 263, 265], these oscillations are characterized by resonance between noise and delay. We introduce this state here, because it will serve in Parts II and III as a particular interesting state to test our theoretical approaches.

1.7.1 Delay-induced oscillations

The delay-induced stochastic oscillations of the particle position, which we have encountered in the linear system with time-delayed feedback (see Sec. 1.4.6), can also be observed in the nonlinear systems. Despite many similarities, the oscillatory behavior is be more rich, due to the additional timescales and the more complex environments. A particularly important difference is that in the nonlinear systems, the oscillations can occur within one potential well, but they can also be *between different potential valleys* (e.g., depending on the barrier heights). This is the dynamical state we will be most interested in throughout the thesis.

Let us first discuss universal features. The aforementioned association between *positive / negative feedback induces oscillations with mean period* $\sim 2\tau$ */* $\sim \tau$*,* is also found to be characteristic in the nonlinear example system [1, 2, 40, 256, 257, 266], for intrawell and interwell oscillations. Further, the oscillations are as well not transient (as typical for DDEs), but manifest in the autocorrelation function of the steady state. They have again stochastic character. The interwell oscillations are a bit easier to see in the trajectories. They randomly set in and pause, and have varying, random period.

Figure 1.15 shows as an example the position-autocorrelation function in the doublewell potential without and with negative feedback. In the latter case, the correlation function has pronounced peaks at multiples of τ. Here, the particle position oscillated between the two valleys, and are thus much clearer visible as in the linear system. The delay-induced oscillations are subsequent escape events with a waiting time between two escape events around $\tau/2$ (or τ in the case of positive feedback). We further observe that the oscillations randomly set in and pause. The oscillations also manifest in the waiting time distributions, as we will later show in Sec. 6.5.4.

We recall that the purely linear, time-delayed (overdamped) system is characterized by two relevant timescales, the relaxation within the well, and the time-delay. The oscillations are pronounced, if the delay is larger than the relaxation time. This is also true for intrawell oscillations in the nonlinear systems. Similarly, particle oscillations between neighboring wells set in, if the delay is large compared to the jump duration times (see above Sec. 1.6). These are delay-induced interwell oscillations. For the latter, yet another timescale is important, that is, the Kramers escape time. This additional timescale is indeed connected to an intricate dynamical

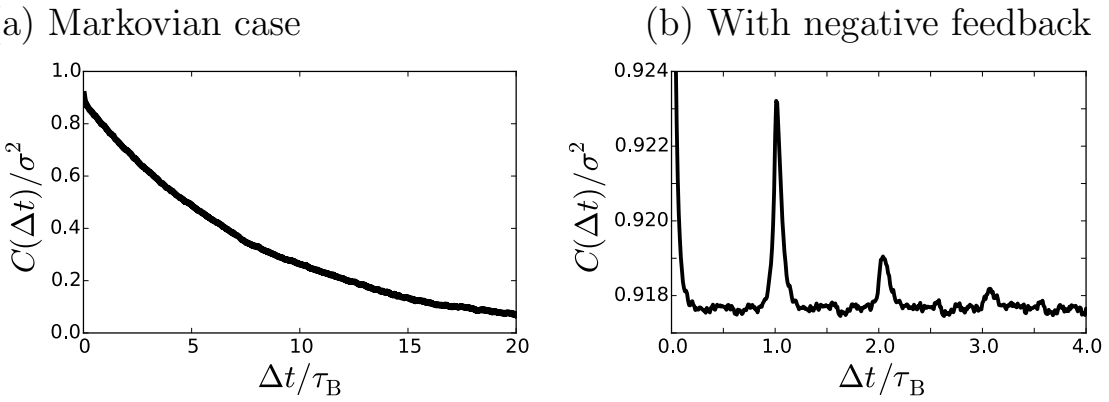

Figure 1.15: *The delay-induced oscillations manifest in the position-autocorrelation function $C(\Delta t) = \langle X(t)X(t+\Delta t)\rangle$. Numerically calculated C in the doublewell with barrier height $V_0 = 4k_B\mathcal{T}$. (a) No delay force, (b) under negative feedback with delay force strength $k = 12k_B\mathcal{T}/\sigma^2$ and delay time $\tau = \tau_B$.*

phenomenon, called *coherence resonance*.

1.7.2 Coherence resonance

If τ is large compared to the jump duration time, delay-induced interwell oscillations of the particle position are visible (e.g., in the position-autocorrelation function). We consider parameter regimes, where in the absence of noise, no (transient) intrawell escape events could be induced by the delay force (i.e., k is somewhat small compared to V_0). For a very low temperature of the system, the oscillations are random events that often stop after two or three oscillation periods. Upon increasing the temperature, the Kramers escape time is drastically reduced, as can be seen from Eq. (1.58). When the Kramers escape time is in the regime of τ, the interwell oscillations are particularly pronounced and regular. For even larger temperature, the motion is dominated by the noise. The particle's thermal energy is so high that it does not really "see" the barrier anymore. Then, the regularity of the jumps decays. The resonant response of the delayed system to the noise for intermediate, finite temperatures is called *coherence resonance* (CR). Importantly, in contrast to *stochastic resonance*, there is no periodic external driving.

These oscillations and CR have already been studied previously. Experimentally confirmations of CR in delayed, bistable models were reported in Refs. [40, 257, 263]. A particularly important theoretical contribution is [256], where an approximate expression for the corresponding position-autocorrelation function was derived, based on a discretization of the dynamics. We will elucidate this 2-state reduction in more detail and employ the result from [256] in Chapter 8. However, a full characterization of the dynamics, in particular, regarding the specific systems considered here, was not presented in earlier literature. This motivates us to further elucidate this dynamical state. We will characterize the oscillations and CR more elaborately in Parts II and III of this thesis, particularly in Chapter 8. To analyze these phenom-

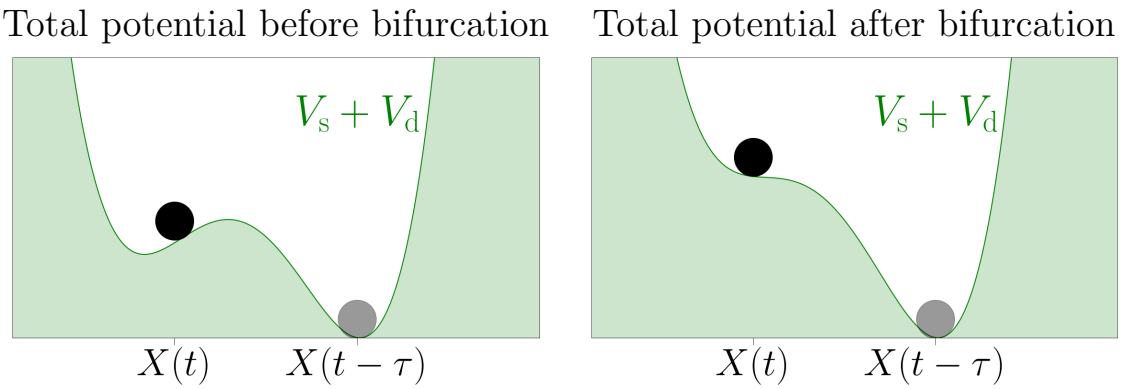

Figure 1.16: *Schematic plots of the history-dependent total potential* $V_s[X(t)] + V_d[X(t), X(t-\tau)]$ *for two different values of* $k\sigma^2/V_0$*. Here* $X(t-\tau) = \sigma$ *is at one potential minimum. Left:* $k < V_0/2\sigma^2$ *is slightly before the bifurcation, where the total potential still has two minima, right:* $k > V_0/2\sigma^2$ *slightly after the bifurcation, where the second minimum has disappeared. Another visualisation of the total potential (1.54) is provided in Fig. 1.12.*

ena, we will consider the position-autocorrelation function (e.g., see Fig. 8.7), as well as waiting time distributions (e.g., Figs. 6.10 and 8.6). Furthermore, we will study how thermodynamic quantities behave around CR (see Chapter 8).

Throughout this thesis, our analysis of CR is focused on the bistable model. However, our numerical investigations indeed suggest that the periodic model as well exhibits CR.

1.7.3 Bifurcation theoretical perspective on delay-induced oscillations

Let us briefly revisit the delay-induced interwell oscillations from the viewpoint of *bifurcation theory* [196]. In general, bifurcation scenarios are best understood for deterministic systems. Thus, we here start with the corresponding deterministic, i.e, noise-free LE ($\mathcal{T} = 0$). Furthermore, for the sake of simplicity, we here focus on the bistable model, and $k > 0$. The idea can be generalized to the washboard potential, and to $k < 0$. We recall that one can define a total potential (1.54) consisting of the static doublewell potential plus the harmonic trap around the delayed position, which is then a history-dependent potential.

Using the terminology of nonlinear dynamics (see, e.g., [196]), the bistable system with linear delay force (and not too small τ) has a bifurcation upon the increase of k. In particular, for any k, the system has two solutions $\{X(t) = \pm\sigma, X(t-\tau) = X(t)\}$, corresponding to $X(t)$ and $X(t-\tau)$ being in the same of the two minima of V_s. These two solutions are always stable. Further, if k reaches a certain critical value (which is around $k \approx V_0/2\sigma^2$), a new stable solution in the form of a *limit cycle* bifurcates. This limit cycle corresponds to X-oscillations with period $\sim \tau$ between both sides of V_s (then, $X(t)$ and $X(t-\tau)$ are in opposite minima of V_s, switching back and forth). The initial conditions alone decide which solution is visible. This bifurcation occurs at precisely the value where, when $X(t)$ and $X(t-\tau)$ are in *opposite* minima

of $V_{\rm s}$, i.e., $\{X(t) = \pm\sigma, X(t-\tau) = -X(t)\}$, the *total* potential $V_{\rm s}(x) + V_{\rm d}(x, x_\tau)$ has only one unique minimum. (For smaller values of k, the total potential has two minima.) A plot of the history-dependent total potential before (left panel) and after the bifurcation (right panel) is shown in Fig. 1.16. However, the whole bifurcation scenario is actually not simple due to the infinite-dimensional nature of the system (x and x_τ are not two independent phase space variables, as we have pretended in this description of the bifurcation).

When we quantify coherence resonance (specifically in Chapter 8), we will focus on parameter regimes, where no delay-induced oscillations would occur, in the absence of noise, i.e., at k values *before* that bifurcation (left panel in Fig. 1.16). Thus, the described oscillations are *delay- and noise-induced*. One can take the perspective that the noise makes this limit cycle "visible" before it actually bifurcates.

2 | Fokker-Planck equations

In the preceding Chapter, we have introduced the Langevin equation, which describes the random processes studied in this thesis on a stochastic level. For Markovian systems, it is well-known that Fokker-Planck equations (FPE) provide a complementary way of description, on the *probabilistic level*. These are deterministic equations, whose solutions are the corresponding probability density functions. In the following, we will briefly introduce this concept, first focusing on the Markovian case. Towards the end of this Chapter, we will review some previous results from the literature on Fokker-Planck (FP) descriptions for systems with time delay. We will keep this discussion short, because the entire Part II of this thesis is dedicated to the probabilistic treatment of delay systems.

2.1 Markovian case

There is a unique FPE for each Markovian LE of type (1.16), where F might depend upon time and the system state X [173], describing the temporal evolution of the probability densities. We have introduced the concept of probability density functions in Chapter 1 [see Eq. (1.7)]. Recall that $\int_y^z \rho_1(x,t)\,\mathrm{d}x$ gives the probability to find the system in a state $x \in [y,z]$ at time t. The FPE for the one-time PDF reads [158, 173]

$$\frac{\partial}{\partial t}\rho_1(x,t) = -\frac{\partial}{\partial x}\left[\frac{1}{\gamma}F(x,t)\rho_1(x,t)\right] + D_0\,\frac{\partial^2}{\partial x^2}\rho_1(x,t). \tag{2.1}$$

Given a normalized initial condition $\rho(x,t=0)$ with $\int_\Omega \rho(x,t=0)\mathrm{d}x = 1$, this equation gives the PDF at all $t > 0$ (assuming that the system is stable). The first term on the r.h.s is the *drift term* describing how the PDF shifts in time due to external forces. The second term on the r.h.s describes the diffusive spreading due to the white noise, and is called *diffusion term*. To illustrate the impact of the latter, we consider a Wiener process performed by the motion of a particle with zero drift ($F = 0$), and a delta-distributed initial condition [the particle is initially fixed at $X(0) = x_0$]. This yields [173] (p. 99)

$$\rho_1(x,t) = \frac{1}{\sqrt{4\pi D_0 t}} e^{-\frac{(x-x_0)^2}{4D_0 t}}. \tag{2.2}$$

Thus, the PDF is a zero-mean[1] Gaussian distribution, reflecting the Gaussianity of the white noise and the linearity of the underlying LE. Furthermore, the variance increases linearly by $\langle [X(t) - x_0]^2 \rangle = 2D_0 t$, consistent with the mean square displacement calculated from the LE, see Eq. (1.18).

2.1.1 Natural boundary conditions

Throughout this thesis, we will consider natural boundary conditions, i.e.,

$$\lim_{x\to\pm\infty} \rho_1(x,t) = 0, \qquad \lim_{x\to\pm\infty} \partial_x \rho_1(x,t) = 0, \tag{2.3}$$

for $\Omega = \mathbb{R}$. Thus, the PDF and its spatial derivative both decay for large $|x|$. This boundary condition is appropriate in many scenarios, where the overall extent of the system is large compared to the length scales of interest.

Here and in the following, we occasionally use ∂_t, ∂_x, and ∂_{xx}, as short notations for $\frac{\partial}{\partial t}$, $\frac{\partial}{\partial x}$, and $\frac{\partial^2}{\partial x^2}$, respectively. We will further denote improper integrals $\lim_{r\to\infty} \int_{-r}^{r}$ simply as $\int_{-\infty}^{\infty}$.

2.1.2 Joint probability densities

Similar to (2.1), one can also find FPEs for *multiple*-time PDFs. As we will later see, they play a particularly important role in the description of non-Markovian processes. Analogously to the one-time PDF [from Eq. (1.7)], one can also define, based on the stochastic process $X(s)$, $s \in [t_0, t_n]$, n-time (joint) PDFs (for any $n \in \mathbb{N}$), as

$$\begin{aligned}\rho_n(x^{(n)}, t_n; x^{(n-1)}, t_{n-1}; ...; x^{(1)}, t_1) \\ = \langle \delta[X(t_n) - x^{(n)}] \delta[X(t_{n-1}) - x^{(n-1)}] ... \delta[X(t_1) - x^{(1)}] \rangle.\end{aligned} \tag{2.4}$$

This gives the *joint* PDF for multiple random *events* $\left\{x^{(i)}, t_i\right\}$ at different times t_i, with $i \in \{1, n\}$. For example,

$$\int_x^{y+\Delta y} \int_{y'}^{y'+\Delta y'} \int_y^{y+\Delta y} \rho_3 \left(x^{(3)}, t_3; x^{(2)}, t_2; x^{(1)}, t_1\right) \mathrm{d}x^{(3)} \mathrm{d}x^{(2)} \mathrm{d}x^{(1)}$$

gives the joint probability to find the particle at times t_1 and t_3 within the interval $[y, y + \Delta y]$ *and* at time t_2 in $[y', y' + \Delta y']$. Since the overall probability to find the particle somewhere in Ω at all three times must be one, the normalization is $\iiint_\Omega \rho_n \mathrm{d}x^{(1)} ... \mathrm{d}x^{(n)} = 1$.

Marginalizing the joint n-time PDF w.r.t. one variable, yields the respective $(n-1)-$time PDF. For instance,

$$\int_\Omega \rho_2 \left(x^{(2)}, t_2; x^{(1)}, t_1\right) \mathrm{d}x^{(2)} = \rho_1 \left(x^{(1)}, t_1\right). \tag{2.5}$$

[1] this is due to the spatial symmetry of the system.

A further connection between joint and marginal PDF is given by the *conditional* PDF, ρ_c, defined as

$$\rho_2\left(x^{(2)},t_2;x^{(1)},t_1\right)=\rho_c\left(x^{(2)};t_2\middle|x^{(1)},t_1\right)\rho_1\left(x^{(1)},t_1\right), \tag{2.6}$$

here again using the case $n=2$ as an example. $\int_y^{y+\Delta y}\rho_c\left(x^{(2)};t_2\middle|x^{(1)},t_1\right)\mathrm{d}x^{(2)}$ tells us how likely it is the find the particle at time t_2 around y, given that it was at position $x^{(1)}$ at time t_1.

For Markovian processes it is, in fact, not necessary to consider n-time PDFs with $n>2$. This is because of the following simplification. For the three-time PDF with $t_3>t_2>t_1$, *Markovianity* implies

$$\rho_c\left(x^{(3)},t_3\middle|x^{(2)},t_2;x^{(1)},t_1\right)=\rho_c\left(x^{(3)},t_3\middle|x^{(2)},t_2\right). \tag{2.7}$$

The last expression represents a common *definition of Markovianity*. In the present context, Eq. (2.7) further implies

$$\rho_3\left(x^{(3)},t_3;x^{(2)},t_2;x^{(1)},t_1\right)=\rho_c\left(x^{(3)},t_3\middle|x^{(2)},t_2\right)\rho_c\left(x^{(2)},t_2\middle||x^{(1)},t_1\right)\rho_1\left(x^{(1)},t_1\right). \tag{2.8}$$

In the same way, the four-time PDF can be expressed using the conditional PDFs, and so on.

Indeed, the conditional PDF $\rho_c(x,t|x',t')$, which is also called "transition probability", provides the full information about the probability evolution of a Markovian system[2], in contrast to the one-time PDF alone. As a matter of fact, the conditional PDF $\rho_c(x^{(1)},t^{(1)}|x^{(2)},t^{(2)})$ is just the one-time PDF $\rho_1(x^{(1)},t^{(1)})$ from the FPE (2.1) given the special initial condition $\rho_1(x,t^{(2)})=\delta(x-x^{(2)})$.

2.1.3 The probability current and steady states

Now we will introduce another important probabilistic quantity, the probability current. This will also bring us to the notion of equilibrium on the level of the Markovian FPE.

As the total probability to find the particle is a conserved quantity, the FPE (2.1) can be rewritten in the form of a continuity equation

$$\partial_t\rho_1(x,t)=-\partial_x J(x,t), \tag{2.9}$$

with the *probability current*

$$J(x,t)=\left[\gamma^{-1}F(x,t)-D_0\,\partial_x\right]\rho_1(x,t). \tag{2.10}$$

[2]Which is as well immediately apparent in the so called Chalpman-Kolmogorov equation [173] $\rho_c(x^{(3)};t_3|x^{(1)},t_1)=\int_\Omega\rho_c(x^{(3)};t_3|y^{(2)},t_2)\rho_c(y^{(2)};t_2|x^{(1)},t_1)\mathrm{d}y^{(2)}$.

We recall that steady states are defined by $\partial_t \rho_{1,\mathrm{ss}}(x,t) = 0$, which means the l.h.s. of the FPE is naturally zero. As immediately follows from Eq. (2.9), the probability current is then automatically constant in x and t. Steady states can be distinguished between two cases: $J = 0$, and $J \neq 0$, which are, for Markovian systems, associated with EQ and NESS, respectively [267]

$$J = 0 \Leftrightarrow \mathrm{EQ} \tag{2.11}$$

$$J = \text{const.} \neq 0 \Leftrightarrow \mathrm{NESS}. \tag{2.12}$$

To better understand this connection, it is interesting to look at the *particle current* $\langle \dot{X} \rangle$, which is closely related to J via $\langle \dot{X} \rangle = -\int_\Omega J(x,t)\mathrm{d}x$ [267]. Thus, for Markovian systems in steady states, a nonzero probability current implies net *particle transport*, which is clearly a nonequilibrium property.

As a first example, we consider a situation where all external forces stem from a potential $F = -V_\mathrm{s}'$. Then, the system approaches a steady state, where the one-time PDF is given by the *Boltzmann distribution* [173]

$$\rho_{1,\mathrm{ss}}(x) = Z^{-1}\, e^{-V_\mathrm{s}(x)/(\gamma D_0)} \tag{2.13}$$

with normalization constant Z, as readily follows from the FPE (2.9). Furthermore, plugging the Boltzmann distribution (2.13) into Eq. (2.10) yields a vanishing probability current, implying zero particle transport and thermal equilibrium.

Second, a prototypical example for a one-dimensional, Markovian system in a NESS is a colloidal particle on a ring (i.e., a finite spatial domain with periodic boundaries) that is driven by a constant force F in one direction of the ring. Here, for symmetry-reasons $\rho_{1,\mathrm{ss}}$ is constant in x, which implies $\partial_x \rho_1 = 0$ and further $J \propto F$ [from Eq. (2.10)]. Thus, due to the preferred direction of motion aligned with the external drive F, this system exhibits a finite probability and particle current (with the same sign as F). Also see Ref. [125].

Multidimensional case We have used above a notation indicating *one*-dimensionality of the process, where the different random events are denoted $\{x^{(i)}, t_i\}$ with an index i referring to the specific event, i.e., instance in time $t = t_i$ and one-dimensional value of $X = x^{(i)}$. However, the formulae can as well be applied to multidimensional systems, where the system state itself is multi-dimensional, for example by replacing X by a vector $\underline{X} = \{X_0, X_1, ..., X_n\}$. This could be used to describe Brownian motion in three-spatial dimensions (then $n = 2$). Then, every stochastic event is a multi-dimensional object as well, e.g., $\left\{\underline{x}^{(n)}, t_n\right\} = \left\{\{x_0^{(i)}, x_1^{(i)}, x_2^{(i)}\}, t_i\right\}$.

The Markovian $(n+1)$-variate FPE reads

$$\partial_t \rho_{n+1}(\underline{x}, t) = -\sum_{j=0}^{n} \partial_{x_j} \underbrace{\left[\frac{F_j}{\gamma_j} - \frac{k_\mathrm{B}\mathcal{T}_j}{\gamma_j}\,\partial_{x_j}\right]\rho_{n+1}}_{=J_j(\underline{x},t)}, \tag{2.14}$$

with the $(n+1)$-point *one*-time PDF $\rho_{n+1}(\{x_0, x_1, ..., x_n\}, t)$, and the various $(n+1)$-dimensional probability currents $J_j(x_0, x_1, ..., x_n, t)$. The natural boundary conditions amount to $\lim_{x_i \to \pm\infty} \rho_1(x_0, x_1, ..., x_n, t) = 0 \; \forall i \in \{0, 1, .., n\}$. We will consider multi-variate FPEs various times in Parts II and III.

2.2 Introduction to Fokker-Planck descriptions of systems with time delay

Above, we have seen that for Markovian systems, the full probabilistic information is encoded in the transition probabilities. Furthermore, the probability current (2.10) is a sufficient criterion to distinguish thermal equilibrium from nonequilibrium steady states. Both, the transition probabilities and the probability current are given by the FPE for the one-time PDF, which is a closed equation (2.1). Now we turn to the systems with *time delay*, where the situation completely changes, and is found to be literally "infinitely more" complicated. In this Chapter, we review some literature results and hint to open problems. Later, in Part II, we present new results for the probabilistic treatment of systems with delay.

The FPE for the one-time PDF for the delay LE (1.42) was introduced in [82, 85], and reads

$$\frac{\partial}{\partial t}\rho_1(x,t|\phi) = -\frac{\partial}{\partial x}\left[\frac{1}{\gamma}\int_\Omega \mathrm{d}x_\tau F(x,x_\tau)\rho_2(x,t;x_\tau,t-\tau|\phi)\right] + D_0\frac{\partial^2}{\partial x^2}\rho_1(x,t|\phi), \tag{2.15}$$

with $x_\tau \in \Omega$ being a second spatial variable needed due to the two involved particle positions in the delay force. The initial condition ϕ, must now be specified on the interval $[-\tau, 0]$ and can either be a constant sharp trajectory $X(-\tau \leq t \leq 0)$, or a distribution of such paths. The initial condition dependency of the PDFs will be omitted in the following.

Comparing (2.15) to an ordinary FPE for Markovian systems (2.1), one notes the appearance of an integral over all possible past positions weighted with ρ_2. Thus, there is a "delay-averaged" drift term (instead of the usual drift term). This is a direct consequence of the delay in the underlying equation of motion. For this reason, Eq. (2.15) is not self-sufficient, as opposed to the Markovian FPE. Even worse, the corresponding FPE for ρ_2, which we will discuss in Chapter 4, is found to explicitly involve the three-time PDF, ρ_3, and so forth. In this way, an infinite hierarchy of coupled equations emerges, whose nth member depends on the $(n+1)$-time probability density $\rho_{n+1}(x, t; x_\tau, t-\tau; ...; x_{n\tau}, t-n\tau)$.

Please note that we use here subscripts $i\tau$ for x to indicate different instances in time $t - i\tau$ of the stochastic event, which is somewhat inconsistent with Eq. (2.14), where subscripts denote dimensions (and superscripts denote times). However, as we explicitly give the time arguments as well, the notion should not be ambiguous.

Linear delay systems – time-convolutionless transformation The hierarchical structure of the FPE is in general a major issue in the analytical treatment.

An exception is again the case of *linear* delay Langevin equations, where the (steady-state) PDFs are all *Gaussian*. This makes it possible to obtain rigorous results for the various n-time PDFs in the steady state [77, 84], or if the initial condition is Gaussian or Delta-distributed [89]. We we will give explicit closed-form expressions for the steady-state PDFs in Sec. 4.2. Notably, in this special case, the infinite Fokker-Planck hierarchy can, in fact, be alternatively represented by a *finite* set of equations. This was explicitly shown in [92], were a time-convolutionless transform of the Langevin equation was employed. The latter is a technique to convert a time-nonlocal equation to a time-local one by means of Green's functions (which is not possible for nonlinear delayed LEs).

Side note: Hierarchical problems in probabilistic descriptions Similar hierarchy problems also occur in other contexts in statistical physics. A well-known example is the Bogoliubov-Born-Green-Kirkwood-Yvon (BBGKY) hierarchy, in which the time evolution of the one-particle density of an interacting multi-particle system depends upon the two-particle density and so on. For this problem, various closing strategies have been proposed, such as the simple mean-field (factorization) approximation and the more sophisticated dynamical density functional theory (involving an adiabatic approximation) [268, 269]. Another closure example is the mode coupling theory for glassy systems [270, 271]. We stress, however, that in all these examples the hierarchical structure emerges due to (conservative) particle–particle interactions. In contrast, the hierarchy of Fokker-Planck equations appearing in time-delayed systems – although having a similar structure – arises due to delay-induced *temporal* "interactions". Thus, the aforementioned strategies are not readily applicable in the present context. One might try to adopt some ideas, but to us, no obvious, useful connection occurred. This is a possible direction of future Research.

2.2.1 Earlier approximation schemes

Part II of this thesis is dedicated to the apparent problems in the probabilistic treatment induced by the fact that (2.15) is not closed. Besides explicitly considering the higher members of the hierarchy, a main objective is to discuss and compare different strategies to find closed, approximate expressions for the one-time PDF of nonlinear delay LE. Here, we will briefly review the main ideas of the most prominent approaches proposed in earlier literature. In Part II, we will explicitly apply them to our example systems.

Small delay expansion The unarguably most popular scheme is the small delay approximation on the level of the LE, which was originally introduced in [82, 83]. This is essentially a Taylor expansion up to linear order of the total deterministic forces (and of the noise intensity, if multiplicative, i.e., system-state dependent noise is present), in powers of τ. For the delay LE (1.42) this yields the approximate

Markovian LE [82, 83]

$$\dot{X}(t) = \gamma^{-1} F[X(t), X(t)] \left[1 - \tau \frac{\partial F(x, x_\tau)}{\partial x_\tau} \Big|_{\substack{x=X(t) \\ x_\tau=X(t)}} \right] + \sqrt{2D_0}\, \xi(t), \tag{2.16}$$

where τ only appears as a parameter.

Considering its simplicity in terms of the resulting expressions and the underlying idea, it is no wonder that this approximation is so popular [9, 38, 86, 89, 112]. However, it should be emphasized that this approach is appropriate for very small delay times only, which is, at the same time, a regime where the dynamics is qualitatively very similar to a Markovian one. This means that many effects induced by the memory are not captured by this approximation.

Perturbation theory Another approximation scheme introduced in earlier literature involves a first-order perturbation-theoretical ansatz for the density on level of the FPE. We refer the interested Reader to Ref. [86], where this *perturbation theory* (PT) was introduced, and only outline the main idea here.

Within the PT approach, one makes the assumption $|F_\mathrm{d}| \ll |F_\mathrm{s}|$ for the total deterministic force appearing in the LE [recall $F[X(t), X(t-\tau)] = F_\mathrm{s}[X(t)] + F_\mathrm{d}[X(t), X(t-\tau)]$, Eq. (1.43)]. This justifies that the delay force F_d is regarded as a small perturbation to the non-delayed, i.e., Markovian dynamics. As typical for a perturbation theory, an ansatz for the PDF with additive terms of increasing order in the perturbation is plugged into the Fokker-Planck equation.

The resulting first-order equation [86]

$$\frac{\partial}{\partial t}\rho_1(x,t) = -\frac{\partial}{\partial x}\left[\frac{1}{\gamma}\int_\Omega \mathrm{d}x_\tau F(x, x_\tau)\rho_c^{F_\mathrm{d}\equiv 0}(x_\tau, t-\tau|x,t)\right]\rho_1(x,t) + D_0\frac{\partial^2}{\partial x^2}\rho_1(x,t), \tag{2.17}$$

involves the conditional probability density with respect to the unperturbed, i.e., non-delayed system, $\rho_c^{F_\mathrm{d}\equiv 0}$ denoted here. The latter follows a closed *Markovian* FPE, and consequentially, the combination of the resulting two FPEs is self-sufficient. In this way, one obtains an approximate one-time PDF.

Disadvantages of existing approximations One conceptual disadvantage of these approximations is that they both effectively render Markovian descriptions. Thus, they may not capture the essence of the process with memory. Furthermore, they both rely on assumptions which are only appropriate for small delays, or small delay forces. To tackle this problem, we will explore new schemes in Part II, going beyond the "quasi-Markovian" regime.

2-state reduction It is noteworthy that specifically for the bistable system with delay, there is yet another type of approximation known from earlier literature (see Sec. 1.5.1), namely, the 2-state reduction proposed by Tsimring and Pikovsky [256]. Please note that this approach is not on equal footing with the other approaches discussed here, as it does not yield an approximate one-time PDF. In contrast, the

2-state reduction is based on the corresponding Master-equation, and aims for the position-autocorrelation function. We will elucidate this approach in more detail and apply it to our example model in Chapter 8. By construction, the 2-state reduction is appropriate in parameter regimes, where the dynamics is dominated by jumps between wells, as the intrawell fluctuations are fully neglected. It would be very interesting to extend this idea to other nonlinear systems, e.g., the washboard potential (which we have introduced in Sec. 1.5.2).

2.2.2 Probability current and apparent equilibrium of time-delayed systems

We have mentioned above that, for Markovian systems, the probability current J is an important quantity, which enables an easy distinction between EQ and NESS. Here we consider the probability current for systems with delay.

Due to the time delay, the probability current is a functional of the two-time PDF

$$J(x,t) = \left[\gamma^{-1}\int_\Omega \mathrm{d}x_\tau F(x,x_\tau)\rho_\mathrm{c}(x_\tau, t-\tau|x,t) + D_0\partial_x\right]\rho_1(x,t), \tag{2.18}$$

which is typically very hard to calculate [the current is defined analogously to the Markovian case by writing (2.15) in the form (2.9)]. It contains the delay-averaged drift term. In regard to Markovian systems, we have argued above that J is constant in steady states, where $J = 0$ signals EQ.

How is the situation in the presence of time delay? In particular, let us consider the three example systems (the linear, bistable and periodic one), introduced in the last Sections of Chapter 1. An apparent shared feature of the systems is their *coordinate inversion symmetry*, i.e., the equations are invariant under the transformation $x \to -x$. In other words, the spatial symmetry of the systems is not broken, even in the presence of delay force. As a consequence, the probability current naturally vanishes in steady states. This can be easily seen from the FPE, which then reads $\partial_x J = 0$, yielding a constant current which would switch its sign under the coordinate inversion. Of course, a constant current with $J = -J$ must be zero. Besides the main example applications that we will consider throughout this thesis, in fact, various systems fall into this very class. Naively employing the aforementioned criterion of EQ established for Markovian systems, such systems with time delay appear to be in thermal equilibrium. We will prove this idea wrong in Part III.

2.2.3 Side note: Delay in the probability density

In this thesis, we consider systems which are subject to a feedback force that depends on the time-delayed position $X(t-\tau)$. Thus, the delay appears in a *stochastic* quantity X in the LE. As we have mentioned above, the corresponding Fokker-Planck equation for the one-time PDF is not closed. In earlier research [70], we have also

considered the situation that a time delay is instead introduced in an ensemble-averaged, and hence *deterministic*, quantity. Specifically, we have considered feedback that depends on the delayed *mean* particle position $\langle X(t-\tau)\rangle$. In this case, one readily finds a closed FPE, which then involves the delayed probability density $\rho(t-\tau)$. Thus, this case is much easier to handle analytically. The underlying reason is that then, there is no direct interplay between delay and noise, and both can be treated separately. As we have shown in [70], such a delay force may indeed also yield interesting dynamics, e.g., induce a ratchet effect. However, this type of feedback is harder to realize in an experimental setup, and, more importantly, a LE with $\langle X(t-\tau)\rangle$ has much fewer applications. Therefore, we focus on systems with a time-delayed position-dependence in this thesis.

3 | Stochastic thermodynamics

Finally, we introduce the third great theoretical framework that will be employed in this thesis: stochastic thermodynamics. The latter is a quite modern theory, which aims at a generalization of thermodynamic notions towards small-scale systems dominated by thermal fluctuations and towards nonequilibrium systems [126]. On this scale, individual fluctuating trajectories can transiently violate the second law of thermodynamics, but in the limit of large systems (or long times), one shall always recover the well-established laws of macroscopic thermodynamics.

Because this theory is applicable to *nonequilibrium*, *stochastic* systems, it is a promising candidate to also describe stochastic processes with time delay. However, as we will see in the following, many fundamental concepts in stochastic thermodynamics are based on the *Markov assumption*[1]. In this Chapter, we will first lay out the basic concepts in brief, focusing on Markovian, overdamped systems consisting of a single colloidal particle coupled to a heat bath. Towards the end of the Chapter, we will turn to the question how stochastic thermodynamics applies to time-delayed systems, reviewing some previous results and forumlating expectations and apparent problems. These considerations provide the basis for our investigations in Part III of this thesis.

As in the preceding Chapters, we will use a notation indicating that the system is one-dimensional, where X denotes the stochastic system state. However, the definitions can readily be applied to multidimensional systems as well, by simply replacing X by a vector $\underline{X} = \{X_0, X_1, ..., X_n\}$ ($n+1$ being the spatial dimension). Further, we will again focus on steady states.

An educational introduction to this field, specifically to the energetics of stochastic systems, is given in Sekimoto's book [124]. Some ideas of our introduction are inspired by the latter. Further, the review article by Seifert [125] summarizes the elegant and consistent theoretical concepts, which have been developed in the last decades, focusing on the notion of fluctuating entropy.

Before we begin with introducing the main concepts, we give a brief historical overview, locating this theory within the overall framework of theoretical physics. This is a side note and not necessary to understand the remainder of this thesis.

[1] The main problem is rooted in the notion of *total entropy production*, which lies at the heart of this theory.

3.1 Side note: Some historical notes & Where is stochastic thermodynamics located in theoretical physics?

Macroscopic thermodynamics goes back to the seventeenth century[2]. Its importance for physics and engineering is undeniable and well-known. This traditional, macroscopic thermodynamics is essential for the development of macroscopic machines, and made possible the industrial revolution. It describes many complex phenomena in the world on large scales, e.g., on length scales comparable to the size of a human. The laws of macroscopic thermodynamics are, however, purely empirical. Further, many concrete statements (like Carnot efficiency) only hold for systems that are close to *equilibrium*, or quasistatically driven. In Sec. 1.3.5 we have mentioned the corresponding notion of macroscopic reversibility. A key concept is the formulation of a purely phenomenological state-function, called *entropy*, which can never decrease if a system is isolated (the second law).

Later, the framework of *statistical mechanics* was founded, which connects thermodynamic laws to *first principles* by unraveling the routes to the microscopic, Newtonian laws. These ideas base on a picture, where the world is made of particles (atoms or molecules) that are ruled by the *laws of classical mechanics*. A central result is the *statistical interpretation of entropy* by, e.g., Ludwig Boltzmann, whose meaning for modern physics and our recent understanding of the world is monumental. The (Shannon) entropy, which was before just an empirical concept, is now linked to microscopic states and their probabilities. Loosely speaking, the entropy is proportional to the logarithm of the number of possible microstates (i.e., microscopic configurations of the individual particles) that could all correspond to the observed macroscopic state (defined via volume, temperature, pressure,...) [115]. Due to this interpretation, the second law of thermodynamics loses the phenomenological character and gains mathematical rigor[3]. This theoretical advancement also introduces a link between physics and information theory [274], e.g., the immediate relation between Shannon entropy and information discussed in the works by Jaynes [275].

Later on, a new way of describing observed phenomena was developed using *stochastic* models, which we have introduced in the last Chapter. We recall that this description is a coarse-grained, mesoscopic one, sitting in between the microscopic, mechanical laws and the macroscopic ones. It introduced a new formulation of statistical mechanics [159]. Important theoretical concepts include the Langevin and Fokker-Planck equation [173] (also the Smoluchowski equation, and dynamical density functional theory), and the fluctuation-dissipation relations. Taking thermal fluctuations explicitly into account, it is suitable for systems that are dominated by "random" perturbations, typically due to the chaotic motion of molecules in the surrounding gas or fluid. The stochastic way of modeling (for, e.g., colloidal suspen-

[2] An example are the works by Robert Boyle, who lived 1627-1691 [272].

[3] Here one should also mention the long debate it has cased, including *Lochschmidt's paradox*, i.e., how can an irreversible thermodynamics be derived from reversible, time-symmetric fundamental laws, and the ongoing question why the initial state of the universe had an exceptionally low entropy [273].

sions) is the best description we have to date to describe complex phenomena on the nano- to micrometer scale. This is the length scale where one can find macromolecules, viruses, individual and small clusters of cells, which is why stochastic modeling is also particularly important for the understanding of biological systems. Rigorous connections between stochastic modeling and the microscopic world, i.e., first principles, were later established. Important cornerstones are the projection method by Mori and Zwanzig (see Sec. 1.3.1), and toy models like, e.g., the Rouse model [14, 15] (suitable to model viscoelastic fluids).

Despite these great successes, we are far away from a full understanding of the complex world that we are part of. Specifically, some important links between the aforementioned frameworks remain, and there are still lots of open questions. For example, there are the two following issues: First, *the generalizing of thermodynamic laws, and the frameworks of statistical physics, to systems far from equilibrium.* After all, the real world is mostly out of thermal equilibrium, and many phenomena cannot even be considered as a relaxation towards it. Also the traditional concept of considering passive systems subject to (quasistatic) external driving is limited; and to capture the essence of complex processes, one often needs models that are *intrinsically* driven. This is particularly true for biological, i.e., living matter, and artificial, energy-consuming systems, like computers or robots. It is also a useful way to treat the feedback-controlled systems with time-delay considered in this thesis. A second issue is that among all theories and connections mentioned above, an important link is missing, which is *the generalization of traditional, macroscopic thermodynamics to the stochastic level of description.*

Both problems are addressed by stochastic thermodynamics, which aims at finding laws and suitable definitions of thermodynamic properties on the stochastic level. This new theory is currently under development. Significant theoretical advances include the path integral representation of the total entropy production, stochastic energetics, and the fluctuation theorems, which we all briefly introduce in the following. In this framework, thermodynamic properties are investigated on the level of individual fluctuating trajectories. Further, basic premises are not limited to equilibrium situations[4], making this theory particularly promising to gain deeper physical understanding of complex, real-world systems, including time-delayed systems. Stochastic thermodynamics thus provides an essential piece of the puzzle to the overall understanding, and has immediate importance for, e.g., biological systems [131, 133] and computer science [276]. Moreover, it also provides important insight into quantum systems [134, 135, 277], and deepens the connection between physics and *information theory*, which was already introduced by Boltzmann and Jaynes and which might play an increasing role in the future [131].

3.2 Stochastic energetics

Let us start introducing the concrete frameworks. In this Section, we review the energetic level of description given by the framework of Sekimoto [99, 124]. We start by considering the energy (ex-)change between an (overdamped) particle and the

[4]accept for the important assumption that the "bath" is in equilibrium!

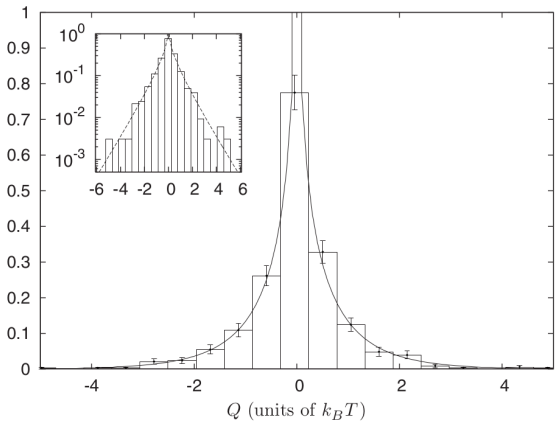

Figure 3.1: *A figure from Ref. [137], showing the probability distributions of heat Q generated along trajectories of length 0.5 sec of a system described by a linear LE in thermal equilibrium. The lines plot the analytical expression given by Eq. (17) of [137], while the bars show experimental results from a colloidal particle in a harmonic optical trap generated by a laser. Inset: same data, with logarithmic y-axis.*

bath consisting of the surrounding fluid particles, along a trajectory $\mathbf{X} = \{X(s)\}$, $s \in [t, t+\mathrm{d}t]$. The bath exerts a force on the particle through random kicks, namely, the Langevin force $\sqrt{2k_\mathrm{B}\mathcal{T}\gamma}\,\xi(s)$. On the contrary, the particle also exerts a force onto the bath, $-\gamma\dot{X}(s)$, the friction[5] When yielding a particle displacement $\mathrm{d}X$, each of these forces is inevitably associated with energy exchange between particle and bath, called heat. Accordingly, the fluctuating heat δq flowing to the reservoir during the infinitesimal time $\mathrm{d}t$, is thus given by[6]

$$\delta q(t) = \left[\gamma \mathrm{d}X(t)/\mathrm{d}t - \sqrt{2k_\mathrm{B}\mathcal{T}\gamma}\,\xi(t)\right] \circ \mathrm{d}X(t). \tag{3.1}$$

The $\circ$-symbol indicates usage of Stratonovich calculus. Note that as is commonly done, we assume the idealized heat bath to remain at all times in equilibrium, at a constant temperature, unaffected by possible external forces or the feedback. Further note that the sign convention of the heat employed here is consistent with [125], but different from [124].

Second, the particle may have an internal energy u. Since we consider the overdamped limit here, we do not properly resolve its kinetic energy. We further assume the particle to have no internal degrees of freedom. Consequentially, u is simply given by the potential energy, i.e., the position of the particle within an external potential V. This means the increment of internal energy $\mathrm{d}u$ reads

$$\mathrm{d}u(t) = V'\left[X(t)\right] \circ \mathrm{d}X(t). \tag{3.2}$$

Finally, in the presence of external, *non*-conservative forces $F_{\mathrm{non-con}}$, these may perform work on the particle given by

$$\delta w(t) = F_{\mathrm{non-con}} \circ \mathrm{d}X(t). \tag{3.3}$$

In total, we have identified three types of possible energy (ex-)change of a particle along a random, erratic trajectory of infinitesimal short time, $\mathrm{d}t$. Accordingly, the

[5]Note that the relation between both forces is usually given by the Fluctuation-dissipation relation, see Sec. 1.3, Eq. 1.35.

[6]Please note the usage of δ symbols to express inexact differentials [278], which is because the heat and work are path-functions (not state-functions).

fluctuating energy changes during a time interval $[t_s, t_f]$ of finite length, are given by

$$u = \int_{t_s}^{t_f} \mathrm{d}u, \tag{3.4}$$

$$e = \int_{t_s}^{t_f} \delta e, \text{ with } e \in \{q, w\}, \tag{3.5}$$

while the *rates* are given by $\dot{u} = \mathrm{d}u/\mathrm{d}t$, $\dot{e} = \delta e/\mathrm{d}t$.

These definitions seem to be very reasonable and consistent with our physical intuition. When combining them with the LE (1.16), i.e., the equation of motion (or force-balance), one readily recovers the *first law of thermodynamics*

$$\delta w(t) = \mathrm{d}u(t) + \delta q(t). \tag{3.6}$$

This is the generalization of the first law to the stochastic level. Energy conservation must be valid along *each* fluctuating trajectory (of arbitrary length), as expected.

The description of the fluctuations of heat and work is one central objective of stochastic thermodynamics. An important contribution was the work [137], where the heat fluctuations in equilibrium and in a NESS of a linear, Markovian system were analytically calculated and experimentally confirmed, see Fig. 3.1 for an example.

Special case: A free particle If no external potential or non-conservative forces are present, i.e., the LE reads $\gamma \dot{X} = \sqrt{2k_\mathrm{B}\mathcal{T}\gamma}\xi(t)$, all energy contributions trivially vanish, as follows from (3.1 -3.3, 3.6). In such situations, the force exerted by the bath as a random kick does not "yield" energy exchange, but is *immediately* balanced out by friction. Thus, there are no heat, inner energy, or work fluctuations. This reflects the meaning of the Markov assumption on an energetic level (in the overdamped limit).

3.2.1 Steady states

Many systems approach steady states in the long-time limit ($t \to \infty$), where the transient initial condition effects have decayed away (see Sec. 1.2.2). In steady states, the ensemble average of the internal energy must be a conserved quantity, as is physically clear for stability reasons[7], and as can also be seen from (3.2), which implies

$$\begin{aligned} \Delta U = \langle \Delta u \rangle &= \left\langle \int_{t_s}^{t_f} \mathrm{d}u(t) \right\rangle = \left\langle \int_{t_s}^{t_f} V'\left[X(t)\right] \circ \mathrm{d}X(t) \right\rangle \\ &= \langle V\left[X(t_f)\right] \rangle - \langle V\left[X(t_s)\right] \rangle \overset{\text{ss}}{=} 0. \end{aligned} \tag{3.7}$$

Here and in the following, we denote ensemble averages by capital letters: $\Delta U = \langle \Delta u \rangle$, $\delta W = \langle \delta w \rangle$, $\delta Q = \langle \delta q \rangle$. Thus, when considering steady-state ensemble

[7] No system can remain stable when its internal energy constantly and for ever increases.

averages, one is only interested in the heat and work, which must, according to the first law (3.6), fulfill

$$\delta W \overset{\text{ss}}{=} \delta Q. \tag{3.8}$$

Further, due to the time-invariance of steady states, the mean heat and work *linearly* increase with the length of the considered trajectory. For this reason, we will focus on the mean rates

$$\dot{Q} = Q/(t_f - t_s), \quad \text{and} \quad \dot{W} = W/(t_f - t_s). \tag{3.9}$$

In the following, we further distinguish the steady states into two different kinds.

Thermal equilibrium In the absence of non-conservative external forces, the work is trivially *zero* (has no fluctuations) for all times, as given by (3.3). In such a situation, the system is expected to approach thermal equilibrium in the long-time limit, as no external driving prevents it. Then, the work, internal energy, and heat all have zero mean

$$\text{EQ}: \quad \delta W = \delta Q = 0. \tag{3.10}$$

Internal energy and heat may nevertheless exhibit non-trivial fluctuations, see Fig. 3.1 for an example.

NESS When external non-conservative forces act on the particle, they prevent equilibration, because they continuously pump energy into the system. Ultimately, this energy is dissipated in the form of heat, as stated by the first law (3.6). Thus, a finite net heat flow is a signature of non-equilibrium. When the driving is time-independent, the system may nevertheless approach a stationary state, which is then a *nonequilibrium steady state* (NESS) characterized by a steady energy flow through the system[8]

$$\text{NESS}: \quad \delta W = \delta Q \neq 0. \tag{3.11}$$

This particular system state is very important for the framework of stochastic thermodynamics [125]. First of all, it is the normal operating regime of many nonequilibrium systems, and is hence relevant for experiments and applications. Further, it is technically somewhat easier to handle than other nonequilibrium states, as the PDFs are time-independent. Various exact relations are known for the fluctuation of entropy and energy (ex-)changes of systems in a NESS [125], including the existence of *fluctuation theorems* (which we will come to below), and exact expressions of the distributions in linear systems [137].

[8] A particular case of a NESS is approached in the absence of an external potential. Then, the inner energy (3.2) is trivially zero (has no fluctuations), and the fluctuations of heat and work, which still may have finite mean values, are *identical*, i.e., $\delta w = \delta q$ along each trajectory (3.6).

3.3 Fluctuating entropy

So far, we have discussed the framework of stochastic energetics, where we encountered the generalization of the first law of thermodynamics to the fluctuating world. Now we consider the notion of fluctuating entropy, yielding a generalization of the second law. We will begin with the definition of the *total entropy production (EP) functional*, which is a central concept with deep physical implications. We will see that it is inextricably linked to the *thermodynamic arrow of time*. This definition will further automatically bring us to the fluctuating Shannon entropy and information.

The framework of stochastic thermodynamics reveals that the total EP of a process can literally be calculated via its *irreversibility*, i.e., the time-reversal symmetry breaking. Let us first give an illustration inspired by [279], and afterwards make things precise.

Imagine a movie of a recorded event, e.g., the periodic motion a pendulum swinging back and forth, or the mixing of coffee and milk in a glass of latte macchiato [279]. The total EP measures how "sure" an observer can be of the fact that the movie is played forward, and not reversed. In the first case, a short movie of the oscillating pendulum does not enable the observer to clearly decide on this question[9]. In sharp contrast, everybody could immediately tell[10] that the recorded movie of the latte macchiato, which would *de*mix in the reversely played movie, is certainly played forwards. The difference between the two processes is that the mixing of coffee and milk yields a high increase of entropy, signaling a fast relaxation towards equilibrium (where coffee and milk are well mixed). On the contrary, the pendulum does not render a high increase in entropy, and hence, has a less pronounced arrow of time. In other words, the demixing of the coffee would violate "more clearly" the second law of thermodynamics (which we give below). The *decisiveness* of the imaginative observer is really what the total EP measures. For these macroscopic examples, the decisiveness appears to be a rather binary thing: The process is, or is not, reversible (corresponding to equilibrium or nonequilibrium). However, in situations where the observer considers small-scale processes (e.g., small numbers of colloids or molecules), one can already guess that it will not always be so clear. For example, the process can transiently appear reversible, although it is on average irreversible (or vice versa). Thus, the total EP is a fluctuating quantity.

To be more precise, the total EP compares the following processes: The *forward* process is the actual, observable system state evolution. As thermodynamics tells us, this evolution is characterized by the system's attempt to relax towards equilibrium, possibly prevented by some external driving that might be time-dependent. The *backward* process is an imagined process, where the time is running backwards and the system is driven by the same external driving protocol, but, if time-dependent, in reversed temporal order. Further, for a realization of the process, i.e., a single

[9] In a very long movie, the observer could also decide if the movie of the pendulum is played forward, as the pendulum would eventually stop due to friction, a process that produces entropy.

[10] even without being aware of thermodynamics, as our brains are trained in a world which fulfills the second law

fluctuating trajectory $\mathbf{X} = \{X(t')\}, t' \in [t_s, t_f]$, the time-reversal operation amounts to replacing t by $-t$, and flipping the signs of velocities (which have "odd" parity under time-reversal [117]), but keeping the signs of positions X and accelerations $\ddot{X}$ the same (which are "even" quantities). In Markovian, memory-less systems, the backward process is typically as well realizable in experiments, and can further be accessed in simulations.

Comparing the chances to see a particular random trajectory in the forward process and the corresponding time-reversed trajectory in the backward process, is the very definition of total EP [125, 280]

$$\Delta s_{\mathrm{tot}}[\mathbf{X}] = k_{\mathrm{B}} \ln \frac{\mathcal{P}[\mathbf{X}]}{\hat{\mathcal{P}}[\hat{\mathbf{X}}]}, \tag{3.12}$$

involving the path probabilities $\mathcal{P}$ and $\hat{\mathcal{P}}$ for forward and backward process, respectively. The functional (3.12) assigns to a specific realization of the random process $\mathbf{X}$ a certain amount of total EP, which, in turn, depends on the path probabilities of the entire ensemble. As we will elaborate below, Δs_{tot} by construction fulfills an integral fluctuation theorem [see Eq. (3.22)], which is a statement about the symmetry properties of the fluctuations. Furthermore, the ensemble average of the total EP (3.12) automatically fulfills the *second law of thermodynamics*

$$\Delta S_{\mathrm{tot}} \geq 0, \tag{3.13}$$

(as follows from the fluctuation theorem, see below). The capital letter is again used to denote the ensemble average. Equation (3.13) states that the mean total EP due to any spontaneous process can never be negative. This law breaks the symmetry between past and future, and defines the thermodynamic arrow of time. The fact that (3.12) implies the second law (3.13) exemplifies that the definition of total EP employed in stochastic thermodynamics is a generalization of the concept of entropy to the fluctuating level, from which one can recover (on average) results from traditional thermodynamics.

At this point, we consider Markovian systems only, e.g.,our delay LE (1.16) at $\tau = 0$ where $F[X(t), X(t-\tau)] \equiv F[X(t)]$. Using the conditional path probability

$$\mathcal{P}[\mathbf{X}] = \mathcal{P}[\mathbf{X}|x_s]\rho_1(x_s, t_s), \tag{3.14}$$

the total EP from definition (3.12) can be rewritten as the *entropy balance* equation

$$\Delta s_{\mathrm{tot}}[\mathbf{X}] = \Delta s_{\mathrm{sh}} + k_{\mathrm{B}} \ln \frac{\mathcal{P}[\mathbf{X}|x_s]}{\hat{\mathcal{P}}[\hat{\mathbf{X}}|x_f]}, \tag{3.15}$$

involving the change of the *fluctuating Shannon entropy*

$$s_{\mathrm{sh}}[X, t] = -k_{\mathrm{B}} \ln[\rho_1(X, t)] \tag{3.16}$$

of the one-time probability density function ρ_1. Its mean value is

$$S_{\mathrm{sh}} = \langle s_{\mathrm{sh}} \rangle = -k_{\mathrm{B}} \int_{\Omega} \rho_1(x, t) \ln[\rho_1(x, t)] \mathrm{d}x. \tag{3.17}$$

The change of Shannon entropy only depends on the start and end points[11] of the trajectory $\mathbf{X}$.

Let us now consider the second term on the r.h.s. of Eq. (3.15). As mentioned in Chapter 1, $\mathcal{P}$ conditioned on the starting point $X(t_s) = x_s$ is proportional to the exponential of the Onsager-Machlup action [117, 125]. The corresponding $\mathcal{P}[\mathbf{X}|x_s]$ is provided in Eq. (1.27). In such a Markovian situation, the path probability of the *backward process* is analogously given by

$$\hat{\mathcal{P}}[\hat{\mathbf{X}}|x_f] = \mathcal{J}\, e^{-\int_{t_s}^{t_f} \left(-\dot{X}(t')-\gamma^{-1}F[X(t')]\right)^2/(4k_{\mathrm{B}}\mathcal{T}\gamma)\mathrm{d}t'}. \tag{3.18}$$

Importantly, the sign of velocities $\dot{X}$ is flipped in the backward process, while the position X keeps its sign, corresponding to odd and even time-reversal parity, respectively. Using the last expression and $\mathcal{P}$ from Eq. (1.27), one can immediately identify the remaining term in (3.15) as the *medium entropy production*, i.e,

$$\Delta s_{\mathrm{m}} = k_{\mathrm{B}} \ln \frac{\mathcal{P}[\mathbf{X}|x_s]}{\hat{\mathcal{P}}[\hat{\mathbf{X}}|x_f]} = \frac{\delta q}{\mathcal{T}}, \tag{3.19}$$

which is proportional to the fluctuating heat, as defined in Sekimoto's framework given in Eq. (3.1). Contrary to the Shannon entropy change, this entropic contribution depends on the entire trajectory, i.e., the specific path that the particle takes, not only on the starting and end point[12]

In total we have seen that, for a Markovian system (with position-dependent, time-dependent or constant) driving, the total EP is given by the change of Shannon entropy plus the medium entropy production. The first term quantitifies entropy production due to configurational changes of the system, while the second term quantifies entropy production in the medium associated with heat dissipation.

It is well-known that the ensemble average of the total EP of a Markovian stochastic system can equivalently be derived by means of the Fokker-Planck equation [125] [instead of path integrals (3.15)], yielding the alternative expression

$$\dot{S}_{\mathrm{tot}} = \int \frac{\gamma J(x,t)^2}{k_{\mathrm{B}}\mathcal{T}\,\rho_1(x,t)}\,\mathrm{d}x \geq 0. \tag{3.20}$$

We will explicitly show this in Chapter 9.

3.3.1 Thermal equilibrium & Nonequilibrium steady states

In the previous Section, we have considered the energy flows characteristic for nonequilibrium steady states vs. thermal equilibrium. After defining the total entropy production (3.15), we can now revisit these two types of steady states from an entropic perspective.

[11] It can therefore be considered as a "temporal boundary term". Typically, it does not increase with the length of the trajectory and, thus, has a negligible contribution along very long trajectories.

[12] It can be considered a "temporal bulk term", which typically increases with the overall length of the trajectory.

First we notice that in steady states, which are defined by $\partial_t \rho = 0$, the mean Shannon entropy (3.16) is, by construction, a conserved quantity. Thus, $\Delta S_{\rm tot} = \delta Q/\mathcal{T}$. Further, we can now finally give the fundamental, thermodynamic definition of equilibrium, that is, $\Delta S_{\rm tot} = 0$. This notion is consistent with the term equilibrium which we used above (e.g., when we discussed the heat bath, or the probability current in Chapter 2).

In all other cases, the LE yields $\Delta S_{\rm tot} > 0$, as follows from (3.15, 3.19) and the second law. Thus,

$$\Delta S_{\rm tot} = \delta Q/\mathcal{T} \begin{cases} = 0, \ \text{EQ}, \\ > 0, \ \text{NESS} \end{cases} . \tag{3.21}$$

According to Eq. (3.21), a finite medium entropy production signals nonequilibrium. Further, the medium EP and the corresponding net heat flow are *strictly positive* quantities. This means, in Markovian systems, external driving (such as a constant or position-dependent force, see [125] for examples), can only cause a net energy flow from the particle to the bath, and never a reversed one[13]. In other words such an ("uninformed", no feedback/control) external force ultimately yields a heating of the environment, and never cools it down. This is consistent with our physical intuition[14].

One should note that in our theoretical model, where the bath is assumed to be infinitely big, see Sec. 1.3.1, such a heating or cooling can never yield a change of the bath's *temperature*, which is artificially kept fixed by to the construction of the model.

Canonical ensemble We aim to mention that the Langevin equation generally describes a system at constant temperature, volume, and particle number, thus, the thermal equilibrium describes a canonical ensemble situation with a *free energy* $\mathcal{F} = U - \mathcal{T} S_{\rm sh}$ w.r.t. the colloid [125]. Here, the term "system" refers to the colloidal particle at position $X(t)$.

Let us recall that among the derivation of the LE from mechanical laws, the constant temperature was introduced by hand when the noise is "derived", see the given arguments between Eq. (1.34) and (1.35) in Sec. 1.3.1. This is noteworthy, as the starting point of this (Mori-Zwanzig) derivation was a system with constant energy, but fluctuating temperature. Thereby, the term "system" refers to the colloidal particle *plus* surrounding "bath particles".

Information gain The fluctuating Shannon entropy (3.16) has a direct connection to information theory. Specifically, $s_{\rm sh}$ is equivalent[15] to the *information* i gained by (i.e., the surprisal value of) a measurement of $x = X(t)$ [274]. Correspondingly, the

[13] Be aware that this statement refers to ensemble averages only, *individual* random trajectories can transiently violate the second law and hence have a negative heat flow, as explicitly stated by the fluctuation theorems like (3.26).

[14] Heat is the energy exchange due to the friction between particle and solvent initiated by the external force.

[15] Here we refer to information measured in units of $k_{\rm B}$. More generally $i \propto -\ln[\rho_1(X,t)]$.

mean Shannon entropy is equal to the expected information of one measurement, giving the *uncertainty* about the system [275]. Thus, it can be considered as the information that is needed to find out in which state the system is. The EQ state is the state with the maximal uncertainty under given external constraints (like constant temperature).

3.4 Fluctuation theorems

A particularly important result of stochastic thermodynamics are the various celebrated *fluctuation theorems* [211]. They represent general symmetry constraints for the distributions of thermodynamic quantities. Depending on the system at hand, there are fluctuation theorems for, e.g., the total EP, for the heat, and the work, which can often be derived from quite general assumptions. As a particularly important example, we here consider the *integral fluctuation theorem* (IFT)

$$\left\langle e^{-\Delta s_{\mathrm{tot}}/k_{\mathrm{B}}} \right\rangle = 1, \tag{3.22}$$

which is valid for the total EP along arbitrarily long trajectories, and holds in and outside of steady states or equilibrium. It can be derived from the definition (3.12), as follows

$$\begin{aligned} \left\langle e^{-\Delta s_{\mathrm{tot}}/k_{\mathrm{B}}} \right\rangle &= \int e^{-\Delta s_{\mathrm{tot}}[\mathbf{X}]/k_{\mathrm{B}}} \mathcal{P}[\mathbf{X}] \, \mathfrak{D}\mathbf{X} \\ &= \int e^{\ln \frac{\hat{\mathcal{P}}[\hat{\mathbf{X}}]}{\mathcal{P}[\mathbf{X}]}} \mathcal{P}[\mathbf{X}] \, \mathfrak{D}\mathbf{X} \\ &= \int \hat{\mathcal{P}}[\hat{\mathbf{X}}] \, \mathfrak{D}\mathbf{X} = 1. \end{aligned} \tag{3.23}$$

Thus, we only had to assume normalization of the paths, which is generally given[16]. The IFT imposes a universal constraint on the symmetry of the total entropy fluctuations. This reasoning can readily be generalized to multidimensional systems (as we will do in Chapter 9).

Furthermore, the IFT (3.22) readily implies the second law of thermodynamics (3.13). This can be shown using the general relation

$$\ln(\sigma) \leq \sigma - 1, \quad \forall\, \sigma > 0, \tag{3.24}$$

which can be applied to the present case, when expressing the ensemble average as sum over $N \to \infty$ realization of the random process, as follows

$$\begin{aligned} \langle -\Delta s_{\mathrm{tot}}/k_{\mathrm{B}} \rangle &= \sum_{i=1}^{N} \frac{-\Delta s_{\mathrm{tot,i}}/k_{\mathrm{B}}}{N} = \sum_{i=1}^{N} \frac{\ln(e^{-\Delta s_{\mathrm{tot,i}}/k_{\mathrm{B}}})}{N} \\ &\leq \sum_{i} \frac{e^{-\Delta s_{\mathrm{tot,i}}/k_{\mathrm{B}}} - 1}{N} = \left\langle e^{-\Delta s_{\mathrm{tot}}/k_{\mathrm{B}}} \right\rangle - 1 \overset{\mathrm{IFT}}{=} 0. \end{aligned} \tag{3.25}$$

[16]Please note that the backward path probabilities sum up to 1, also when integrated over all forward paths.

More information about the entropy fluctuations are provided by the *detailed fluctuation theorem* [125, 281, 282]

$$\frac{P(\Delta s_{\mathrm{tot}})}{P(-\Delta s_{\mathrm{tot}})} = e^{\Delta S_{\mathrm{tot}}/k_{\mathrm{B}}}, \tag{3.26}$$

which holds along arbitrarily long trajectories in a steady state (NESS or EQ) of a Markovian system. Loosely speaking, it states that observing "violations of the second law" is exponentially less likely than observing "fulfillment". In particular, the probability P of observing a particular *negative* value of total entropy production, $-\Delta s_{\mathrm{tot}}$, along a fluctuating trajectory, is exponentially less likely then observing the corresponding positive value, $+\Delta s_{\mathrm{tot}}$, where the exponent is given by the average value of produced total EP. Thus, in equilibrium (where $\Delta S_{\mathrm{tot}} = 0$), negative and positive EP are equally likely, and the EP distribution is thus symmetric w.r.t. the origin. On the contrary, in a NESS, the mean total EP linearly increases with length of the trajectory. Therefore, the chances of observing negative EP exponentially decrease with the length of the trajectory.

There are various further fluctuation theorems. They are not only very important for the theoretical understanding, but also useful for applications [165, 283–286], like the advancement of simulation techniques. For example, an immediate practical application of the IFT is that it can be used to test numerical integration schemes, by checking how well the generated random trajectories fulfill (3.22). Another prominent example of a fluctuation theorem is the *Jarzynsky relation* [287]

$$\left\langle e^{-w/(k_{\mathrm{B}}\mathcal{T})} \right\rangle = e^{-\Delta\mathcal{F}/(k_{\mathrm{B}}\mathcal{T})}, \tag{3.27}$$

which enables reconstruction of the free-energy difference $\Delta\mathcal{F}$ between two EQ states upon experimental measurements of the fluctuating work w while driving the system from one of the EQ states to the other (along a non-equilibrium path). This relation has later been refined by Crooks [288], additionally taking the backwards process into account, in particular [125]

$$\frac{P(w)}{P(-w)} = e^{-(w-\Delta\mathcal{F})/\mathcal{T}}. \tag{3.28}$$

At this point, we aim to briefly revisit the discussion about causality, which we have started before in Chapter 1, Sec. 1.3.5.

3.4.1 Route to first principles – Axiom of causality

How deep is the physical insight provided by the integral fluctuation theorem?

While (3.23) is a "derivation" of the second law starting from the stochastic LE, one should emphasize that the LE itself already breaks time-reversal symmetry. Thus, this is not a derivation from first principles! Importantly, the route to first principles always requires at some point the introduction of an element of asymmetry by hand.

This can be done by making suitable assumptions about the initial condition, or by imposing an *axiom of causality*. At this point, we resume the discussion from Sec. 1.3.5.

We recall that, in contrast to the underlying microscopic dynamics, the Langevin equation [that fulfills the FDR (1.35)] does not obey mechanical time-reversal symmetry. This is because, during the derivation, initial condition correlations must be neglected, or a causality principle must be assumed, see Sec. 1.3.5 for further details.

Indeed, the mathematical derivation of the fluctuation theorem from first principles, specifically the second law from [211], requires causality. To emphasize this, [211] explicitly shows that when assuming an anti-causal principle, one can derive – from the correct mechanical laws – a fluctuation theorem which implies an "anti-second law", i.e., decrease of entropy in the course of time (inconsistent with our observations of the real world). This clearly demonstrates the necessity of an axiom of causality to derive macroscopic irreversibility (thus, the second law), from time-reversible mechanical laws.

While the second law is hence rather imposed in the stochastic model than proven by (3.22), it should nevertheless be stressed that the symmetry properties uncovered by the fluctuation theorems are indeed nontrivial and important. In particular, fluctuation theorems represent one of the rare universal, rigorous results for *nonequilibrium* systems. By telling us, e.g., how unlikely transient "violations" of the second law are, they offer an intriguing view on thermal fluctuations.

However, as the fluctuation theorems are not readily applicable to time-delayed systems (which are, e.g., non-Markovian, have no notion of free energy,...), we will not go into further detail here, but refer the interested Reader to [125] and references therein.

3.5 Information

We have already mentioned a couple of times that concepts from information theory also play a crucial role in statistical physics, and especially in (stochastic) thermodynamics. This is particularly true in the context of control. In addition to energy and entropy, information represents a third perspective on thermodynamic processes [274]. In Sec. 3.3 we have already given the connection between Shannon entropy and information gain. Here, we will introduce some further information-theoretic quantities, which will be relevant for our investigations.

As opposed to the considerations up to this point, we now consider a system with multiple d.o.f. (subsystems), such that we can discuss the information flow *between* these subsystems. Lets say, for the sake of simplicity, we have a system with $n+1$ coupled d.o.f., X_0, X_1,..., X_n. We again focus on the case of Markovian systems only. (The d.o.f. $X_{j\neq 0}$ will later represent the controller, as explained in detail in Chapter 9.)

But let us first introduce yet another, more established quantity in statistical physics, that is, the mutual information. In a second step, we will show their connection.

Mutual Information The mutual information between all d.o.f. is defined as

$$\mathcal{I}(\underline{x}) = \int \rho_{n+1}(\underline{x}) \ln \frac{\rho_{n+1}(\underline{x})}{\rho_1(x_0)\rho_1(x_1)...\rho_1(x_n)} \,\mathrm{d}\underline{x}, \tag{3.29}$$

with the $(n+1)$-point joint probability distribution $\rho_{n+1}(\underline{x}, t)$ and $\underline{x} = \{x_0, x_1, ..., x_n\}$. To have a shorter notation, we omit the time-dependencies in the formulae here.

In other words, the mutual information among a set of variables is the Kullback-Leibler divergence between the joint and the marginal PDFs. It quantifies the "amount of information" obtained about one set of d.o.f. through measurement of the other degrees of freedom. It is, by construction, a *symmetric* quantity, i.e., the mutual information does not distinguish which of the two sets is the observed, and which is the unobserved one.

In steady states, the mutual information is conserved, i.e., $\dot{\mathcal{I}} = 0$ (because $\partial_t \rho_{n+1} = 0$). However, the individual information flows, which are closely related, are not necessarily zero.

Information flow The information flow to X_j from all other d.o.f. is defined as

$$\dot{I}_{\to j} = \iint \ln \frac{\rho_1(x_j)}{\rho_{n+1}(\underline{x})} \partial_{x_j} J_j \,\mathrm{d}\underline{x}. \tag{3.30}$$

While this quantity is widely used to investigate discrete systems, it is less common for time- and space-continuous systems. It should further be noted that there are other notions of information flows and information exchanges, see [289] for an educational overview.

To derive the connection between information flow and mutual information, we start with considering the total derivative of $\mathcal{I}$ from Eq. (3.29), that is,

$$\begin{aligned}\dot{\mathcal{I}} = &\int \underbrace{\partial_t \rho_{n+1}(\underline{x})}_{=(*)} \ln \frac{\rho_{n+1}(\underline{x})}{\rho_1(x_0)...\rho_1(x_n)} \,\mathrm{d}\underline{x} \\ &+ \int \rho_{n+1}(\underline{x}) \left\{ -\frac{\partial_t \rho_{n+1}(\underline{x})}{\rho_{n+1}(\underline{x})} - \frac{\partial_t [\rho_1(x_0)\rho_1(x_1)...\rho_1(x_n)]}{\rho_1(x_0)\rho_1(x_1)...\rho_1(x_n)} \right\} \mathrm{d}\underline{x}.\end{aligned} \tag{3.31}$$

To bring the last expression into another form, we utilize the multivariate FPE (2.14) $\partial_t \rho_{n+1} = -\sum_{j=0}^{n} \partial_{x_j} J_j$ to substitute $(*)$, and find

$$\begin{aligned}\dot{\mathcal{I}} =& \sum_{j=0}^{n} \int \partial_{x_j} J_j(\underline{x}) \ln \frac{\rho_1(x_0)...\rho_1(x_n)}{\rho_{n+1}(\underline{x})} \mathrm{d}\underline{x} - \int \underbrace{\partial_t \rho_{n+1}}_{\to 0} \mathrm{d}\underline{x} \\ &- \int \frac{\rho_{n+1}}{\rho_1(x_0)...\rho_1(x_n)} \underbrace{\partial_t [\rho_1(x_0)...\rho_1(x_n)]}_{\to 0} \mathrm{d}\underline{x} \\ =& \sum_{j=0}^{n} \int \partial_{x_j} J_j(\underline{x}) \ln \frac{\rho_1(x_0)...\rho_1(x_n)}{\rho_{n+1}(\underline{x})} \mathrm{d}\underline{x}.\end{aligned} \tag{3.32}$$

Let us now consider the individual summands. By application of basic properties of the logarithm and the natural boundary conditions, we find

$$\begin{aligned}\int \partial_{x_j} J_j(\underline{x}) \ln \frac{\rho_1(x_0)\rho_1(x_1)..\rho_1(x_n)}{\rho_{n+1}(\underline{x})} \,\mathrm{d}\underline{x} &= \iint \ln \frac{\rho_1(x_j)\rho_1(x_{i\neq j})}{\rho_{n+1}(\underline{x})} \partial_{x_j} J_j \,\mathrm{d}\underline{x} \\ &\quad - \iint \ln \rho_1(x_i) \underbrace{[J_j]_{-\infty}^{\infty}}_{\to 0} \,\mathrm{d}\underline{x}_{\neq j} \\ &= \iint \ln \frac{\rho_1(x_j)}{\rho_{n+1}(\underline{x})} \partial_{x_j} J_j \,\mathrm{d}\underline{x} = \dot{I}_{\to j}. \end{aligned} \tag{3.33}$$

Thus, the change of mutual information is given by the sum over all information flows,

$$\dot{\mathcal{I}} = \sum_{j=0}^{n} \dot{I}_{\to j}. \tag{3.34}$$

The last identity further implies that in steady states, the information flows between all components of the entire network nullify (thus, from an information-theoretical point of view, the network as a whole is "closed"). In Chapter 2, we have mentioned that EQ is associated with vanishing probability current, while a NESS has constant, nonzero J. Combining this with the definition (3.30), one can see that in EQ all individual information flows vanish, whereas *finite information flows are characteristic for a NESS*. This resembles the aforementioned net energy flows (heat and work), which likewise signal nonequilibrium.

In Chapter 9, we will derive rigorous relations between information flows and other thermodynamic quantities. Further, we will show that one can derive a generalized second law valid for the individual mesoscopic subsystems, based on this notion of information flow.

3.6 Previous results, expectations and apparent problems for systems with time delay

After having introduced the basic concepts, focusing on Markovian systems, we now turn to systems subject to time-delayed feedback. We will review some previous results and point out open problems. In addition, we will formulate expectations, apparent technical challenges, and relevant research questions, in this way preparing the considerations presented in Part III.

While we will later in Part III consider both, systems with discrete and distributed time delay, we will here focus on the former case.

3.6.1 Energetics in the presence of delay

We start with considering the energetic perspective. For Markovian systems, we have introduced the framework of stochastic energetics in Sec. 3.2. Form a formal

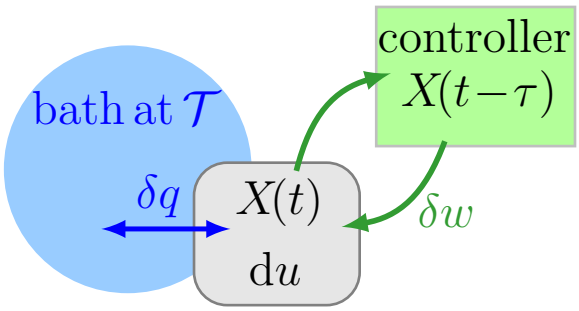

Figure 3.2: *Fluctuating energy (ex-)changes of the system $X(t)$: the controller can perform work δw on the system, heat δq might flow to the bath, and the internal energy of the system might change $\mathrm{d}u$. While in Markovian systems, the direction of the mean heat flow is dictated by the second law (3.21), a control with delay might induce a reversed net heat flow.*

perspective, the incorporation of time-delay delayed is straightforward: the fundamental definitions of the energy (ex-)changes (3.1 - 3.3) are still valid in the very form introduced here, as well as the first law. After all, energy conservation must generally hold.

According to this framework, a position-dependent delay force $F[X(t-\tau)]$, which is *per se* a non-conservative force, is expected to perform work (3.3) on the system (see Fig. 3.2). However, as we have already mentioned in Chapter 2 in the context of probabilistic treatments, there are a significant differences between this history-dependent force and usual time-dependent or constant (non-conservative) Markovian forces. First, a formal difference is related to symmetry breaking and probability currents. Specifically, the feedback force considered here has at each instance in time *on ensemble average* no preferred direction. In other words, it does not break the spatial symmetry of the system, and, therefore, it does not induce a net probability current. In turn, the absence of a probability current is a signature of thermal equilibrium, as is well-known for Markovian processes (see Chapter 2). From this point of view, it is hence not immediately clear whether our feedback force will automatically yield nonequilibrium and introduce finite *average* energy flows.

Furthermore, in Sec. 3.3.1 we have seen that a simple external, position-dependent driving can never induce a negative mean heat flow in a NESS of a Markovian system [see (3.21)]. This is because of the second law, which forbids net energy extraction from a single heat bath in such a stationary situation. However, based on physical intuition, one may guess that it will make a substantial difference that our force is the result of a *control*, which involves something like a measurement and control action/decision[17]. It is a form of driving realizable by some external "intelligence", or controller; different from, lets say, some interaction forces between mutually coupled, mechanical degrees of freedom[18]. In principle, it is well-known that an ("informed") control force can have a very different effect than usual ("uninformed") drivings. For example, control forces may enable *energy extraction of a single heat bath*, as illustrated by the "Maxwell demon" thought experiment [136, 148, 149]. This demon is a fictive creature which can, by track-

[17] We will later in Chapter 9 make this distinction rigorous, by connecting it to associated information flow.

[18] As an example, compare with the (friction) force applied by the surrounding heat bath oscillators derived in Sec. 1.3.1, which illustrates that not every history-dependent force is necessarily a control.

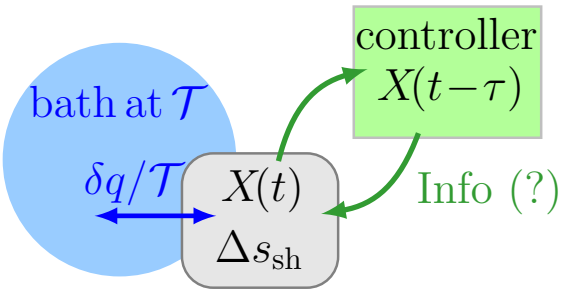

Figure 3.3: *Fluctuating entropy balance of the controlled system: the Shannon entropy s_{sh} may vary upon a change of the PDF or $X(t)$. The medium entropy might change due to a heat flow δq. Due to measurement and control action, we additionally expect an information-theoretic term in the entropy balance. Figure 3.2 shows the corresponding energetic picture. The unknown term "(?)" will be discussed in Sec. 9.5.1.*

ing individual particle velocities and making decisions based on the measurement, "defeat the second law", i.e., extract useful work from a single heat bath. Another realization is a Szilárd engine [148, 150], which involves a position-dependent feedback. One can also consider such devices "information to energy" converter [131]. In our setting, energy extraction of a single heat bath would imply a net *negative heat current*.

These considerations motivate us to investigate whether the feedback inevitably results in a net heat flow (despite the absence of a probability current), and second, whether it is possible to induce a *reversed heat flow* with a time-delayed feedback (which, although being very different from a Maxwell demon or Szilárd engine, likewise involves a measurement operation).

So far, no explicit general expressions for the thermodynamic quantities are worked out for time- and space-continuous systems with a continuously operating time-delayed feedback control. However, for the case of a *linear* delay LE, a couple of earlier studied can be found in the literature. We aim to highlight the theoretical contributions to this field by Rosinberg and Munakata [93, 99, 116, 152]. This previous research has established that delayed feedback indeed induces a net heat flow [99] (despite the absence of a probability current). The delay indeed pushes the system out of equilibrium, and, in the absence of time-dependent drivings, the system may approach a NESS in the long-time limit [99]. Moreover, for underdamped systems, it has been shown that a delayed feedback loop can indeed induce a reversed (negative) heat flow. In contrast, the study [99] for overdamped systems suggests that the delay force can *not* induce a reversed heat flow; but the arguments used there are based on an incomplete entropic treatment (see below). Furthermore, there are two other open problems concerning the NESS energy flows. Specifically, the previous result [99] fails to correctly describe the limit of vanishing delay time. In [99], an offset of the heat rate at $\tau \to 0$ (where the heat flow should nullify) is detected, which is attributed to an Itô vs. Stratonovich inconsistency. Second, the approaches taken in [93, 99, 116, 152] are only applicable to linear systems.

This motivates us to revisit in this thesis the delay-induced steady-state heat flow, investigating linear and *nonlinear* delay systems (Chapter 8).

3.6.2 Entropic description

Now we turn to the entropic treatment of our system in the presence of time-delayed feedback.

First, we aim to revisit the discussion from the previous Section, where we have mentioned the connection to "Maxwell demon"-devices, and the possibility of steady heat extraction from a single heat bath (see Sec. 3.6.1).

The long, historic discussion about the "Maxwell demon" has revealed that, from a formal perspective, the reason for the apparent "violation of the second law" implied by a negative steady-state heat flow, are *additional entropic terms* that need to be taken into account in the entropy balance (3.15) (in addition to the medium EP and the Shannon entropy change, which is on average zero in steady states). In particular, it has been established that measurement, storing, and deleting of information, which are associated with the control, have entropic cost, see Fig. 3.3. This is the essence of the famous principle by Landauer [148, 153]. As a consequence, the total average entropy production $\Delta S_{\rm tot}$ of controlled systems in a NESS may differ from the medium entropy $\Delta S_{\rm m} = \delta Q/\mathcal{T}$ (contrary to the uncontrolled case described above), such that the second law does not impose nonnegativity on $\Delta S_{\rm m}$ alone.

While this statement is generic, how to write down explicit formulae for the additional terms is not at all obvious. Specifically, the incorporation of continuously-operating, time-delayed feedback turns out to be highly nontrivial, as we will see in the following.

It is noteworthy that the treatment is much easier, when the controller only interacts with the system X at some (predefined) *solitary* instances in time. This is the scenario considered in various earlier studies, e.g., [127, 290, 291], a educational example being the feedback ratchet from Ref. [151]. The treatment of such systems is much more straightforward. In particular, one can chop the trajectories in a way that the control actions are at the boundaries only. On these time intervals, one has then a process that is Markovian, and can be treated by the standard approach described at the beginning of this Chapter. For example, one can in principle evaluate the the total EP functional on the short trajectory pieces, and in a second step sum over all of the pieces. In particular, the backward image of this process does not violate causality, contrasting the situation considered here, as will elaborate in the next Section. To emphasize this difference, the type of control considered here was named *non-Markovian feedback control* in Ref. [215], and the considered continuous, stochastic system is denoted "autonomous system" [116, 292]. We aim to stress that for many physical or biological mechanisms such a continuous, non-Markovian feedback constitutes the more realistic model.

3.6.3 The acausality issue

Above we have shown that the entropy balance readily follows from evaluating the logarithm of the ratio of path probabilities of forward and backward process (3.15). One crucial issue in the thermodynamic description of delayed systems, or more

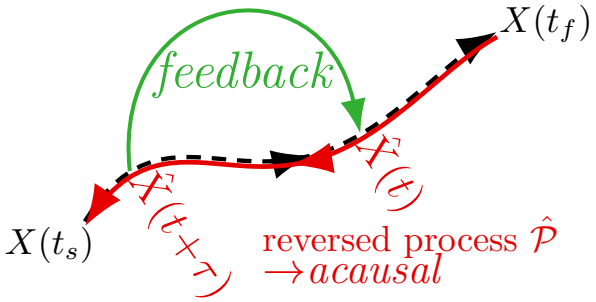

Figure 3.4: *Sketch of a trajectory (dotted line), and its time-reversed picture (red solid line). In the presence of a feedback force, which connects the time instances t and $t-\tau$ (green arrow), the backward process becomes acausal, as it depends on its own future.*

generally non-Markovian systems, is related to the time-reversed process. Recall that the backward process with probability $\hat{\mathcal{P}}$ shall be the time-reversed image of the process, where the driving is carried out in opposite direction in time (and the velocities are flipped).

For systems with *time delay*, where the driving involves $F_{\rm d}[X(t-\tau)]$, the forward process depends on the *previous* trajectory, i.e., the history. Thus, when flipping the time, one needs to think about how to flip the history-dependence: shall the time-reversed process depend on its own history (i.e., the future of the forward process), or shall it depend on its own future?

As argued by Rosinberg and Munakata [93, 116, 293], the time-reversal operation shall imply that the backwards process is not delayed, but *anticipating*, i.e, $F_{\rm d}[\hat{X}(t+\tau)]$. In other words, it is a stochastic process that depends on its own *future*, as sketched in Fig. 3.4. It is important to note that the latter is in principle unforseeable, due to the stochastic nature of the process. When following this idea, the backward process is *acausal*. As we have discussed in Secs. 1.3 and 3.4.1, such a process is in violation with the general formulation of statistical physics. From a technical point of view, the main problem is that the standard path integral formalism is not readily applicable, and instead one encounters a Jacobian $\mathcal{J}[\mathbf{X}]$ which is a highly nontrivial, path-dependent functional, see [93, 116, 152] (at least for any nonlinear system).

At this point it is further interesting to note that delayed differential equations have no unique backward solution[19] (in sharp contrast to ordinary differential equations), even in the absence of noise, see Sec. 1.4.4. We could not find a discussion about how this affects the total EP for delayed systems in the literature, but it might add an interesting argument to the general debate on this problem.

For Markovian systems, the Fokker-Planck equation represents an alternative route to calculate the EP [125]. We will demonstrate this in Chapter 9. However, because there is also no closed FPE for the systems with delay (see Chapter 2 and Part III), this alternative standard approach to the EP is as well not readily applicable.

More generally, entropic concepts appear to be quite difficult for non-Markovian systems. We think that this is indeed not only a minor problem of stochastic ther-

[19] Thus, the log-ratio of the probabilities to observe a certain forward and backward process is not zero in the absence of noise.

modynamics, given that the Markov assumption, which requires a clear timescale separation, is a very strong simplification only valid in a specific limit and inappropriate in many realistic systems. A framework that is restricted to such cases is of somewhat limited use to capture the essence of real-world systems.

Next, we will briefly consider a recent approach to the problem of calculating irreversibility measures alternative to the total EP for in non-Markovian systems.

3.6.4 Short review on Effective thermodynamics

Recent studies explore strategies to define reasonable irreversibility measures for non-Markovian systems based on the trajectories of X alone, similar to the total EP for Markovian systems (3.15). These approaches aim for an "effective thermodynamic" description, e.g., see Ref. [93, 105, 107, 109, 116, 152, 154]. To this end, one may start with a Markovian description of the process. This can be done by considering the probabilities of the joint paths $\mathcal{P}[\mathbf{X}, \mathbf{Y}]$, where $\mathbf{Y}$ could represent the *entire relevant past* of the process $\mathbf{X}$ (and not only the starting point), see Ref. [116] for a more detailed description. Alternatively, $\mathbf{Y}$ could represent appropriate auxiliary variables, for example a variable that represents correlated fluctuations, as in [107]. We will give explicit examples for possible auxiliary variables for different non-Markovian systems in Part III. In the literature, there are different possibilities discussed, how to return to the non-Markovian case (without $\mathbf{Y}$).

As suggested in [116], one can first consider the log-ratio of probabilities of the joint paths, something like

$$R[\mathbf{X}, \mathbf{Y}] = k_\mathrm{B} \ln \frac{\mathcal{P}[\mathbf{X}, \mathbf{Y}]}{\hat{\mathcal{P}}[\hat{\mathbf{X}}, \hat{\mathbf{Y}}]}. \tag{3.35}$$

Then, in a second step, R is integrated over all possible histories (i.e., tracing out $\mathbf{Y}$).

Alternatively, in [107] the authors decide to first integrate out $\mathbf{Y}$ already in the path probabilities (and making some assumptions about the correlations of the initial values of $\mathbf{X}$ and $\mathbf{Y}$), and then consider the log-ratio of

$$\widetilde{R}[\mathbf{X}] = k_\mathrm{B} \ln \frac{\mathcal{P}[\mathbf{X}]}{\hat{\mathcal{P}}[\hat{\mathbf{X}}]}. \tag{3.36}$$

The different strategies yield different irreversibility measures.

We have shown above that the IFT of the total EP can be derived from quite general assumptions. Indeed, the IFT is quite robust and also holds for other functionals that are similar to the total EP (this is also true for some other fluctuation theorems, see [125]). Hence, it is not too surprising that also effective irreversibility measures like R, or an integral over R (where $\mathbf{Y}$ is traced out), or $\widetilde{R}$, fulfill an IFT. The fulfillment is explicitly shown in [93, 107, 116]. Hence, the existence of some fluctuation theorems alone is not a sufficient criterion to confirm the meaningfulness of an reversibility measure.

While these effective thermodynamics approaches are interesting and promising, they still comprise many open questions and ambiguities. For example: at which

point should one marginalize, i.e., trace out $\mathbf{Y}$? Besides conceptual issues, there are also technical challenges. First, a generalization to other types of memory is not straightforward. So far, many studies are focused on a special memory kernel, often exponential or delta-correlated memory; or are only available for non-Markovianity introduced by colored noise. Second, the generalization to *nonlinear* systems is extremely difficult. For example, the approach in [107] relies on the Green's function method, which is only applicable for linear systems. Also, the entropy-like functional from [93, 116], could so far not be calculated for any nonlinear system. This is somewhat problematic, as in real-world systems, nonlinear models are often necessary to capture many important effects [196]. It is a great challenge for future research in the field of stochastic thermodynamics to sort out this issue.

Our approach to the problem (pursued in Chapter 9) In contrast to approaches which try to define meaningful irreversibility measures based on $\mathbf{X}$ alone (like the "effective thermodynamics" descriptions), we will explore a somewhat different strategy in Chapter 9, in line with the earlier works [117, 118]. In particular, we explicitly keep all relevant d.o.f. (i.e., consider the full EP of the joint process $\{\mathbf{X}, \mathbf{Y}\}$), and see what we can learn from that. This approach allows us to use the standard definition of total EP, and is in a sense more general w.r.t. the mathematical details of the control (memory kernel, nonlinear forces). In turn, it makes crucial assumptions on the physical picture behind the control, and is, in this sense, much more restrictive.

From the viewpoint of effective thermodynamics, one can also regard our investigation as a first step. Building up on our findings, one might, in a second step, construct again an effective thermodynamics, by some sort of marginalization. This is beyond the scope of this thesis, but provides an interesting perspective for future research in this direction.

Moreover, without further reasoning, the approach employed in Chapter 9 is appropriate for the following physical picture: Assume that the delayed force stems from physical *external subsystems*, which generate the feedback loop and interacts with the particle. We call them "external", as they are only implicitly accounted for in our original non-Markovian description [in the form of a delayed force in the LE (1.42)]. One can also say that they are *hidden*. These "external subsystems" might belong to a specifically designed agent, a *feedback controller*, or, more generally, could be a part of a more complex biological system (for example part of a chemical reaction network). In any case, we assume that this external system as well consists of physical components, which cannot perform forces without energetic and entropic cost, i.e., we assume that the external agent is a physical system. Furthermore, we assume that these external subsystems interact with the particle in an instantaneous (Markovian) manner, (e.g., the involved measurement operation is instantaneous). The delay comes through a coupling of multiple subsystems to each other in a specific way (in the form of a unidirectional ring).

We will explain all of this in detail in Chapter 9. There we will show that such a setup is only possible with *non-reciprocal* coupling among the controller subsystems.

Role of hidden degrees of freedom Indeed, it is well-known that *hidden degrees of freedom*, i.e., physical subsystems which interact with the studied systems but are not explicitly accounted for in the employed model, can have a significant impact on the EP, see for example [294, 295]. This provides an additional motivation for the study discussed in Chapter 9.

From a more general point of view, this problem is connected to the question of correct coarse-graining procedures, which, ideally, preserve as much of the thermodynamic properties as possible. We will only very briefly comment on this issue in Part III, but refer the interested Reader to [125, 126, 154, 294–298].

3.7 Side note: Active particles & Non-reciprocal interactions

By investigating the random trajectories of a feedback-controlled, non-Markovian particle, one notices differences to the trajectories of a normal, passive colloidal system subject to e.g. external "uninformed" forces, $F[X,t]$. This is for example evident by the correlation functions, which reveal the memory in the system (compare Figs. 1.4 and 5.3). However, the differences become much more evident, when we consider the thermodynamic properties (see Part III). An important distinction is the nonexistence of a thermal equilibrium state. As we have already mentioned before, one may thus consider the feedback-controlled colloids as a "novel" type of particle, in comparison to the passive particles usually studied in traditional statistical mechanics. Indeed, there is another, yet very different "novel" type of particle, which is nowadays studied by many researchers by means of statistical physics and specifically stochastic thermodynamics. The latter have been denoted as *active systems*.

Active systems A historic curiosity is that Robert Brown in his original paper from 1827 [161], in which Brownian motion was described for the first time, speculated that the observed motion of the pollen might be due to their liveliness, and called them "active molecules". He therefore dedicated his investigation mainly on how the motion is different between a living and dead plant, e.g., by burning them. Of course, the observed stochastic motion is also present in the pollen of dead plants, and is, in this case, independent of this question.

In contrast, the stochastic motion of bacteria, sperm cells, and many other, actually living objects is indeed very different from *passive Brownian motion* (of, e.g., a pollen, or dead bacteria), because these systems actively swim, i.e., propel themselves. The *self-propulsion* is possible because these organisms convert other forms of energy, like chemical energy stored in ATP, into motion (just as we humans do). So far, we have introduced the Langevin equation as a model for passive particles who may be subject to external forces (e.g., feedback forces). Besides this, Langevin equations can indeed also be used to model such an active motion.

Active motion is often characterized by *persistence* [19–21]. In regard to the diffusive behavior this means, it has a ballistic regime on intermediate timescales,

while the motion is typically like normal diffusion on very short and on very long timescales. This behavior is described by all of the three common models [299]: active Brownian particles, active Ornstein-Uhlenbeck particles, and run-and-tumble motion.

Besides the aforementioned examples, such active colloids can also be individual parts of a larger organism, like the hair bundles of auditory cells [300, 301]. Further, they don't have to be (parts of) living creatures, but can also be artificial, an obvious example being robots. Another well-know realization are Janus particles with unihemispherical coating [302], or vibrating screws or rods [303]. Studying the statistical properties of active motion is subject of ongoing research in the field of *active matter* [304].

A defining criterion of active particles is an inherent (independent) energy input [18, 102, 103, 305]. Thus, these systems are intrinsically out of equilibrium. While this statement is generic, the development of a thermodynamic description of active systems is subject of an ongoing debate, and the objective of numerous latest research articles in the field of stochastic thermodynamics, see, e.g., [104–108, 302, 305–310]. Here, key problems lie in the search for appropriate definitions of irreversibility measures and entropy production, and the interpretation of heat flux and work. A further open question is the time-reversal parity of the involved variables [105], which is relevant for the total EP functional (3.12). These problems are closely connected to a more general question in active matter, that is, the search for suitable models of active systems.

3.7.1 Active Ornstein-Uhlenbeck particles

The simplest and therefore very popular model is the *Active Ornstein-Uhlenbeck particle* (AOUP) [105, 107, 109, 307, 308]. It can be expressed using two d.o.f. following the equations

$$\gamma\dot{X}(t) = Y(t) + F + \sqrt{2k_\mathrm{B}\mathcal{T}\gamma}\,\xi(t), \tag{3.37a}$$

$$\tau_\mathrm{p}\dot{Y}(t) = -Y(t) + \xi_\mathrm{p}(t), \tag{3.37b}$$

with the two, zero-mean white noises: $\langle\xi(t)\xi(t')\rangle = \delta(t-t')$, $\langle\xi_\mathrm{p}(t)\xi_\mathrm{p}(t')\rangle = \delta(t-t')$, $\langle\xi(t)\xi_\mathrm{p}(t')\rangle = 0$. Thereby, Y is a d.o.f. that models the effect of the flagella or a bacterium, or the asymmetric flow field around a Janus colloid, in a very simplistic, one-variable manner. However, X which corresponds to the particle position, is the d.o.f. of main interest. One can thus also employ a one-variable description focusing on X, obtained by *projecting*[20] the second equation with the d.o.f. Y into the X equation. To this end, one formally solves the equation for Y [assuming $Y(0) = 0$], which yields $Y(t) = \int_0^t e^{-(t-t')/\tau_\mathrm{p}}\xi_Y(t')\mathrm{d}t' = \nu(t)$. Then one substitutes the result into the first equation of (3.37) obtaining

$$\gamma\dot{X}(t) = \nu(t) + F + \xi(t),$$

$$\langle\nu(t)\nu(t+T)\rangle = \int_0^t e^{-(t-t')/\tau_\mathrm{p}}\xi_\mathrm{p}(t')\mathrm{d}t' =: C_\nu(T). \tag{3.38a}$$

[20] Analogously to the projection in the derivation of Mori and Zwanzig, see Sec. 1.3.1, and our approach in Chapter 5.

We have renamed Y to ν, to emphasize the different characters of the variable, which is a dynamical d.o.f. in (3.37) and a colored noise in (3.38). This is a simple, non-Markovian description of an active particle's motion with persistence.

It is noteworthy that this very model is in fact also used to describe a passive (i.e., "dead") colloid in an *active bath*, i.e., a heat bath consisting of self-propelled particles. However, here we are going to take the aforementioned, alternative viewpoint of an active colloidal swimmer in a passive bath.

Both Eqs. (3.38) and (3.37) equally describe the dynamics of a *microswimmer* at position X. The auxiliary d.o.f. Y in (3.37) introduces correlated fluctuations in the X dynamics with characteristic timescale $1/\tau_{\mathrm{p}}$, and in this way models the effect of the flagella of a bacterium, or the asymmetric flow field around a Janus colloid, responsible for the propulsion force on the swimmer [302, 311]. Thus, here the memory models the persistence of the motion. Importantly, in (3.38) there is no direct connection between memory in noise correlations and in the friction, hence, no FDR (see Sec. 1.3.1). This signals *nonequilibrium*, as one shall indeed expect for an active system. However, the broken FDR alone does not quantify the distance from equilibrium, nor does it provide further insights into the thermodynamic properties of this nonequilibirum system. This motivates us to take a closer look at this model as we do in Part III. Another motivation is the connection between active matter and feedback-controlled systems.

3.7.2 Connection between active matter and time-delayed systems

We have already mentioned one important, common feature of active and feedback systems, i.e., they are intrinsically out of equilibrium already on the level of individual constituents. Conventionally, active systems are mostly associated with "energy consumption" of individual constituents [102, 103, 310], and are hence rather studied from an energetic perspective (or from a purely dynamical perspective), while feedback-controlled systems are considered to be "information fueled" [151]. Therefore, the second type of system is usually studied from an entropic and information-theoretical perspective.

However, it has already been recognized in earlier literature that also active systems involve nontrivial information flows [60, 107]. On the other hand, one clearly expects that a feedback controller "consumes energy" in order to operate. Thus, these systems might not be too different after all. Recent experiments and theoretical works indeed expose the connections between both types of systems, see e.g. [38, 50, 59, 66]. It is also noteworthy, that some active systems, like *E.coli* bacteria, use *intrinsic feedback loops* to create a motion towards food sources, for example chemotaxis [59, 312, 313].

In fact, there is another –less obvious– shared feature of active and feedback systems, that is, the presence of memory. While the discussion of memory, or delay, in feedback loops has a long tradition [1–3, 9, 88, 93, 116, 152] (where it stems e.g. from finite signal processing speed), this viewpoint is less common in the context of active motion. However, the presence of history-dependence becomes apparent in

the form of *persistence* of active swimmers [19–21] as one can see above in Eq. (3.38). Also the viscoelastic properties of active gels [18], or the giant Kovacs-like memory effect in the Vicsek model [111] give a hint to the importance of memory in active systems.

We believe that a fundamental understanding of the two classes of systems from *both* perspectives, the energetic and the entropic/information-theoretic one, is very important for the thermodynamic understanding and their statistical properties. An apparent advantage of combining these fields are the various established results for the collective behavior of active systems on larger scales [20, 38]. In this spirit, we investigate an approach which naturally unifies both perspectives, as we will discus in detail in Chapter 9. A key ingredient is *non-reciprocal coupling*.

3.7.3 Non-reciprocal interactions

The set of equations (3.37) defines a two-dimensional network with *unidirectional* coupling, because X is influenced by Y, whereas Y does not "see" X, i.e.,

$$\begin{aligned} X &\leftarrow Y, \\ X &\not\leftarrow Y. \end{aligned} \tag{3.39}$$

This unidirectional coupling is a special type of *non-reciprocal* (NR) coupling, which denotes *interactions that cannot be represented by an interaction Hamiltonian and violate Newtons third law actio = reactio.*

Such interactions also arise in intelligent systems [314], like pedestrian dynamics, or in collections of self-driving cars, but can also be found in chemical [155] and quantum systems [315–319]. At this point one may also note that in the community of nonlinear dynamics and network science, studying the interplay between non-reciprocal (symmetry-broken) coupling and dynamical states is already a well-established research field. For example, the existence of chimera states [320], a special type of clustering, was linked to symmetry-broken coupling [321], and shown to persist in the presence of discrete delay [47] and Gamma-distributed memory [235].

As we have seen in Chapter 1, the mechanical laws are the foundations of the stochastic models. In fact, various derivations in statistical physics rely on them. Thus, in the absence of Newton's third law, one needs to revisit many well-known results. Only recently, researchers have started to investigate the consequences of NR coupling on statistical properties of many particle systems [110, 155]. However, the thermodynamic implications of the non-reciprocity have, to the best of our knowledge, not been considered in earlier literature.

In Chapter 9, we will examine in detail the heat and information flows induced by NR coupling, considering the AOUP model, but also more general types of network topologies. For example, the very same set of equations as (3.37) with reversed unidirectional coupling (i.e., $X \rightarrow Y$ and $X \not\leftarrow Y$), was recently suggested as a model for a *cellular sensor* [322, 323]. More importantly in the context of this thesis, systems with time-delayed feedback can also be represented as networks with NR coupling among X and auxiliary variables.

In the following two Parts of this thesis, we present novel results for the probabilistic description (Part II) and thermodynamic treatment (Part III). Therewith, we aim for two theoretical concepts, which are cornerstones of statistical physics, already established for passive particles, which are not subject to feedback ("uninformed" particles).

Part II

Probabilistic descriptions for systems with time delay

"*[T]he true Logic of this world is the Calculus of Probabilities*"

– James Clerk Maxwell, 1871 [324]

This Part II is dedicated to the probabilistic description of time-delayed stochastic systems. In Chapter 2, we have already introduced Fokker-Planck equations, and mentioned that for time-delayed systems, this type of description is given by an *infinite hierarchy*. Here we will elucidate this issue in more detail and explore different strategies to tackle the apparent problems in the probabilistic treatment arising due to this hierarchical structure. We will further explicitly inspect higher members of this hierarchy, which have so far received only little attention in earlier research.
On the one hand, we will review earlier derivation strategies that have been used to derive the first, not self-sufficient member, and see if they can be extended towards the higher members. Only one of these approaches results in a complete FP description in the form of the infinite hierarchy, namely the one basing on *Novikov's theorem*. We use this theorem to explicitly derive the second member, and find that it involves some unknown functional derivatives. Moreover, we will introduce a new approach via a Markovian embedding, and, in this way, derive a complete probabilistic description with an alternative representation of the hierarchy.
On the other hand, we will discuss and compare different ways to find closed, approximate equations. We will consider approaches from earlier literature, and further develop new ones, particularly, the force-linearization closure, and an approximation on the level of the second FP hierarchy member.

4 | Infinite Fokker-Planck hierarchy

In this Chapter, we will examine the infinite Fokker-Planck hierarchy. We will explicitly consider the higher members, which play a crucial role to understand the non-Markovian dynamics. Later in Chapter 5, we will present a new derivation of the FP description using a Markovian embedding technique [97, 117–121]. To better understand the technical and conceptual difference of our approach, we will here also review some earlier work, in particular, Refs. [87, 325, 326].

4.1 Derivation of Fokker-Planck hierarchy from Novikov's theorem

We begin with reviewing the approach [87] based on Novikov's theorem (4.3). While [87] only considers the first member, we here further present a generalization towards the higher members of the hierarchy. We have already reported the resulting equation for the two-time PDF[1] $\rho_2(x,t;x_\tau,t-\tau)$ in Ref. [1].

Derivation of the second member We start with the definition $\rho_2(x,t;x_\tau,t-\tau) = \langle\delta[x-X(t)]\delta[x_\tau - X(t-\tau)]\rangle$ [as given in Part I in Eq. (2.4)], and consider its temporal derivative

$$\begin{aligned}\frac{\partial}{\partial t}\rho_2 =& \left\langle \left\{\frac{\partial}{\partial t}\delta[x-X(t)]\right\}\delta[x_\tau - X(t-\tau)]\right\rangle + \left\langle \delta[x-X(t)]\,\frac{\partial}{\partial t}\Big\{\delta[x_\tau - X(t-\tau)]\Big\}\right\rangle \\ =& -\frac{\partial}{\partial x}\left\langle \delta[x-X(t)]\delta[x_\tau - X(t-\tau)]\,\frac{\partial X(t)}{\partial t}\right\rangle \\ & -\frac{\partial}{\partial x_\tau}\left\langle \delta[x-X(t)]\delta[x_\tau - X(t-\tau)]\,\frac{\partial X(t-\tau)}{\partial t}\right\rangle. \end{aligned} \tag{4.1}$$

[1] We recall that a list of all abbreviations and symbols is given on pages VIII and IX, respectively.

Substituting $\partial X(t)/\partial t$ and $\partial X(t-\tau)/\partial t$ with the LE (1.42), we obtain

$$\frac{\partial}{\partial t}\rho_2(x,t;x_\tau,t-\tau) = -\frac{\partial}{\partial x}\Big\langle \delta[x-X(t)]\,\delta[x_\tau - X(t-\tau)]\,\gamma^{-1}F\left[X(t),X(t-\tau)\right]\Big\rangle \tag{A1}$$

$$-\frac{\partial}{\partial x_\tau}\Big\langle \delta[x-X(t)]\,\delta[x_\tau - X(t-\tau)]\,\gamma^{-1}F\left[X(t-\tau),X(t-2\tau)\right]\Big\rangle \tag{A2}$$

$$-\sqrt{2D_0}\,\frac{\partial}{\partial x}\Big\langle \delta[x-X(t)]\,\delta[x_\tau - X(t-\tau)]\,\xi(t)\Big\rangle \tag{B1}$$

$$-\sqrt{2D_0}\,\frac{\partial}{\partial x_\tau}\Big\langle \delta[x-X(t)]\delta[x_\tau - X(t-\tau)]\,\xi(t-\tau)\Big\rangle. \tag{B2}$$

Let us now try to simplify the terms on the r.h.s. The terms (A1, A2) solely involve deterministic forces (and no random Langevin forces). Hence, the ensemble average can readily be calculated, yielding

$$\begin{aligned}(\text{A1}) &= -\frac{\partial}{\partial x}\iint_\Omega \delta(x-y)\,\delta(x_\tau-y_\tau)\,\gamma^{-1}F(y,y_\tau)\rho_2(y,t;y_\tau,t-\tau)\mathrm{d}y\mathrm{d}y_\tau \\ &= -\frac{\partial}{\partial x}\left\{\gamma^{-1}F(x,x_\tau)\rho_2(x,t;x_\tau,t-\tau)\right\}\end{aligned}$$

and

$$(\text{A2}) = -\frac{\partial}{\partial x_\tau}\left\{\int_\Omega \gamma^{-1}F(x_\tau,x_{2\tau})\rho_3(x,t;x_\tau,t-\tau;x_{2\tau},t-2\tau)\mathrm{d}x_{2\tau}\right\}.$$

In order to evaluate (B1, B2), we need to deal with the correlations between the noise and the particle position, X. To this end, we introduce the functional of the random processes

$$\Lambda[\xi] := \delta[x-X(t)]\,\delta[x_\tau - X(t-\tau)], \tag{4.2}$$

which emphasizes by its form that X can be considered a functional of the random process ξ. At this point, we can employ Novikov's theorem.

Novikov's theorem Novikov's theorem connects the variational derivative of any functional $\Lambda[\xi]$ w. r. t. the Langevin force ξ with the correlations between that functional and the Langevin force. For the zero-mean Gaussian white noise with $\langle\xi(t)\xi(t')\rangle = \delta(t-t'),\ \forall t,t'$, it reads [327]

$$\left\langle\, \Lambda[\xi]\xi(t)\,\right\rangle = \left\langle \frac{\delta\Lambda[\xi]}{\delta\xi(t)}\right\rangle. \tag{4.3}$$

A short derivation of (4.3) is given in the Appendix A.2 by means of path integrals. Using this theorem, and simple rules for variational derivatives, we find

$$\begin{aligned}
\text{(B1)} &= -\sqrt{2D_0}\,\frac{\partial}{\partial x}\left\langle \frac{\delta\Lambda[\xi]}{\delta\xi(t)}\right\rangle \\
&= -\sqrt{2D_0}\,\frac{\partial}{\partial x}\left\{\left\langle \frac{\delta\Lambda[\xi]}{\delta X(t)}\frac{\delta X(t)}{\delta\xi(t)}\right\rangle + \left\langle \frac{\delta\Lambda[\xi]}{\delta X(t-\tau)}\underbrace{\frac{\delta X(t-\tau)}{\delta\xi(t)}}_{\to 0}\right\rangle\right\}, \\
\text{(B2)} &= -\sqrt{2D_0}\,\frac{\partial}{\partial x_\tau}\left\langle \frac{\delta\Lambda[\xi]}{\delta\xi(t-\tau)}\right\rangle \\
&= -\sqrt{2D_0}\,\frac{\partial}{\partial x_\tau}\left\{\left\langle \frac{\delta\Lambda[\xi]}{\delta X(t)}\frac{\delta X(t)}{\delta\xi(t-\tau)}\right\rangle + \left\langle \frac{\delta\Lambda[\xi]}{\delta X(t-\tau)}\frac{\delta X(t-\tau)}{\delta\xi(t-\tau)}\right\rangle\right\}.
\end{aligned}$$

The second term of (B1) vanishes since it is *acasual*: The particle position cannot depend on the future noise. We have discussed this causality argument in Secs. 1.3 and 3.4. We will apply it several times in the following. Further, using the integral form of the LE (1.42), i.e., $X(t') = \int_0^{t'} \left\{\gamma^{-1}F\left[X(t'), X(t'-\tau)\right] + \sqrt{2D_0}\xi(t')\right\}\mathrm{d}t' + X(0)$, we generally find

$$\begin{aligned}
\frac{\delta X(t-\tau)}{\delta\xi(t-\tau)} = \frac{\delta X(t)}{\delta\xi(t)} &= \frac{\delta}{\delta\xi(t)}\int_0^t \left[\frac{F[X(s), X(s-\tau)]}{\gamma} + \sqrt{2D_0}\xi(s)\right]\mathrm{d}s \\
&= \sqrt{2D_0}\int_0^t \delta(s-t)\,\mathrm{d}s = \sqrt{\frac{D_0}{2}},
\end{aligned} \tag{4.4}$$

independently of the specific form of F. This means, the effect of varying the noise realization on the resulting path X, is given by the diffusion constant alone, and is not influenced by the energy landscape. This finding severely simplifies (B1) and (B2).

On the contrary to (4.4), the functional derivative of $X(t)$ w. r. t. the *earlier* noise $\xi(t-\tau)$, which also appears in (B2), cannot be treated in this manner. In particular, evaluating $\delta X(t)/\delta\xi(t-\tau)$ requires the formal solution of the delayed LE (1.42), which is not known for general F (an exception being linear systems, which we discuss in Sec. 4.2.1). For general (nonlinear) cases, we need to keep this term explicitly in the equation.

For the remaining terms in (B1, B2), we can immediately write down

$$\frac{\delta\Lambda[\xi]}{\delta X(t)} = -\frac{\partial}{\partial x}\delta[x - X(t)]\delta[x_\tau - X(t-\tau)] \tag{4.5}$$

$$\frac{\delta\Lambda[\xi]}{\delta X(t-\tau)} = -\frac{\partial}{\partial x_\tau}\delta[x - X(t)]\delta[x_\tau - X(t-\tau)], \tag{4.6}$$

which finally yields the FPE[2]

$$\begin{aligned}\frac{\partial}{\partial t}\rho_2(x,t;x_\tau,t-\tau) = &-\frac{\partial}{\partial x}\left[\frac{1}{\gamma}F(x,x_\tau)\rho_2(x,t;x_\tau,t-\tau)\right]\\ &-\frac{\partial}{\partial x_\tau}\left[\int_\Omega \frac{1}{\gamma}F(x_\tau,x_{2\tau})\rho_3(x,t;x_\tau,t-\tau;x_{2\tau},t-2\tau)\mathrm{d}x_{2\tau}\right]\\ &+D_0\left[\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial x_\tau^2}\right]\rho_2(x,t;x_\tau,t-\tau)\\ &+\sqrt{2D_0}\,\frac{\partial^2}{\partial x_\tau \partial x}\left[\frac{\delta X(t)}{\delta\xi(t-\tau)}\Big|_{\substack{X(t)=x\\X(t-\tau)=x_\tau}}\rho_2(x,t;x_\tau,t-\tau)\right]. \quad (4.7)\end{aligned}$$

Equation (4.7) is the second member of the Fokker-Planck hierarchy for a system with a delayed (nonlinear) force $F(x,x_\tau)$ and additive Gaussian white noise. It involves the three-time PDF, indicating the emergence of an infinite hierarchy of FPEs. In contrast to the first member, it further contains a variational derivative which needs to be separately calculated for the specific force. In Sec. 4.2.1, we give explicit results for these missing terms specifically for linear systems [where $F(x,x_\tau)$ is linear in x and x_τ], and further elucidate the meaning of this term.

We note that Eq. (4.7) resembles a Markovian FPE of a two-dimensional (or a two-particle) system in the variables x and x_τ, with the exception of the diffusion term which involves derivatives w. r. t. *both* variables. This term induces in the solutions correlations between particle positions at different times, which we describe in detail in Sec. 4.2.

First member from Novikov's theorem The first FP member can be derived in an analogous way, as shown in Ref. [87]. To this end, one starts with the temporal derivative of ρ_1, instead of ρ_2 [like in Eq. (4.1)]. Moreover, as one should expect, integrating (4.7) over the delayed variable, i.e., $\int_\Omega \mathrm{d}x_\tau$ on both sides of (4.7), as well yields the first member of the FP hierarchy, confirming the consistency of the derived equations. In order to show this, one may use $\int_\Omega \rho_2(x,x_\tau)\mathrm{d}x_\tau = \rho_1(x)$, and apply the natural boundary conditions.

4.1.1 Alternative approach with two time arguments

An alternative derivation of the FPE for the one-time PDF ρ_1 was introduced in [325, 326]. We will here review and extended it towards the next member of the hierarchy. This derivation bases on a description with *two* time arguments. In particular, an integer $j = 0, 1, 2, ..$ is counting the number of intervals of length τ that have passed since the beginning $t = 0$, and a continuous variable $z \in [0, \tau]$ measures the time within the current interval.

[2]We have assumed that the variational derivative w. r. t. the delayed noise ξ is deterministic, which should generally be the case for additive white noise considered here.

Then, auxiliary phase-space variables $X_{j\tau}(z) := X(z+j\tau)$ are introduced, which, by construction, all follow a LE of the identical form

$$\mathrm{d}X_{j\tau}(z) = F[X_{j\tau}(z), X_{(j-1)\tau}(z)]\mathrm{d}z + \sqrt{2D_0}\ \mathrm{d}W_{j\tau}(z). \tag{4.8}$$

The steady-state dynamics can be accessed if $j \to \infty$. A sketch of the construction of the auxiliary variables is given in Fig. 4.1, upper panel. In the probability space belonging to the phase space $\{X_{0\tau}(z), ..., X_{N\tau}(z)\}$, the process is Markovian, and the corresponding FPE is given by

$$\begin{aligned}\partial_z \rho_{N+1} = &\sum_{j=1}^{N} \left[-\partial_{x_{j\tau}} \left\{F(x_{j\tau}, x_{(j-1)\tau})\rho_{N+1}\right\} + D_0 \partial^2_{x_{j\tau}} \rho_{N+1}\right] \\ &- \partial_{x_{0\tau}} \left[F(x_{0\tau}, x_{1\tau})\rho_{N+1}\right] + D_0 \partial^2_{x_{0\tau}} \rho_{N+1},\end{aligned} \tag{4.9}$$

with $\rho_{N+1}(x_{0\tau}, x_{1\tau}, ..., x_{N\tau}, z) = \langle \delta[x_{0\tau} - X(z)]\delta[x_{1\tau} - X(z+\tau)]...\delta[x_{N\tau} - X(z+N\tau)]\rangle$. By integrating the equation over the entire domain of all variables but the first one, this many-variable FPE can be used to derive the well-known FPE for ρ_1. By integrating over all variables but the first and last one, further yields a FPE for ρ_2, reading

$$\partial_t \rho_2 = -\partial_x \left[F(x, x_\tau)\rho_2\right] - \partial_{x_\tau} \int_\Omega \left[F(x, x_\tau)\rho_3\right] \mathrm{d}x_{2\tau} + D_0 \left[\partial^2_x + \partial^2_{x_\tau}\right] \rho_2, \tag{4.10}$$

with $\rho_2(x, t; x_\tau, t-\tau)$ and $\rho_3(x, t; x_\tau, t-\tau; x_{2\tau}, t-2\tau)$. As opposed to the corresponding equation from Novikov's theorem, Eq. (4.7), this equation lacks of the term with the unknown functional derivative.

We think that this is because the stochastic set of Eqs. (4.8) is, in fact, only equivalent to the delayed process, if one further imposes *additional constraints* $X_j(\tau) = X_{j+1}(0)$, $\forall j$, ensuring that the end point of the stochastic trajectories on each $[j\tau, (j+1)\tau]$ matches the starting point on $[(j+1)\tau, (j+2)\tau]$. Translating these constraints to the level of PDFs is a non-trivial task on its own. However, the comparison between Eq. (4.7) from Novikov's theorem and (4.10) suggests that the terms with the unknown functional derivatives in (4.7) indeed represent such terms ensuring the additional constraints on the trajectories.

While appearing quite elegant, this two-time arguments approach seems not to be appropriate to derive higher members, due to the need for additional terms which do not come automatically.

4.2 Exact probabilistic solutions for linear systems with time delay

Now we again turn to the purely linear case, which we have introduced in Sec. 1.4.5, where more analytical results can be obtained. Importantly for the present context, the missing terms in the second member of the FP hierarchy (4.7) can be analytically calculated for this case, as well as the steady state solutions, i.e., the one-time PDF and the two-time PDF.

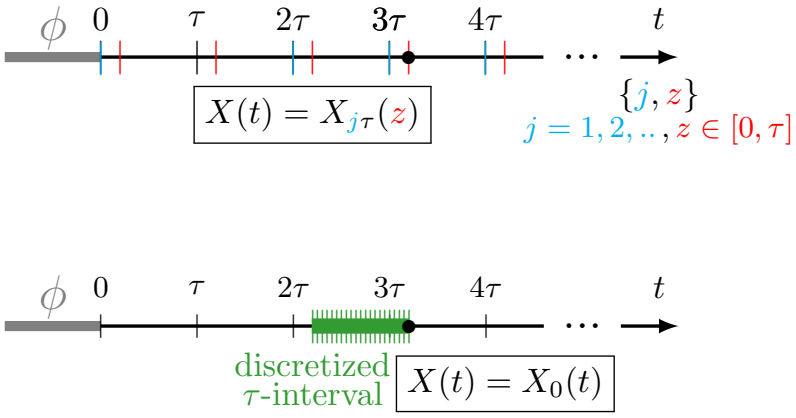

Figure 4.1: *Sketch of the different ways to introduce auxiliary variables, upper panel: within the approach based on two time arguments [325, 326] (mentioned here), lower panel: within the Markovian embedding approach discussed later in Sec. (5.1). These figures stem from [3].*

4.2.1 Derivation of the second member of the Fokker-Planck hierarchy

We aim to calculate the variational derivatives of the particle position with respect to the delayed Langevin force, in order to specify the FPE for ρ_2 from Eq. (4.7) to the linear case

$$F(x, x_\tau) = -\alpha x - \beta x_\tau. \tag{4.11}$$

The corresponding LE is given in (1.49). For this purpose, we will need a formal solution of the delayed Langevin equation. To obtain the latter, we will employ the method of steps.

Method of steps The linear LE with delay can be formally solved in an iterative manner, considering each interval $[(n-1)\tau, n\tau]$ after another, with increasing $n \in \mathbb{N} \cup \{0\}$. This technique is well-known for delay differential equations and is called *method of steps* [85]. We call the formal solution on the nth interval $\mathcal{Y}_n$, such that

$$X(t) = \mathcal{Y}_n(t), \quad \text{for} \quad t \in [(n-1)\tau, n\tau].$$

We recall that, in contrast to deterministic equations, these solutions explicitly contain the random force ξ, hence, they can only be regarded as "formal" solutions. Please further note that the intervals are all closed, such that they overlap at the discrete times $t = n\tau$. However, this is not problematic because $\mathcal{Y}_n(n\tau) = \mathcal{Y}_{n+1}(n\tau)$ [see Eq. (4.13)].

Starting with the initial condition ϕ on the zeroth interval $t \in [-\tau, 0]$

$$n = 0: \quad \mathcal{Y}_0(t) = \phi(t), \quad \forall t \in [-\tau, 0], \tag{4.12}$$

we find for the first interval: $\mathrm{d}\mathcal{Y}_1(t)/\mathrm{d}t = -\alpha\mathcal{Y}_1(t) - \beta\phi(t-\tau) + \sqrt{2D_0}\,\xi(t)$. Since ϕ is a given function, this is a stochastic *non-delayed* differential equation with the

formal solution

$$n=1: \quad \mathcal{Y}_1(t) = \phi(0)e^{-\alpha t} + \int_0^t e^{-\alpha(t-t')}[\sqrt{2D_0}\,\xi(t') - \beta\phi(t'-\tau)]\mathrm{d}t'.$$

Hence $X(t) = \mathcal{Y}_1(t)$, $\forall t \in [0,\tau]$. In the same manner, we can express $\mathcal{Y}_2$ with the help of $\mathcal{Y}_1$, and so on. Consequently, given that the solution $\mathcal{Y}_n$ is known, the non-delayed stochastic differential equation for $\mathcal{Y}_{n+1}$ reads:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{Y}_{n+1}(t) = -\alpha\mathcal{Y}_{n+1}(t) - \beta\mathcal{Y}_n(t-\tau) + \sqrt{2D_0}\,\xi(t),$$

with the (iterative) formal solution [85]

$$\mathcal{Y}_{n+1}(t) = \mathcal{Y}_n(n\tau)e^{-\alpha(t-n\tau)} + \int_{n\tau}^t e^{-\alpha(t-t')}[\sqrt{2D_0}\,\xi(t') - \beta\mathcal{Y}_n(t'-\tau)]\mathrm{d}t'. \tag{4.13}$$

Variational derivatives With the help of Eq. (4.13) we can readily calculate the variational derivative of the particle position w. r. t. the instantaneous, or w. r. t. the delayed Langevin force, $\delta X(t)/\delta\xi(t)$ and $\delta X(t)/\delta\xi(t-\tau)$, which appear in Eq. (4.7). We have already calculated the first one in (4.4) for the general case of (nonlinear) forces. We nevertheless reconsider it here, as a consistency check.

To this end, we write the arbitrary time t as $t = n\tau + \Delta t$, with $\Delta t < \tau$ without loss of generality. Second, we plug in the formal solution $X(t) = \mathcal{Y}_{n+1}(n\tau + \Delta t)$ for $t \in [n\tau, n\tau+\tau]$, such that

$$\begin{aligned}\frac{\delta X(t)}{\delta\xi(t)} = \frac{\delta\mathcal{Y}_{n+1}\big(n\tau+\Delta t\big)}{\delta\xi(n\tau+\Delta t)} &= \underbrace{\frac{\delta\mathcal{Y}_n\big(n\tau\big)}{\delta\xi(n\tau+\Delta t)}}_{\to 0} e^{-\alpha\Delta t} \\ &+ \int_{n\tau}^{n\tau+\Delta t} e^{-\alpha(n\tau+\Delta t-t')}\left[\sqrt{2D_0}\,\underbrace{\frac{\delta\xi(t')}{\delta\xi\big(n\tau+\Delta t\big)}}_{(\times)} - \beta\,\underbrace{\frac{\delta\mathcal{Y}_n(t'-\tau)}{\delta\xi\big(n\tau+\Delta t\big)}}_{\to 0}\right]\mathrm{d}t'. \end{aligned} \tag{4.14}$$

The first and last terms both vanish since they are *acausal*[3]. The remaining term $(\times)$ gives a delta-distribution, and thus we have[4]

$$\frac{\delta X(t)}{\delta\xi(t)} = \sqrt{2D_0}\int_{n\tau}^{n\tau+\Delta t} e^{-\alpha(n\tau+\Delta t-t')}\delta(n\tau+\Delta t-t')\mathrm{d}t' = \sqrt{\frac{D_0}{2}}, \tag{4.15}$$

in agreement with Eq. (4.4), where we have shown that this results also holds for the general (nonlinear) case.

[3]See Secs. 1.3 and 3.4 for a brief discussion concerning the meaning of causality in this context.

[4]We use $\int^{t_0} f(t)\delta(t-t_0)\mathrm{d}t = f(t_0)/2$. To see where this comes from, one may regard the delta-distribution as the limit of a normalized Gaussian distribution with variance ϵ, for $\epsilon \to 0$.

For the variational derivative w. r. t. the delayed Langevin force, we analogously utilize (4.13) which yields

$$\frac{\delta X(t)}{\delta\xi(t-\tau)} = \frac{\delta\mathcal{Y}_{n+1}\big(n\tau+\Delta t\big)}{\delta\xi(n\tau-\tau+\Delta t)} = \underbrace{\frac{\delta\mathcal{Y}_n\big(n\tau\big)}{\delta\xi(n\tau-\tau+\Delta t)}}_{(*)}\, e^{-\alpha\Delta t}$$
$$+\int_{n\tau}^{n\tau+\Delta t} e^{-\alpha(n\tau+\Delta t-t')}\left[\sqrt{2D_0}\,\delta([n\tau-\tau+\Delta t]-t') - \beta\,\underbrace{\frac{\delta\mathcal{Y}_n(t'-\tau)}{\delta\xi\big(n\tau-\tau+\Delta t\big)}}_{(**)}\right]\mathrm{d}t'. \tag{4.16}$$

Again applying the causality argument, $(**)$ is found to have a nonzero value at solely one single point of the integral, i.e., when $t' = n\tau+\Delta t$ (upper bound). Since, furthermore, the integrand is finite at this unique point, in particular, $\delta X(t)/\delta\xi(t) = \sqrt{D_0/2}$ [Eq. (4.15)], the entire integral vanishes. In the remaining term $(*)$, we plug in, once again, the formal solution Eq. (4.13) (now for $t\in[n\tau-\tau, n\tau]$), and find

$$(*) = \underbrace{\frac{\delta\mathcal{Y}_{n-1}\big(n\tau-\tau\big)}{\delta\xi\big(n\tau-\tau+\Delta t\big)}}_{\to 0}\, e^{\alpha\tau}$$
$$+\int_{n\tau-\tau}^{n\tau} e^{-\alpha(n\tau-t')}\left[\underbrace{\frac{\sqrt{2D_0}\,\delta\xi(t')}{\delta\xi\big(n\tau-\tau+\Delta t\big)}}_{(\#)} - \beta\,\underbrace{\frac{\delta\mathcal{Y}_n(t'!-\tau)}{\delta\xi\big(n\tau-\tau+\Delta t\big)}}_{\to 0}\right]\mathrm{d}t',$$

where two terms again vanish due to their acausality. In contrast, the first term in the integral $(\#)= \delta(n\tau-\tau+\Delta t-t')$ gives a finite contribution. We thus obtain

$$\frac{\delta X(t)}{\delta\xi(t-\tau)} = \sqrt{2D_0}\,e^{-\alpha\tau}. \tag{4.17}$$

Hence, this functional derivative is [just like (4.15)] independent of whether the delay force is present, or not ($\beta=0$ or $\beta\neq 0$). This is somewhat surprising, because it means that the presence of delay does not influence how the trajectories change, when the (past) noise process is varied. We think, however, that this independence should be specific to the linear case.

From (4.7, 4.17), we find the FPE for the linear case

$$\frac{\partial}{\partial t}\rho_2(x,t;x_\tau,t-\tau) = -\frac{\partial}{\partial x}\left[(-\alpha x-\beta x_\tau)\,\rho_2(x,t;x_\tau,t-\tau)\right]$$
$$-\frac{\partial}{\partial x_\tau}\left[\int_\Omega(-\alpha x_\tau-\beta x_{2\tau})\,\rho_3(x,t;x_\tau,t-\tau;x_{2\tau},t-2\tau)\;\mathrm{d}x_{2\tau}\right]$$
$$+D_0\left[\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial x_\tau^2}\right]\rho_2(x,t;x_\tau,t-\tau)$$
$$+\sqrt{2D_0}\,e^{-\alpha\tau}\frac{\partial^2}{\partial x\,\partial x_\tau}\rho_2(x,t;x_\tau,t-\tau), \tag{4.18}$$

which we have also reported in [1]. The occurrence of the three-time PDF in (4.18) confirms that the linear system is described by an *infinite* hierarchy [as already discussed for the general case, see Eq. (4.7)]. It is interesting to note that, for the *linear* case, however, a corresponding closed set of FP equations can be found, as shown in [92] using a time-convolutionless transform.

The fact that the linear case can equivalently be described by a closed set of (two) equations, in a sense, suits to the fact that the (steady-state) solutions are Gaussian distributions, which are in turn fully specified by the first two moments. We will discus the Gaussian solutions in the next Section 4.2.2.

Markovian limit To understand the meaning and origin of the different terms in (4.18), it is instructive to consider Markovian limits. To this end, we take the limit $\beta \to 0$ of (4.18)[5]. We first note that τ, which represents a meaningful parameter defining the time lag in the delay force if $\beta > 0$, is, in this Markovian limit, only an (arbitrary) time difference. It specifies where we evaluate the two-time PDF, but has no further meaning. Second, we observe that the coupling to the higher members of the hierarchy vanishes, as expected. This is because the integral in the drift term can then be readily evaluated, yielding $\alpha x_\tau \int_\Omega \rho_3(x,t;x_\tau,t-\tau;x_{2\tau},t-2\tau)\ \mathrm{d}x_{2\tau} = \alpha x_\tau \rho_2(x,t;x_\tau,t-\tau)$. Finally, we realize that the diffusion terms stay unchanged in this limit. In particular, the mixed, $\partial_x \partial_{x_\tau}$ diffusion term, which quantifies some sort of correlation between the two instances in time (t and $t-\tau$), is *not* specific to the non-Markovian case. Thus, this term describes a type of "memory" which is also present in the absence of the delayed force (recall that the position-autocorrelation takes some time to relax to zero even in Markovian systems, representing a temporal correlation, see e.g. Fig. 1.4). For large time differences τ, this term completely vanishes as $e^{-\alpha\tau} \to 0$, and the variables x and x_τ decouple. This is in sharp contrast to the non-Markovian case, where x, x_τ, $x_{2\tau}$, ..., are all additionally coupled via the drift term.

4.2.2 Steady-state solutions

Considering the steady states (i.e., the limit $t \to \infty$), the FPE for ρ_2 (4.18) and ρ_1 (2.15) have (multivariate) Gaussian PDF solutions [77] of the form

$$\rho_{\mathrm{n,ss}} = \frac{1}{\sqrt{(2\pi)^n \det\{\mathbf{D_n}\}}} e^{-(1/2)(\boldsymbol{x}-\langle\boldsymbol{x}\rangle)\mathbf{D_n}^{-1}(\boldsymbol{x}-\langle\boldsymbol{x}\rangle)}, \tag{4.19}$$

with $\boldsymbol{x} = (x, x_\tau, .., x_{(n-1)\tau})^{\mathrm{T}}$, $\mathbf{D_1} = C(0)$, and

$$\mathbf{D_2} = \begin{pmatrix} C(0) & C(\tau) \\ C(\tau) & C(0) \end{pmatrix}, \quad \mathbf{D_3} = \begin{pmatrix} C(0) & C(\tau) & C(2\tau) \\ C(\tau) & C(0) & C(\tau) \\ C(2\tau) & C(\tau) & C(0) \end{pmatrix}. \tag{4.20}$$

The Gaussianity of the PDFs is a direct consequence of the linearity of the underlying LE with Gaussian noise, and can be shown on the basis of the LE alone [77].

[5] Please note that the limit $\tau \to 0$ as well yields a Markovian system, but this limit affects, at the same time, the arguments of the PDF.

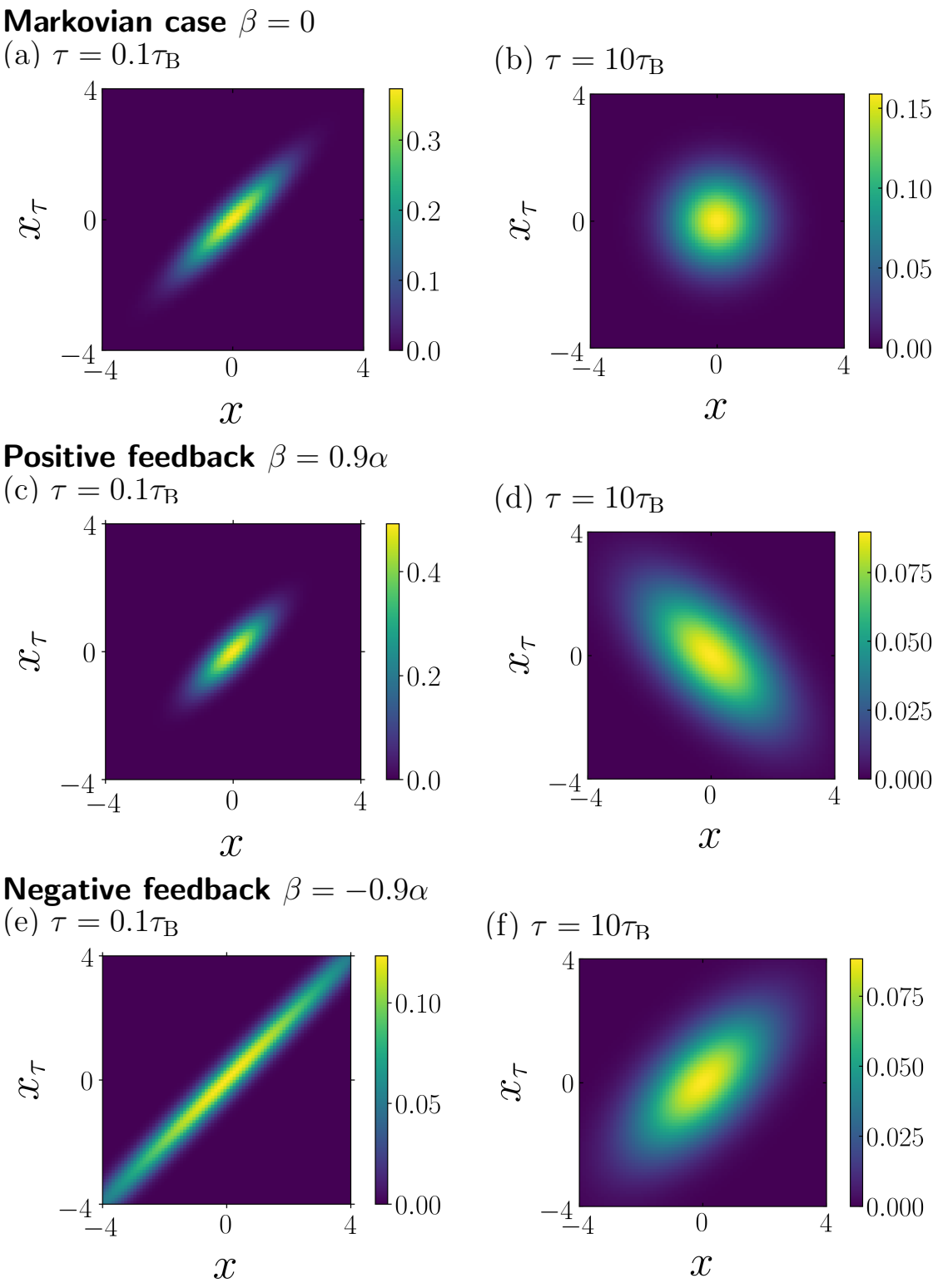

Figure 4.2: *Plots of the analytical results (4.24) for the steady-state two-time probability density function $\rho_2(x,t;x_\tau,t-\tau)$ of the linear LE (1.49) with $\alpha = \gamma/\tau_B$ and different values of β and τ. Left / right panels: Short / large delay times; Upper / central / lower panels: $\beta = 0$ / $\beta = 0.9\alpha$ / $\beta = -0.9\alpha$. The Markovian system always has a radial-symmetric PDF for large τ, the non-Markovian system has nontrivial correlations between x and x_τ for arbitrary τ. x and x_τ are in units of σ, the PDF (color code) is in units of $1/\sigma^2$.*

The covariance matrices $\mathbf{D_n}$ involve the steady-state temporal position-autocorrelation function between the systems state at time t and time $t+z$ (1.10), i.e., $C(z) := \langle X(t)X(t+z)\rangle_{\rm ss}$. Thus, the n-time PDF is fully known, once $C(z)$ with $z = 0, \tau, ...$ up to $z = n\tau$ are known.

In Sec. 1.4.5, we have given exact expressions for $C(0)$ and $C(\tau)$ as functions of α, β, τ and D_0. We recall

$$C(0) = D_0 \frac{1 + (\beta/\omega)\sinh(\omega\tau)}{\alpha + \beta\cosh(\omega\tau)},$$
$$C(z) = C(0)\cosh(\omega z) - (D_0/\omega)\sinh(\omega|z|),$$

with $\omega = \sqrt{\alpha^2 - \beta^2} \in \mathbb{C}$, as given in Eqs. (1.50, 1.51). This readily implies the exact one- and two-time steady-state PDFs.

To obtain the three-time steady-state PDF, we derive an explicit expression for $C(2\tau)$. We do this by generalization of the calculation from Ref. [85].

Extension to larger time differences Now we extend the result (1.50) towards higher z. Due to the symmetry property $C(z) = C(-z)$, we only need to consider non-negative z. One can directly deduce from the LE a differential equation for $C(z)$ for any $z > 0$ [85]

$$\frac{{\rm d}C(z)}{{\rm d}z} = \left\langle X(t)\frac{{\rm d}X(u)}{{\rm d}u}\Big|_{u=t+z}\right\rangle_{\rm ss} = -\alpha C(z) - \beta C(z-\tau) + \sqrt{2D_0}\langle X(t)\xi(t+z)\rangle. \tag{4.21}$$

For $z > 0$, the last term vanishes, because it is *acausal*, thus

$${\rm d}C(z)/{\rm d}z = -\alpha C(z) - \beta C(z-\tau). \tag{4.22}$$

This equation can be solved iteratively using the *method of steps*.

To be more specific, we plug in for the term $C(z-\tau)$ the solution on the interval $z \in [0, \tau]$, which is given in Eqs. (1.50, 1.51). This yields a closed equation for the correlation function on the interval $z \in [\tau, 2\tau]$, with the solution

$$\begin{aligned} C(z) = \frac{2D_0}{\beta[\alpha + \beta\cosh(\omega\tau)]}\Big(& e^{\alpha(\tau-|z|)}[\beta\cosh(\omega\tau) + \alpha] - \alpha\cosh[\omega(\tau - z)] \\ & - \frac{\beta}{2}\cosh[\omega(2\tau - z)] + \frac{\beta^2 - 2\alpha^2}{2\omega}\sinh\left[\omega(\tau - |z|)\right] \\ & - \frac{\alpha\beta}{2\omega}\sinh\left[\omega(2\tau - |z|)\right]\Big). \end{aligned} \tag{4.23}$$

We have reported this solution in [1]. Again, when ω becomes imaginary if $|\alpha| < \beta$, the hyperbolic functions convert to trigonometric ones, as in the case of $C(0 \leq z \leq \tau)$. Eq. (4.23) generalizes the calculation of Ref. [85] towards larger time differences z. In the same manner, one can extend the result further towards higher z. Figure 4.3 shows the analytically calculated autocorrelation function for different values of α, β and τ.

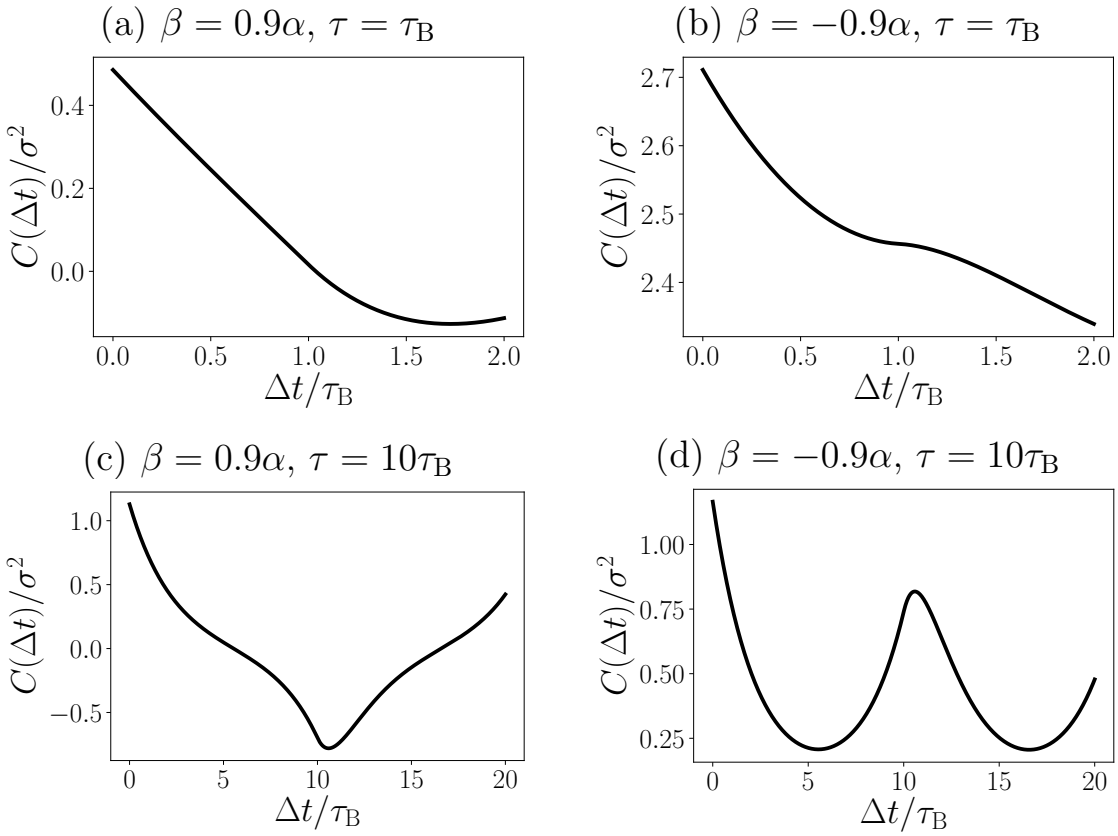

Figure 4.3: *Position-autocorrelation function on $t \in [0, 2\tau]$ from Eqs. (4.23, 1.51), of the linear system $\gamma\dot{X} = -\alpha X(t) - \beta X(t-\tau) + \sqrt{2D_0}\,\xi(t)$ with $\alpha = \gamma/\tau_\mathrm{B}$, and different values of β, τ. Upper panels, (a+b): $\tau = \tau_\mathrm{B}$, lower panels, (c+d): $\tau = 10\tau_\mathrm{B}$. Left panels, (a+c): $\beta/\alpha = 0.9$, i.e., positive feedback, Right panels, (b+d): $\beta/\alpha = -0.9$, i.e., negative feedback. These results also already displayed in Figs. 1.10 and 1.11.*

To avoid confusion, we note that in Ref. [85], the there derived expression (1.51) is erroneously stated to be valid for all $z > 0$. Along their derivation, the equation with number (18) in [85] turns out to be not valid for $|z| > \tau$, although they claim otherwise. A correlation described by Eq. (1.51) $\forall z \in \mathbb{R}$ would, in fact, be unphysical, as it yields unbounded correlations, even in situations where the system is stable and confined (e.g., for $\beta < \alpha$).

Green's function method An alternative strategy to obtain the correlation functions is via the Green's function method. We explain how this works and derive the needed formula in Appendix A.3. In principle, for various linear systems this strategy yields analytical expressions, which can, e.g., be numerically integrated. However, explicit *closed-form solutions* are only available for specific cases, where some required integrals are known (as well as the back-transformation of the Green's function to real space). See Refs. [89, 99] for some explicit results. For the LE (1.49) considered here, we could not find general closed-form expression for the position-autocorrelation function on this way.

Using (4.19) and (1.51), the one-time and two-time PDFs can be written in quite compact form as [85]

$$\rho_{1,\mathrm{ss}}(x) = \frac{1}{\sqrt{2\pi K(\tau)}} e^{-\frac{x^2}{2K(\tau)}}$$
$$\rho_{2,\mathrm{ss}}(x,t;x_\tau,t-\tau) = \frac{d_1(\tau)\sqrt{1-d_2(\tau)}}{\pi} e^{-d_1(\tau)[x^2+x_\tau^2-2x\,x_\tau d_2(\tau)]} \tag{4.24}$$

with variance $K(\tau) = C(0)$ given in Eq. (1.50), and

$$d_1(\tau) = \frac{1}{2}\frac{\beta^2 K(\tau)}{[\beta K(\tau)]^2 - [D_0 - \alpha K(\tau)]^2}, \qquad d_2(\tau) = \frac{D_0 - \alpha K(\tau)}{\beta K(\tau)}. \tag{4.25}$$

In the limit $\beta \to 0$, these results converge to the well-known Markovian PDFs for the Ornstein-Uhlenbeck process, given for example in Ref. [173], p. 101,

$$\rho^{(0)}_{1,\mathrm{ss}}(x) = \frac{1}{\sqrt{2\pi(D_0/\alpha)}} e^{-\frac{x^2}{2D_0/\alpha}}, \tag{4.26}$$

$$\rho^{(0)}_{2,\mathrm{ss}}(x,t;x_\tau,t-\tau) = \frac{1}{2\pi(D_0/\alpha)\sqrt{1-e^{-2\alpha\tau/\gamma}}} e^{-\frac{x^2+x_\tau^2-2x\,x_\tau e^{-2\alpha\tau/\gamma}}{2(D_0/\alpha)(1-e^{-2\alpha\tau/\gamma})}}. \tag{4.27}$$

Figure 4.2 shows the two-time PDF (4.24) of linear systems with delay force, as well as, the Markovian limit (4.26). For very short τ, all depicted PDFs appear as (uprising) diagonals, indicating high probabilities to find $X(t)$ in the vicinity of $X(t-\tau)$. For very large τ, only the Markovian process always approaches a radial-symmetric PDF [this can directly be seen from Eq. (4.26)], due to a vanishing correlation $C(\tau) \to 0$ for increasing τ. More specifically, the symmetric PDF is approached once τ is large compared to the intrawell relaxation time[6] $\tau_{\mathrm{ir}} \approx \gamma/\alpha$, as one can see from (4.26).

On the contrary, the non-Markovian system has a nontrivial correlation between x and x_τ for arbitrarily large τ, resulting in a two-time PDF with broken radial symmetry. The respective value of $C(\tau)$ can be extracted from Fig. 4.3, or Eq. (1.50).

4.2.3 The notion of effective temperature

Based on the one-time PDF alone, the non-Markovian system cannot be distinguished from a Markovian one at a different temperature. This can be seen by comparing (4.24) with (4.26). The delay force only changes the variance

$$D_0/\alpha \;\to\; C(0) = (D_0/\alpha)\,\frac{1+(\beta/\omega)\sinh(\omega\tau)}{1+(\beta/\alpha)\cosh(\omega\tau)}$$

[as we have given in Eq. (1.50)], just as well as a changed temperature

$$\frac{\gamma D_0}{k_{\mathrm{B}}} = \mathcal{T} \;\to\; \mathcal{T}_{\mathrm{eff}} = \frac{\gamma D_0}{k_{\mathrm{B}}}\,\frac{1+(\beta/\omega)\sinh(\omega\tau)}{1+(\beta/\alpha)\cosh(\omega\tau)}$$

[6] The timescales have been introduced in Chapter 1.6. We further recall that a list of all symbols can be found on page IX.

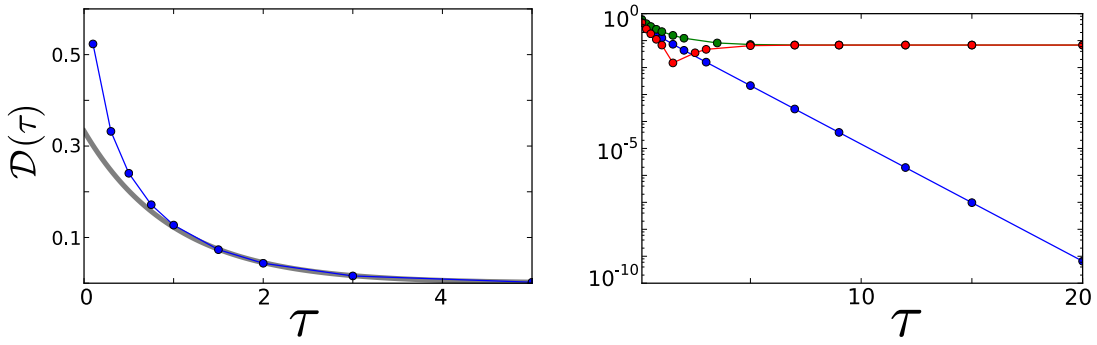

Figure 4.4: *Kolmogorov-like distance $\mathcal{D}$, as defined in (4.28), between two-time PDF $\rho_2(x, x_\tau)$ and the product of the one-time PDFs $\rho_1(x)\rho_1(x_\tau)$ at $\alpha = \gamma D_0/\sigma^2$. Left panel: Markovian case ($\beta = 0$, blue lines) plus an exponential fit: $(1/3)e^{-\alpha\tau/\gamma}$ as a guide to the eye (gray line). Right panel: Again the Markovian case ($\beta = 0$, blue lines), here in a logarithmic plot, supplemented by the results for non-Markovian cases ($\beta = -0.4\,\alpha$ green line, and $\beta = 0.4\,\alpha$ red line).*

would do. However, this viewpoint breaks down, if we consider the higher order PDFs, which are crucially different than the one in a Markovian system. To illustrate this point, we recall Fig. 4.2, which shows that the two-time PDF significantly differ.

Furthermore, when we consider *nonlinear* systems, the idea that the impact of time-delayed force can be captured by an effective temperature will even breaks down on the one-time PDF level. We will come back to this thought in Chapter 6.

4.2.4 Markovian vs. non-Markovian two-time probability density

We turn back to the aforementioned Markovian limit $\beta \to 0$ [discussed below Eq. (4.18)], and reconsider the meaning of the $e^{-\alpha\tau}$ decaying "memory" (induced by the mixed diffusion term) on the basis of the exact solutions. To this end, we study the Kolmogorov-like distance

$$\mathcal{D}(\tau) = \frac{1}{2}\iint \mathrm{d}x\mathrm{d}x_\tau \left|\rho_{2,\mathrm{ss}}(x, x_\tau; \tau) - \rho_{\mathrm{ss}}(x)\rho_{\mathrm{ss}}(x_\tau)\right| . \tag{4.28}$$

which was previously also used in [328] to quantify (non-)Markovianity. This article [328] further provides a systematic comparison between different non-Markovianity measures, and between quantum and classical non-Markovian dynamics. We find that it decays to zero in the Markovian case, as can be seen in Fig. 4.4. For $\tau \gg 0$, the decay is exponential, with the same exponent as the term that scales the mixed diffusion term in Eq. (4.18). Thus, the Markovian process has exponentially decaying correlations between the times t and $t - \tau$, such that the two-time PDF $\rho_2(x, t; x_\tau, t-\tau)$ decomposes to the product of the two identical one-time PDF for large time differences τ.

On the contrary, the distance is not decaying to zero if $\beta > 0$. The non-Markovian system is thus characterized by a correlation between t and $t - \tau$, for

arbitrary τ, as one would expect. As one can further see from Fig. 4.4, the distance instead approaches a constant finite value. This is because in fact, both, the one-time and two-time PDF *saturate*, when τ is larger than the other characteristic timescales of the process (which is here the intrawell relaxation time, discussed in Chapter 6), see Eq. (4.24). This saturation is also observed in nonlinear systems, and we will come back to it in Sec. 6.4.

Interestingly, the approached saturation value of $\mathcal{D}$ identical for $+\beta$ and $-\beta$. This observation also holds for other values of α and D_0 (not shown here).

Consistency check of results We have plugged in the PDFs ρ_2 and ρ_3 (which were derived on the basis of the LE alone) into the FPE (4.18), and found that they indeed solve the equation. This confirms the consistency of the different approaches and of our results.

At this point, one might actually question the usefulness of the FPE, which is technically more involved than the LE, but still seems not necessary to find solutions. However, as we will see in the following, the FPE is a valuable tool to obtain exact relations (especially for nonlinear cases). It further represents a starting point for approximation schemes, as discussed in Chapter 6. Also the probabilistic solutions of the linear delay system will prove themselves useful in this regard.

5 | Markovian Embedding – A new derivation of the Fokker-Planck hierarchy

"A clever person solves a problem. A wise person avoids it."

– supposedly by Albert Einstein [329, 330]

In this Chapter, we derive the FP hierarchy by using a Markovian embedding technique, which is inspired by the *linear chain trick* [9, 122, 123]. Markovian embeddings are already a well-established tool to treat stochastic systems with memory, for example in the context of generalized Langevin equations [97, 117–121]. The following considerations are based on Ref. [3], where we have introduced this derivation. Here, we add additional noise terms to the auxiliary variables (different from [3]), which might provide a conceptual advantage for future analytic treatments.

Main idea We introduce $n \in \mathbb{N}$ auxiliary variables $X_{j=1,2,\dots,n}$ whose dynamics is given by a set coupled, linear differential equations without delay. These degrees of freedom will generate the memory. They are coupled to another degree of freedom, X_0, which follows a (nonlinear) stochastic differential equation, which is as well

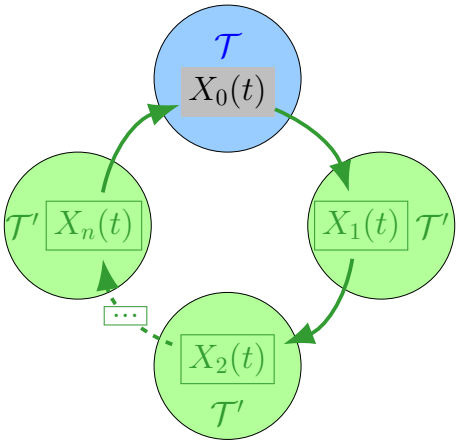

Figure 5.1: *Schematic visualization of the $(n+1)$-dimensional Markovian system (5.1). In the limit $n \to \infty$, the dynamics of X_0 is identical to the time-delayed process (1.42). The variables $X_{j>0}$ represent the memory and form the feedback loop.*

delay-free, but represents the original delayed process. In the limit $n \to \infty$, this set of equations generates the same stochastic process as the delayed equation, where $X_0(t) \to X(t)$. Thus, the delay in the original equation is replaced by a coupling with (infinitely many) additional degrees of freedom.

This $(n+1)$-dimensional Markovian system is described by a closed, *self-sufficient* FPE. By marginalization, i.e., projection onto the one-dimensional subspace X_0, and a subsequent limiting procedure, we derive the well-known delayed FPE for the one-time PDF. Furthermore, by projection onto higher-dimensional subspaces, we find higher FPEs, yielding an infinite hierarchy which has a different form than the one obtained by Novikov's theorem [327] (explicitly considering the second member, we show that the two hierarchies can actually be converted into each other). Interestingly, the limiting procedure contains one step where position-like variables are converted into velocity-like variables.

In a nutshell, the idea is to first employ a Markovian embedding on the level of the Langevin equation (5.1), and then, again integrate out the auxiliary variables from the embedding, on the level of the resulting Markovian FPE. Thus, the trick is to introduce and remove auxiliary variables at the right moments.

Meaning of additional noise terms and Generalization of Ref. [3] Contrary to Ref. [3], where we have introduced this idea, we here generalize the approach by adding noise terms $\sim \xi_j$ to the auxiliary variables $X_{j>0}$, scaled with a temperature $\mathcal{T}'$. The limit $\mathcal{T}' \to 0$ corresponds to the case considered in [3]. In particular, we show that the impact of the additional noise terms on X_0 *vanishes* in the limit $n \to \infty$. Therefore, starting from the $n+1$ LEs *with*, or *without*, the additional noise terms, yields the *same* FPE for ρ_1 (despite the additional source of randomness).

The benefits of this generalization include technical, as well as physical aspects. First, there is a technical advantage of adding noise terms to all d.o.f.. In this way, we "regularize" the noise. More specifically, by adding the noise terms to $X_{j>0}$ we avoid coupling between deterministic and stochastic degrees of freedom. This also implies that the diffusion matrix in the multivariate Fokker-Planck equation (5.27) has full rank. This solves, for the specific case at hand, the long-standing problem of singular diffusion matrices, see [331]. For example, this matrix can now be inverted. It might be interesting to check, whether one can also approach other cases like the ones from [331] with an analogous approach (i.e., adding noise terms to render a diffusion matrix with full rank, and then consider some limit where the impact of the additional noise term vanishes.) Typically this is a very useful property in the search of exact results (e.g., see [173], p. 84ff).

Second, we will in Part III reconsider the network (5.1) from a thermodynamic perspective. In the framework of stochastic thermodynamics, it is problematic to have variables without noise, because some key quantities are not meaningful in the noise-free "limit".

Third, there is a physical argument to add the noise terms $\sim \xi_j$, which takes into account possible underlying microscopic dynamics. Specifically, the additional noise terms allow for an alternative interpretation of the variables in (5.1). In particular, one can take the viewpoint that all equations in (5.1) are (overdamped) equations of motion, expressing a *force balance*, like a Langevin equation. From that perspective,

it is reasonable to balance each "friction force" $\sim -\gamma \dot{X}_j$ with a corresponding noise $\sim \xi_j$. In fact, friction and random force usually have the same origin, that is, a coupling with an (integrated out) heat bath (see Sec. 1.3 for an example). Thus, often one does not come without the other in a stochastic equation.

5.1 Markovian embedding – A different view on Memory

We consider the set of dynamical equations given by

$$\dot{X}_0(t) = \quad \gamma^{-1} F[X_0(t), X_n(t)] \quad + \sqrt{2k_{\mathrm{B}}\mathcal{T}_0/\gamma}\,\xi_0(t) \tag{5.1a}$$

$$\dot{X}_j(t) = \; (n/\tau)[X_{j-1}(t) - X_j(t)] \; + \sqrt{2k_{\mathrm{B}}\mathcal{T}'/\gamma'}\,\xi_j(t), \tag{5.1b}$$

with $j \in \{1, 2, ..., n\}$, $t \geq 0$, $X_j \in \mathbb{R}$ and independent Gaussian white noise terms $\langle \xi_i(t)\xi_j(t')\rangle = \delta_{ij}\delta(t-t')$. This is essentially a generalization of the so-called linear chain trick [9, 122, 123]. Importantly, the n auxiliary variables generally follow *linear* Eqs. (5.1b), while the delayed process itself may involve a nonlinear force, F. The Eqs. (5.1) define a unidirectionally coupled ring network, as depicted in Fig. 5.1. The initial conditions are set to $X_{j\in\{1,..,n\}}(0) \equiv 0$. As will become clear below, this amounts to the initial condition $X(t<0) \equiv 0$ in the non-Markovian LE (see Sec. 5.1.4).

As we explicitly show in the following, *projection* of (5.1b) onto the variable X_0 yields

$$\gamma \dot{X}_0(t) = F\left[X_0(t), \int_{-\tau}^{t} K_n(t-t')X_0(t')\,\mathrm{d}t' + \nu_n(t)\right] + \sqrt{2k_{\mathrm{B}}\mathcal{T}_0\gamma}\;\xi_0(t), \tag{5.2}$$

with the n-dependent Gamma-distributed memory kernel

$$K_n(t-t') = \left(\frac{n}{\tau}\right)^n \frac{(t-t')^{n-1}}{(n-1)!}\, e^{-n(t-t')/\tau}. \tag{5.3}$$

and the n-dependent colored noise ν_n with $\langle \nu_n \rangle = 0$, and

$$C_\nu(T) = \frac{k_{\mathrm{B}}\mathcal{T}'}{\gamma'} \sum_{p=0}^{n-1} \sum_{l=0}^{p} \frac{2^{l-2p}(2p-l)!}{p!(p-l)!l!} \left(\frac{n}{\tau}\right)^{l-1}\; e^{-nT/\tau} T^l. \tag{5.4}$$

5.1.1 Projection, Memory kernel & Colored noise

Here we show how to *hide* the $X_{j>0}$, meaning we carry out the projection onto X_0. In this way, we derive the memory kernel and colored noise from Eq. (5.2). The projecting method is technically similar to the Mori-Zwanzig projection (see Sec. 1.3.1). However, in contrast to the derivation of a generalized LE, we here do not start from microscopic, mechanically coupled particles, but instead from the mesoscopic auxiliary d.o.f..

Projection We transform the equations for $X_{j\in\{1,2,...,n\}}$ into frequency space, and make use of their linearity (we recall that their linearity is irrespective of the question whether the equation for X_0 is linear). First, we apply the Laplace-transformation[1] $\mathcal{L}[X_j(t)](s) = \int_0^\infty X_j(t)e^{-st}\mathrm{d}s$ and obtain for each $j \in \{1, 2, ..., n\}$

$$\begin{aligned} s\hat{X}_j(s) &= (n/\tau)\left[\hat{X}_{j-1}(s) - \hat{X}_j(s)\right] + \sqrt{2k_\mathrm{B}\mathcal{T}'/\gamma'}\,\hat{\xi}_j(s) \\ \Rightarrow \hat{X}_j(s) &= \frac{(n/\tau)\hat{X}_{j-1}(s) + \sqrt{2k_\mathrm{B}\mathcal{T}'/\gamma'}\;\hat{\xi}_j(s)}{s + (n/\tau)}. \end{aligned} \tag{5.5}$$

Here we have already incorporated the initial conditions $X_{j\in\{1,..,n\}}(0) \equiv 0$, which simplify the expression (5.5). One should further note that the Laplace-transformation of the noise is in fact not well-defined. However, we can safely ignore this issue here, as we will transform back at the end. Iteratively substituting the solutions (5.5) into each other yields

$$\hat{X}_n = \frac{(n/\tau)^n}{[s + (n/\tau)]^n}\hat{X}_0 + \sum_{j=1}^{n} \frac{(n/\tau)^{j-1}\,\sqrt{2k_\mathrm{B}\mathcal{T}'/\gamma'}\,\hat{\xi}_{n-j+1}}{[s + (n/\tau)]^j}. \tag{5.6}$$

Now we transform back to the real space via inverse Laplace-transformation. In (5.6) we identify the Laplace-transformation of the Gamma distribution

$$\mathcal{L}\left[K_j(t)\right](s) = \frac{(n/\tau)^j}{[s + (n/\tau)]^j} \tag{5.7}$$

with the Gamma-distributed kernels $K_j(t) = \frac{n^j}{\tau^j(j-1)!}\,t^{j-1}e^{-nt/\tau}$. Furthermore, making use of the convolution theorem as well as the linearity of the Laplace-transformation, we readily find

$$X_n(t) = \int_0^t K_n(t - t')X_0(t')\,\mathrm{d}t' + \nu_n(t) \tag{5.8}$$

with the Gaussian colored noise

$$\nu_n(t) = \sum_{j=1}^{n} \int_0^t \frac{\tau}{n} K_j(t - t')\,\sqrt{\frac{2k_\mathrm{B}\mathcal{T}'}{\gamma'}}\,\xi_{n-j+1}(t')\,\mathrm{d}t'. \tag{5.9}$$

Hence, we have found the expressions for the memory kernel and colored noise in the equation of motion of X_0 (5.2), generated by the variables $X_{j>0}$.

Here we have used a special choice of initial conditions for the variables $X_{j>0}$, which simplified the derivation. For general initial conditions, there are additional terms in the projected equation for X_n, which read $X_n(0)e^{-nt/\tau} + \frac{n}{\tau}X_{n-1}(0)t\,e^{-nt/\tau} + ...$, as we have explicitly shown in Ref. [3] (where we have presented the projection in real space). However, these terms decay exponentially with time. Thus, for the long time limit, which we mostly focus on in this thesis, these terms are anyway irrelevant.

[1] As opposed to Fourier-transformations, the Laplace-transformation is useful to treat initial-value problems, while Fourier-transformations have advantages in steady states, where the initial condition effects have already decayed. Both transformations could be used here.

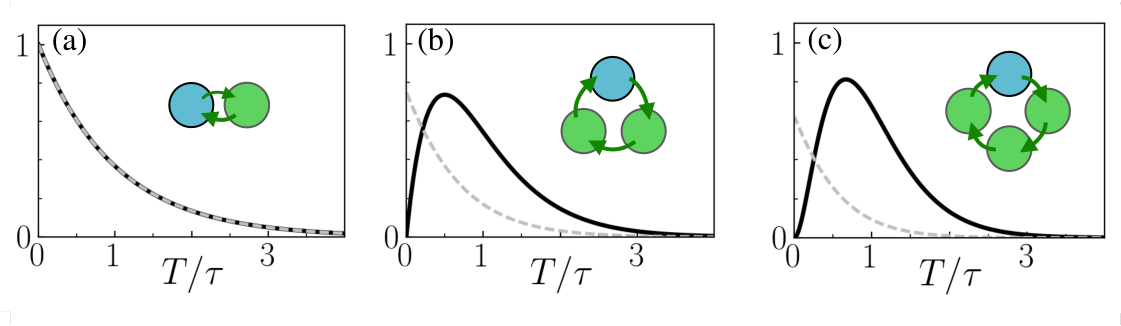

Figure 5.2: *The memory kernel (solid black lines) from Eq. (5.3) and the noise correlations (dashed gray lines) from Eq. (5.4), from $n = 1, 2, 3$ auxiliary variables shown in panel (a), (b), and (c), respectively. The insets show sketches of the respective networks (5.1), similar to Fig. 5.1. Here, $k_{\mathrm{B}}\mathcal{T}' = \sigma^2\,\tau\,\gamma'$.*

Memory kernel The memory kernel is given by a Gamma distribution (5.3). This is a consequence of the linear coupling. Figure 5.2 depicts the kernel and the noise correlations for different numbers of auxiliary variables, i.e, different values of n. For $n = 1$ [Fig. 5.2 (a)], the kernel is a simple exponential decay. For $n > 1$ [Fig. 5.2 (b,c)], the kernel has a maximum around its mean value $\mu = \int_0^\infty K_n(\Delta t)\Delta t\,\mathrm{d}(\Delta t) = \tau$, which is independent of the value of n. By increasing n, the kernel gets sharper [and eventually approaches a delta-distribution (5.20), see below].

Noise correlations The correlations $C_\nu(\Delta t) = \langle \nu_n(t)\nu_n(t+\Delta t)\rangle$ of (5.9) can be calculated exactly for an arbitrary $\Delta t > 0$, as follows. First, we perform the trivial integration over the δ-distributions from the white noise correlations $\langle \xi_{n-j+1}(t')\xi_{n-i+1}(t'')\rangle = 2\gamma' k_{\mathrm{B}}\mathcal{T}'\delta_{i,j}\delta(t'-t'')$, which readily yields

$$\begin{aligned}\frac{\gamma' n^2}{k_{\mathrm{B}}\mathcal{T}'\tau^2\ 2}\,C_\nu(\Delta t) &= \sum_{j=1}^{n}\int_0^t K_j(t-t')K_j(t-t'+\Delta t)\,\mathrm{d}t' \\ &= \sum_{j=1}^{n}\frac{(n/\tau)^{2j}}{(j-1)!^2}\int_0^t (t'^2 + t'\Delta t)^{j-1}e^{-n(\Delta t+2t')/\tau}\,\mathrm{d}t'. \end{aligned} \tag{5.10}$$

Performing the substitution $u = 2t'n/\tau$, we further obtain

$$\frac{\gamma' n^2}{k_{\mathrm{B}}\mathcal{T}'\tau^2\ 2}\,C_\nu(\Delta t) = e^{-n\Delta t/\tau}\sum_{j=1}^{n}\frac{(n/\tau)2^{1-2j}}{(j-1)!^2}\underbrace{\int_0^{2tn/\tau}\left(u^2 + u\frac{2n}{\tau}\Delta t\right)^{j-1}e^{-u}\,\mathrm{d}u}_{=(*)}. \tag{5.11}$$

The integral $(*)$ can be simplified by using the binomial theorem, which yields

$$(*) = \sum_{l=0}^{j-1} \binom{j-1}{l} \int_0^{2tn/\tau} u^{2(j-1-l)} \left(u\frac{2n\Delta t}{\tau}\right)^l e^{-u}\,\mathrm{d}u$$
$$\sum_{l=0}^{j-1} \frac{(2n\Delta t)^l\,(j-1)!}{\tau^l\,l!\,(j-l-1)!} \int_0^{2tn/\tau} u^{2j-l-2} e^{-u}\,\mathrm{d}u. \tag{5.12}$$

Since we focus on steady states and are mainly interested in the noise correlations after the initial condition effects have decayed[2], we now take the limit $t \to \infty$. Then, we perform the integration using the definition of the Gamma-function

$$\int_0^\infty x^p e^{-x}\mathrm{d}x = \Gamma(p+1) = p!, \tag{5.13}$$

which yields

$$(*) = \sum_{l=0}^{j-1} \left(\frac{2n\Delta t}{\tau}\right)^l \frac{(j-1)!(2j-l-2)!}{l!(j-l-1)!}. \tag{5.14}$$

Combining (5.11, 5.14) yields the noise correlation given in (5.4).

Dynamics of X_j With the very same arguments as used above to derive (5.2), one can show that each $X_{j>0}$ follows a similar equation. Specifically, one finds

$$X_j(t) = \int_0^t K_j(t-t')X_0(t')\,\mathrm{d}t' + \nu_j(t) \tag{5.15}$$

with $K_j(t-t') = \frac{n^j\,(t-t')^{j-1}}{\tau^j(j-1)!}e^{-n(t-t')/\tau}$, and the Gaussian colored noise

$$\nu_j(t) = \sum_{l=1}^{j} \int_0^t \frac{\tau}{n} K_l(t-t') \sqrt{\frac{2k_\mathrm{B}\mathcal{T}'}{\gamma'}}\,\xi_{j-l+1}(t')\,\mathrm{d}t'. \tag{5.16}$$

5.1.2 Limit $n \to \infty$

Here we study the limit of the dynamical equation for the variable X_0, Eq. (5.2).

Limit of memory kernel In the limit $n \to \infty$, the memory kernel (5.2) collapses onto a delta peak at τ (5.20). To show this, we consider several properties of the kernel. First, it is clearly a non-negative function $K \geq 0$. Second, its mean value

$$\mu = \int_0^\infty K_n(\Delta t)\Delta t\,\mathrm{d}(\Delta t) = \tau, \tag{5.17}$$

[2]One could treat transient dynamics similarly by using the incomplete Gamma function instead.

independent of the value of n. The same holds for its weight, i.e., the integral over the kernel $\int_0^\infty K_n(\Delta t)\,\mathrm{d}(\Delta t) = 1$, which is unity. In contrast, its variance

$$\int_0^\infty K_n(\Delta t)^2\,\mathrm{d}(\Delta t) - \mu^2 = \frac{\tau^2}{n}, \tag{5.18}$$

vanishes for $n \to \infty$. Finally, for large n, its value at $\Delta t = \tau$ can be estimated by using the Stirling formula $n! \approx \sqrt{2\pi n}(n/e)^n$, yielding

$$K_n(\tau) = \frac{n^n\, e^{-n}}{(n-1)!\tau} \approx \frac{n!}{(n-1)!\sqrt{2\pi n}\tau} = \sqrt{\frac{n}{2\pi}}\frac{1}{\tau} \to \infty. \tag{5.19}$$

In total, these considerations give rise to the limit

$$\lim_{n\to\infty} K_n(\Delta t) = \delta(\Delta t - \tau). \tag{5.20}$$

Colored noise Now we turn to the limit of the noise correlations. To this end, we consider the weight of the correlation C_ν [from Eq. (5.4)], which is given by

$$\begin{aligned}\int_0^\infty C_\nu(\Delta t)\mathrm{d}(\Delta t) =& \frac{k_\mathrm{B}\mathcal{T}'}{\gamma'}\frac{\tau}{n}\sum_{p=0}^{n-1}\sum_{l=0}^{p}\frac{n^l\tau^1}{2^{2p-l}p!}\frac{(2p-l)!}{(p-l)!\,l!}\int_0^\infty e^{-n\Delta t/\tau}\Delta t^l\,\mathrm{d}(\Delta t)\\ =& \frac{k_\mathrm{B}\mathcal{T}'}{\gamma'}\frac{\tau^2}{n^2}\sum_{p=0}^{n-1}\sum_{l=0}^{p}\frac{(2p-l)!}{(p!)(p-l)!}\left[\frac{1}{2}\right]^{2p-l}.\end{aligned} \tag{5.21}$$

Using the binomial theorem, we can further simplify this expression, and find

$$\begin{aligned}\frac{\gamma'}{k_\mathrm{B}\mathcal{T}'}\int_0^\infty C_\nu(\Delta t)\mathrm{d}(\Delta t) &= \frac{\tau^2}{n^2}\sum_{m=1}^{n}\sum_{r=0}^{p}\binom{p}{r}\left[\frac{1}{2}\right]^r\left[\frac{1}{2}\right]^{p-r}\\ &= \frac{\tau^2}{n^2}\sum_{m=1}^{n}\left[\frac{1}{2}+\frac{1}{2}\right]^{2(n-m)} = \frac{\tau^2}{n^2}\sum_{m=1}^{n}1 = \frac{\tau^2}{n}.\end{aligned} \tag{5.22}$$

Hence, the weight *vanishes* as $n \to \infty$. Since C_ν is obviously a non-negative function of Δt, this readily implies

$$\lim_{n\to\infty} C_\nu(\Delta t) = 0. \tag{5.23}$$

Limit of X_0 With (5.20) and (5.23), the limit $n \to \infty$ of the projected equation (5.2) is indeed *identical* to the original delayed LE (1.42), with

$$X_0(t) \to X(t), \quad X_n(\tau) \to X(t-\tau). \tag{5.24}$$

Limit of dynamical equations for X_j So far, we have shown that the colored noise in the dynamical equation for X_0 vanishes in the limit $n \to \infty$, see Eq. (5.23), while the memory kernel approaches a delta peak, Eq. (5.20). By application of the

same arguments, one can analogously show that in the projected equation (5.15) for $X_{j>0}$, K_j collapses onto a delta distribution, and ν_j fades away.

Thus, via the Eq. (5.15) all variables $X_{j>0}(t)$ are determined by the values of $X_0(t')$, $t' \in [t-\tau, t]$ alone, as expected. In the unidirectionally coupled ring network, the information is passed from X_0 to all other variables (see Fig. 5.1). The impact of the noise terms $\xi_{j>0}$ on the variables $X_{j>0}$ vanishes in the limit $n \to \infty$, irrespective of the value of $\mathcal{T}'$.

In Eq. (5.24), we have linked X_0 and X_n to the variable X of the delayed LE (1.42). The same can thus be done for all other $X_{j>0}$. We show this in the following and give another perspective on the meaning of the variables $X_{j>0}$ in this limit.

5.1.3 Interpretation of the X_j variables

We reconsider the dynamical Eqs. (5.1b) for the variables $X_{j>0}$. Using the *difference quotient* to approximate $\dot{X}_j$ yields

$$\frac{n}{\tau}\left[X_{j-1}(t) - X_j(t)\right] \overset{(5.1)}{\approx} \dot{X}_j(t) \approx \frac{X_j(t+\Delta t) - X_j(t)}{\Delta t}. \tag{5.25}$$

As we are interested in the interpretation at $n \to \infty$, it is safe to neglect the noise terms ξ_j here, whose impact on the X_j fades away in this limit, see Eq. (5.15) and below. Equation (5.25) gives rise to the interpretations $\Delta t = \tau/n$ and $X_j(t+\Delta t) \approx X_{j-1}(t)$, and, through iteration, $X_j(t + j\tau/n) \approx X_{j-i}(t)$. In combination with Eq. (5.24), this reveals

$$X_j(t) \to X(t - j\tau/n), \quad \forall j \in \{0, 1, 2, .., n\}. \tag{5.26}$$

The interpretation (5.26) becomes accurate in the limit $n \to \infty$ (at $\Delta t \to 0$ the difference quotient equals the differential quotient). Figure 4.1 (left panel) shows a visualization of the dynamics, which emphasizes the interpretation (5.26), and becomes accurate for $n \to \infty$. Figure 5.1 shows an alternative visualization focusing on the network structure of the coupled Langevin Eqs. (5.1).

5.1.4 Initial condition

We recall the initial conditions $X_{j\in\{1,..,n\}}(0) \equiv 0$. From the previous consideration, specifically Eq. (5.26), we see that this amounts to a constant zero history function $X(t < 0) \equiv 0$ of the delay LE. Thus, it is safe to let the integral with the memory kernel start at time $t = 0$, instead of $t \to -\infty$. As we are mainly interested in steady states throughout this thesis, we can choose this initial condition without loss of generality. However, for general cases, the transient dynamics is only described correctly by the Markovian set of equations, if the history-function of $X(t)$ is correctly translated to the initial values $X_{j\in\{1,..,n\}}$. Equation (5.15) provides the formula for this translation [for the case $\xi_j(t < 0) \equiv 0$].

5.1.5 $(n+1)$-dimensional Markovian Fokker-Planck equation

On a probabilistic level, the Markovian $(n+1)$-dimensional system given by the set of coupled LEs (5.1) can be described by the FPE

$$\underbrace{\partial_t \rho_{n+1}}_{\text{(I)}} = \underbrace{-\partial_{x_0}\left[\gamma^{-1}F(x_0,x_n)\rho_{n+1}\right]}_{\text{(II)}} \underbrace{+\frac{k_\mathrm{B}\mathcal{T}}{\gamma}\,\partial_{x_0}^2\rho_{n+1}}_{\text{(III)}} \underbrace{-\frac{n}{\tau}\sum_{j=1}^{n}\partial_{x_j}\left[(x_{j-1}-x_j)\rho_{n+1}\right]}_{\text{(IV)}} \underbrace{+\frac{k_\mathrm{B}\mathcal{T}'}{\gamma}\sum_{j=1}^{n}\partial_{x_j}^2\rho_{n+1}}_{\text{(V)}} \tag{5.27}$$

for the $(n+1)$-point (joint) PDF $\rho_{n+1}(x_0,t;x_1,t;...;x_n,t) = \langle\delta[x_0 - X_0(t)]\delta[x_1 - X_1(t)]..\delta[x_n - X_n(t)]\rangle$. Equation (5.27) is a self-sufficient FPE on a $(n+1)$-dimensional space, which describes on one subspace the dynamics of the delayed process in a closed form. This description is accurate in the limit $n \to \infty$. Importantly, we will nevertheless always regard n as a finite value when dealing with Eq. (5.27) itself, and only perform the limit at appropriate times (i.e., after marginalization as we explain below).

From a mathematical point of view, the limit $n \to \infty$ of Eq. (5.27) must be treated with special care. Indeed, it not is *per se* clear how to formulate a Fokker-Planck equation on an infinite-dimensional space [123]. Moreover, one drift term in (5.27) is proportional to n itself, making this limit highly non-trivial. Finding a mathematical framework in which the limit $n \to \infty$ of Eq. (5.27) is meaningful and well-defined is subject of ongoing work. Here, we can safely ignore this problem, since we focus on the equations for the marginalized (low-dimensional) PDFs. As shown in the following (Secs. 5.2 and 5.3), they can be derived by keeping in Eq. (5.27) n finite and performing the limit after the marginalization.

Further, we aim to note that Eq. (5.27) again resembles the FPE of a many-particle system, similar to (4.7). However, in sharp contrast to a system of interacting (colloidal) particles, the "interactions" are here *unidirectional* (see Fig. 5.1), which means that the "particle" x_j only feels the "particle" x_{j-1} and *not* vice-versa. Hence, the forces between them are *non-reciprocal* (violating Newtons third law) and cannot be described by an interaction potential. This point will be discussed in more detail in Part III of this thesis.

5.2 Derivation of first member of Fokker-Planck hierarchy

Here, we will use the Markovian embedding to find a probabilistic description of the delayed system, by a "coarse-graining" of (5.27) in the form of a *marginalization*. To be more specific, we integrate out all variables, but the ones relevant for the targeted member of the FP hierarchy. In this way, we obtain an FPE for the marginal PDFs.

We make use of the general relation between joint, $\rho_m(x,y)$, and marginal distributions: $\rho_{m-1}(z) = \int_\Omega \rho_m(y,z)\mathrm{d}y$. Importantly, we take the limit $n \to \infty$ only *after* the marginalization. We further employ again the natural boundary conditions (see Sec. 2.1.1). For sake of a shorter notation, we will omit the time arguments of the PDFs in lengthy expressions. We begin with the first member of the hierarchy, i.e., the delayed FPE for ρ_1.

Marginalization We start with the FPE (5.27) for the joint PDF, and integrate it over the whole domain of *all* n auxiliary variables, i.e., $\iint_\Omega \mathrm{d}x_1\mathrm{d}x_2..\mathrm{d}x_n$. The terms (I – III) in Eq. (5.27) can be readily marginalized to

$$\begin{aligned} \text{(I)} \quad &\to \partial_t \rho_1(x_0,t), & \text{(II)} \quad &\to -\partial_{x_0} \int_\Omega \left[F(x_0,x_n)\rho_2(x_0,t;x_n,t)\right] \mathrm{d}x_n, \\ \text{(III)} \quad &\to \partial_{x_0}^2 \rho_1(x_0,t), & \text{(V)} \quad &\to 0. \end{aligned} \tag{5.28}$$

To show the disappearance of (V), we have employed the natural boundary conditions, which imply $[\partial_{x_j}\rho_2(x_0;x_j)]_{-\infty}^{\infty} \to 0$. The remaining drift terms (IV) in Eq. (5.27) can be simplified by making use of their linearity w.r.t. the various x_j, and treating the remaining integrals by partial integration, which yields

$$\begin{aligned} \text{(IV)} \to \frac{n}{\tau}\Bigg\{ & n\,\rho_1(x_0) - \sum_{j=2}^{n} \iint_\Omega x_j\, \partial_{x_j}\rho_3(x_0;x_{j-1};x_j)\mathrm{d}x_j\mathrm{d}x_{j-1} + x_0 \underbrace{[\rho_2(x_0;x_1)]_{\omega_1}^{\omega_2}}_{=0} \\ & - \int_\Omega x_1\, \partial_{x_1}\rho_2(x_0;x_1)\mathrm{d}x_1 + \sum_{j=2}^{n} \int_\Omega x_{j-1} \underbrace{[\rho_3(x_0;x_{j-1};x_j)]_{\omega_1}^{\omega_2}}_{=0} \mathrm{d}x_{j-1} \Bigg\} \\ = \frac{n}{\tau}\Bigg\{ & n\,\rho_1(x_0) + \sum_{j=2}^{n} \underbrace{[x_j\rho_3(x_0;x_{j-1};x_j)]_{\omega_1}^{\omega_2}}_{=0} - \sum_{j=2}^{n} \int_\Omega \rho_2(x_0;x_j)\mathrm{d}x_j \\ & - \underbrace{[x_1\rho_2(x_0;x_1)]_{\omega_1}^{\omega_2}}_{=0} - \rho_1(x_0) \Bigg\} = 0. \end{aligned} \tag{5.29}$$

Again, we have further employed the boundary conditions multiple times. Hence, the contribution of (IV) to the drift term vanishes, and we obtain in total from (5.28, 5.29)

$$\partial_t \rho_1(x_0,t) = -\partial_{x_0} \int_\Omega \frac{1}{\gamma}\, F(x_0,x_n)\rho_2(x_0;x_n)\mathrm{d}x_n + \frac{k_\mathrm{B}\mathcal{T}}{\gamma}\, \partial_{x_0}^2 \rho_1(x_0,t). \tag{5.30}$$

Limiting procedure The marginalized FPE (5.30) still pertains to a finite (but arbitrary) value of $n \in \mathbb{N}$. As a last step, we finally take the limit $n \to \infty$. To this end, we recall the meaning of the respective stochastic variables in this limit [Eq. (5.24)]. Based on that, we perform the phase space transformation $\{x_0,t\} \to \{x,t\}$ and $\{x_n,t\} \to \{x_\tau, t-\tau\}$, and therewith, obtain[3] the delayed FPE for the one-time PDF (2.15).

[3] independently of the choice of $\mathcal{T}'$

Next, we demonstrate that our procedure can also be applied to find higher members of the FP hierarchy.

5.3 Derivation of higher members via Markovian embedding

To derive a FPE for the two-time PDF $\rho_2(x,t;x_\tau,t-\tau)$, we marginalize the full $(n+1)$-dimensional FPE (5.27) w.r.t. all variables but x_0 and x_n (again keeping n finite until after the marginalization). This time, we do not integrate over x_n as this variable represents $X(t-\tau)$ in the limit $n\to\infty$. In a second step, we will again perform a coordinate transformation to take the limit.

Marginalization The integration of the full FPE (5.27) over the variables $x_1,..,x_{n-1}$, is completely analogous to the marginalization presented in Sec. 5.2. In particular, one again performs the same steps, i.e., partial integrations and subsequent application of the boundary conditions. As opposed to the first member, now the term (IV) gives a non-vanishing contribution, namely

$$\text{(IV)} \to \partial_{x_n}\left[\int_\Omega \frac{n}{\tau}(x_n - x_{n-1})\rho_3(x_0,x_{n-1},x_n)\mathrm{d}x_{n-1}\right], \tag{5.31}$$

which involves a *three*-point PDF. Moreover, now the diffusion term (V) yields a non-zero term, that is

$$\text{(V)} \to \frac{k_\mathrm{B}\mathcal{T}'}{\gamma'}\partial_{x_n}^2\rho_2(x_0,x_n). \tag{5.32}$$

Thus, for the first time, we see a trace of the additional noise terms $\xi_{j>0}$ in the FPE, which yield an additional diffusion term on the level of the two-point PDF in the variable x_n.

In this way, one readily obtains the marginalized FPE

$$\begin{aligned}\partial_t\rho_2(x_0,t;x_n,t) =& -\partial_{x_0}\left[\gamma^{-1}F(x_0,x_n)\rho_2(x_0,t;x_n,t)\right] + \frac{k_\mathrm{B}\mathcal{T}}{\gamma}\partial_{x_0}^2\rho_2(x_0,t;x_n,t)\\ &+\partial_{x_n}\int_\Omega\frac{n}{\tau}(x_n-x_{n-1})\rho_3(x_0,t;x_{n-1},t;x_n,t)\mathrm{d}x_{n-1}\\ &+\frac{k_\mathrm{B}\mathcal{T}'}{\gamma'}\partial_{x_n}^2\rho_2(x_0,x_n).\end{aligned} \tag{5.33}$$

Confirming the consistency of this calculation, we note that integrating (5.33) over x_n yields the first member of the hierarchy (5.30), thus, the exact FPE for ρ_1 of the system with time delay. (By the way, this integration makes the x_n-diffusion term disappear again, independent of $\mathcal{T}'$.) Higher-order equations can be obtained in an analogous manner, yielding the infinite hierarchy of coupled equations, whose

$(m+1)$st member for $\rho_m(x_0; x_{n-m+2}, .., x_n)$, with $m \geq 3$, reads

$$\begin{aligned}
\partial_t \rho_m = &- \partial_{x_0} \left[\gamma^{-1} F(x_0, x_n) \rho_m(x_0; x_{n-m+2}, .., x_n)\right] \\
&+ \sum_{l=0}^{m-3} \frac{n}{\tau} \partial_{x_{n-l}} \left[(x_{n-l} - x_{n-l-1}) \rho_m(x_0; x_{n-m+2}, .., x_n)\right] \\
&+ \frac{n}{\tau} \partial_{x_{n-m+2}} \int_\Omega (x_{n-m+2} - x_{n-m+1}) \rho_{m+1}(x_0; x_{n-m+1}, .., x_n) \mathrm{d}x_{n-m+1} \\
&+ \frac{k_\mathrm{B} \mathcal{T}}{\gamma} \partial_{x_0}^2 \rho_m(x_0, t; x_{n-m+2}, t, .., x_n, t) \\
&+ \frac{k_\mathrm{B} \mathcal{T}'}{\gamma'} \sum_{l=0}^{m-3} \partial_{x_{n-1}}^2 \rho_m(x_0, t; x_{n-m+2}, t, .., x_n, t). \qquad (5.34)
\end{aligned}$$

In contrast to the FPE (5.30) for ρ_1, all higher members of this FP hierarchy (5.33, 5.34) explicitly contain the linear forces stemming from the auxiliary variable Eqs. (5.1b), and further involve diffusion terms scaled with the temperature $\mathcal{T}'$.

We now inspect in more detail the second member (5.33) in the limit $n \to \infty$, to recover the system with discrete delay. We again need to transform the probability space, hence, the phase space and the PDFs defined on it.

Transformation of phase space, limiting procedure In order to take the limit $n \to \infty$ of Eq. (5.33), we will transform the phase space and the PDFs. We perform the transformation in two steps. First we replace the phase space variable x_{n-1} by $\widetilde{x} = (n/\tau)(x_{n-1} - x_n)$, which will later in the limit $n \to \infty$ receive the meaning of a *velocity*. In a second step, we will transform back to the original variables of the delayed LE (1.42). This transformation will mainly affect the time arguments.

As first step, we change the phase space from $\{x_0, t; x_{n-1}, t; x_n, t\}$ to $\{x_0, t; \widetilde{x}, t; x_n, t\}$. To this end, we start with performing the substitution $x_{n-1} \longrightarrow \widetilde{x} = (n/\tau)(x_{n-1} - x_n)$ to rewrite the integral $\int_\Omega \mathrm{d}x_{n-1}$ as $(\tau/n) \int_{\widetilde{\Omega}} \mathrm{d}\widetilde{x}$. The drift term then becomes

$$\begin{aligned}
&\frac{n}{\tau} \partial_{x_n} \int_\Omega (x_n - x_{n-1}) \rho_3(x_0, t; x_{n-1}, t; x_n, t) \mathrm{d}x_{n-1} \\
&\quad = \frac{\tau}{n} \partial_{x_n} \int_{\widetilde{\Omega}} \widetilde{x}\, \rho_3 \left(x_0, t; \frac{\tau}{n}\widetilde{x} + x_n, t; x_n, t\right) \mathrm{d}\widetilde{x}. \qquad (5.35)
\end{aligned}$$

In the next step, we need to transform the PDF $\rho_3 \to \widetilde{\rho}_3$, where $\widetilde{\rho}_3$ is a normalized (joint) PDF of the new phase space variables $\{x_0, \widetilde{x}, x_n\}$. The relation between both PDFs can be found by comparing the normalizations. In particular,

$$\begin{aligned}
1 &= \iiint_\Omega \rho_3 \left(x_0, t; x_{n-1}, t; x_n, t\right) \mathrm{d}x_0 \mathrm{d}x_{n-1} \mathrm{d}x_n \\
&= \frac{\tau}{n} \int_\Omega \int_{\widetilde{\Omega}} \int_\Omega \rho_3 \left(x_0, t; \widetilde{x} + x_n, t; x_n, t\right) \mathrm{d}x_0 \mathrm{d}\widetilde{x} \mathrm{d}x_n, \qquad (5.36)
\end{aligned}$$

and, on the other hand, by definition, $\iiint \widetilde{\rho_3}\,(x_0,t;\widetilde{x},t;x_n,t)\,\mathrm{d}x_0\mathrm{d}\widetilde{x}\mathrm{d}x_n = 1$. Combining the constraints due to both normalizations, yields

$$\rho_3\left(x_0,t;\frac{\tau}{n}\widetilde{x}+x_n,t;x_n,t\right) \overset{!}{=} \frac{\tau}{n}\widetilde{\rho_3}\,(x_0,t;\widetilde{x},t;x_n,t)\,. \tag{5.37}$$

Therewith, we obtain the drift term

$$\begin{aligned}\frac{n}{\tau}\partial_{x_n}\int_\Omega (x_n-x_{n-1})\rho_3(x_0;x_{n-1};x_n)\mathrm{d}x_{n-1} &\longrightarrow \partial_{x_n}\int_{\widetilde{\Omega}}\widetilde{x}\,\widetilde{\rho_3}\,(x_0;\widetilde{x};x_n)\,\mathrm{d}\widetilde{x}\\ &= \partial_{x_n}\left[\left\langle \widetilde{X}(t)\middle| x_0=X_0(t), x_n=X_n(t)\right\rangle \rho_2(x_0,t;x_n,t)\right].\end{aligned} \tag{5.38}$$

In the last line, we have used the conditional average notation, which involves the respective stochastic variables. Now, we recall the meaning of the stochastic variables in the limit $n\to\infty$, i.e., $X_0(t)=X(t)$, $X_n(t)=X(t-\tau)$ from (5.24), and $\dot{X}_n(t)=\frac{n}{\tau}\left[X_{n-1}(t)-X_n(t)\right]+\sqrt{2k_\mathrm{B}\mathcal{T}'/\gamma'}\xi_j$ from (5.1b). Based on this, we perform the transformation

$$\{x_0,t\}\to\{x,t\}, \tag{5.39}$$
$$\{x_n,t\}\to\{x_\tau,t-\tau\}, \tag{5.40}$$
$$\{\widetilde{x},t\}\to\{\dot{x}_\tau,t-\tau\}. \tag{5.41}$$

A nonzero $\mathcal{T}'$ does not affect this transformation, as we will discuss. Further, we transform

$$\begin{aligned}&\partial_{x_n}\int_{\widetilde{\Omega}}\widetilde{x}\,\widetilde{\rho_3}\,(x_0,t;\widetilde{x},t;x_n,t)\,\mathrm{d}\widetilde{x}\\ &\qquad\longrightarrow \partial_{x_\tau}\left[\left\langle \dot{X}(t-\tau)\middle| x=X(t), x_\tau=X(t-\tau)\right\rangle \rho_2(x,t;x_\tau,t-\tau)\right],\end{aligned} \tag{5.42}$$

with $\dot{x}_\tau\in\widetilde{\Omega}=\mathbb{R}$. We note that the appearance of a *velocity*-like variable is unusual, as we deal with overdamped dynamics. Since $\dot{X}$ has no well-defined corresponding PDF for overdamped dynamics [158], we cannot express the conditional average on the right side of (5.42) in the form an integral.

Due to the coordinate transformation and limiting procedure, we finally obtain from (5.33), which is the FPE for the two-point PDF of the Markovian system (5.1), the FPE for the two-*time* PDF of the original non-Markovian system with time delay

$$\begin{aligned}\partial_t\rho_2(x,t;x_\tau,t-\tau) =& -\partial_x\left[\gamma^{-1}F(x,x_\tau)\rho_2(x,t;x_\tau,t-\tau)\right]+\frac{k_\mathrm{B}\mathcal{T}}{\gamma}\partial_x^2\rho_2(x,t;x_\tau,t-\tau)\\ &+\partial_{x_\tau}\left[\langle\dot{X}(t-\tau)|x=X(t),x_\tau=X(t-\tau)\rangle\rho_2(x,t;x_\tau,t-\tau)\right]\\ &+\frac{k_\mathrm{B}\mathcal{T}'}{\gamma'}\partial_{x_\tau}^2\rho_2(x,t;x_\tau,t-\tau).\end{aligned} \tag{5.43}$$

Before we compare Eq. (5.43) to the respective FPE obtained from Novikov's theorem [Eq. (4.7)] we will discuss the impact of the additional noise terms.

Effect of additional noise terms at $\mathcal{T}'$ Let us consider the effect of additional noise terms. It is easy to see that the transformation $\{\widetilde{x}, t\} \to \{\dot{x}_\tau, t-\tau\}$ from (5.41) is exact in the limit $\mathcal{T}' \to 0$. However, because $n \to \infty$, this transformation is as well expected to apply at finite $\mathcal{T}' > 0$ (recall that the additional noise term in the LE vanishes as $n \to \infty$ despite $\mathcal{T}' > 0$). Thus, from that perspective, $\mathcal{T}'$ may be finite. On the contrary, the resulting FPE (5.43) contains the additional diffusion term $\frac{k_\mathrm{B}\mathcal{T}'}{\gamma}\partial^2_{x_\tau}\rho_2(x,t;x_\tau,t-\tau)$, which describes an effect of the additional noise terms on the probabilistic level. This renders a different two-time PDF as at $\mathcal{T}' \to 0$. This is interesting, because the first member of the FP hierarchy was independent of $\mathcal{T}'$. In the following, we will consider the case $\mathcal{T}' = 0$.

5.3.1 Comparison to equation from Novikov's theorem

Comparing the here derived equation (5.43) (at $\mathcal{T}' = 0$)

$$\begin{aligned}\partial_t\rho_2(x,t;x_\tau,t-\tau) = \partial_{x_\tau}\left[\langle\dot{X}(t-\tau)|x=X(t),x_\tau=X(t-\tau)\rangle\rho_2(x,t;x_\tau,t-\tau)\right] \\ - \partial_x\left[\gamma^{-1}F(x,x_\tau)\rho_2(x,t;x_\tau,t-\tau)\right] \\ + \frac{k_\mathrm{B}\mathcal{T}}{\gamma}\partial_x^2\rho_2(x,t;x_\tau,t-\tau),\end{aligned} \tag{5.44}$$

to the FPE (4.7) obtained from Novikov's theorem, reveals that they are in fact quite different. Our approach thus yields an *alternative representation* of the second member. However, as we demonstrate in the Appendix A.4 one can indeed transform (5.44) into (4.7), and *vice versa*. (This further represents a consistency check of our approach.)

The main difference of (5.44) as compared to (4.7) (from Novikov's theorem), is that it involves statistical dependencies between the present and delayed system state and the delayed *velocity*. Instead, Eq. (4.7) implicitly involves the statistical dependencies among the system states at three different times via the three-time PDF $\rho_3(x,t;x_\tau,t-\tau,x_{2\tau},t-2\tau)$. This indicates that, instead of going back even further *into the past* with every member [by first taking into account $X(t-\tau)$, then $X(t-2\tau)$, then $X(t-3\tau)$ ...], the here derived hierarchy rather collects more and more information about the dynamics at time $t-\tau$ [by taking into account $X(t-\tau)$, then $\dot{X}(t-\tau)$]. This suits to the different pictures of the delayed dynamics provided by the two illustrations in Fig. 4.1.

As both representations describe the exact same process, this demonstrates that to (probabilistically) predict the future (after a time t), either knowledge of the system states at all times $t, t-\tau, t-2\tau, ...$ would be needed, or, complete knowledge about the preceding interval $[t-\tau, t]$. On the stochastic level of description, the process which underlies the non-Markovian one, can either be considered as a process of the state vector $\{X(t), X(t-\tau), X(t-2\tau), X(t-3\tau), ...\}$, or, equivalently (and actually more natural), of the state vector of all $X(s) \in \mathbb{R}^\infty$, with $s \in [t-\tau, t[$.

A technical advantage of the representation (5.44) is that it does not involve unknown functional derivatives w.r.t. the noise, as opposed to (4.7), and only contains one diffusion term w.r.t. x. In Chapter 7, we will utilize these technical

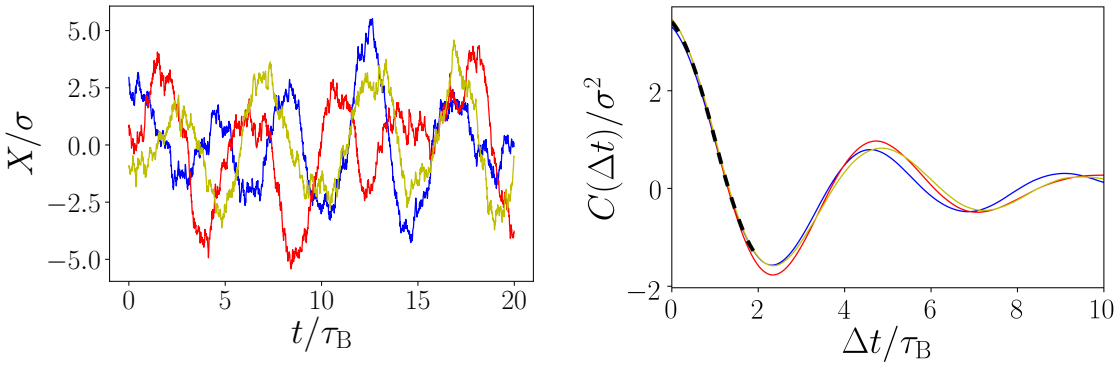

Figure 5.3: *Left panel: Three example trajectories of an Brownian particle subject to a linear time-delayed force $F_\mathrm{d} \propto -X(t-\tau)$ in the steady state. The LE is (5.46) at $\tau = \tau_\mathrm{B}$. Right panel: Corresponding position-autocorrelation function from the three single numerically generated trajectories, and additionally plotted by a dashed black line from the analytical results (4.23, 1.51) for $t \in [0, 2\tau]$, which we will explain in detail in Sec. 1.4.5. The delay force applies positive feedback to the colloid.*

advantages, and use (5.44) as starting point to find an approximate two-time PDF, $\rho_2(x, t; x_\tau, t-\tau)$. To our knowledge, this is the first approach of this kind.

Further perspective on this approach As a final remark, we would like to reconsider the Markovian embedding from a conceptual point of view. Non-Markovianity, in general, and specifically the discrete time delay considered here, is a consequence of an incomplete description of a system, i.e., missing degrees of freedom. The delayed force might either stem from a specifically designed external agent, a *feedback controller*, or, more generally, from any part of a more complex system (the "super-system"), which is not explicitly modeled within the considered description, but interacts with the system at hand in form of delayed forces. The Markovian auxiliary system, in which we here embed the delayed process, allows us to mimic the external agent (or other parts of the "super-system") without explicitly modeling it in detail. In this sense, the auxiliary variables can be interpreted as a "substitute" of the hidden degrees of freedom in the system. On the other hand, it is well-known that hidden degrees of freedom have a significant impact on the entropy production and other thermodynamic notions, see for example [294, 295]. Hence, it is interesting to study the consequences of our treatment of delayed systems from a stochastic-thermodynamical (and information-theoretical) perspective. This motivates the investigations presented in Part III, Chapter 9.

5.4 Side note: Discrete vs. distributed delay

The Markovian embedding from Sec. 5.1 has revealed that the discrete delay, white noise LE can be considered as the $n \to \infty$ limit of a LE (5.2) with distributed delay and colored noise. In particular, the dynamics of the first degree of freedom, X_0, in the $(n+1)$-dimensional network (5.1) is then identical to the time

delayed process (with delta-correlated memory kernel). Keeping n finite amounts to approximating the delayed force $F_\mathrm{d}[X(t), X(t-\tau)]$ in the LE with a force $F_\mathrm{d}[X(t), \int K(t-t')X_0(t')\mathrm{d}t' + \nu(t)]$ that involves a Gamma-distributed memory kernel (5.3) with finite width and colored noise 5.4 (if $\mathcal{T}' > 0$). For $n > 1$, the delay had similar features as the discrete delay case, i.e., a pronounced maximum around τ, and a vanishing kernel $K(t-t')$ at $t' = t$. On the contrary, the case $n = 1$, this yields a LE with an exponentially decaying delay [shown in Fig. 5.2 (a)]. To get an impression of the system with distributed delay, let us here compare the trajectories, focusing on $n = 1$, i.e., the most distinct case. In a second step, we briefly consider the corresponding PDFs.

5.4.1 Trajectories of the delayed Ornstein-Uhlenbeck process

For the comparison between discrete and distributed delay, we consider the simplest possible example, in particular, the linear delay equation

$$\gamma\dot{X} = -\int_{-\infty}^{t} K(t-t')X(t')\mathrm{d}t' + \sqrt{2D_0}\xi(t). \tag{5.45}$$

Discrete delay We start with the discrete delay, which is recovered at $K(t-t') = \delta(|t-t'|-\tau)$, yielding

$$\gamma\dot{X} = -(\gamma/\tau)X(t-\tau) + \sqrt{2D_0}\xi(t). \tag{5.46}$$

This is a special case of the linear delay LE considered in Sec. 1.4.5 and Sec. 4.2.2 (with $\alpha = 0$, $\beta = \gamma/\tau$). In regards to the colloidal particle in the trap with optical feedback, this corresponds to the case $k = -1$, $a = 1$, (i.e., positive feedback).

Figure 5.3 depicts some example steady-state trajectories and the associated position-autocorrelation function C of a Brownian particle subject to a linear feedback force, whose motion is given by (5.46).

As expected from the considerations in Secs. 1.4.5 and 4.2.2, the correlation function C shows damped oscillation, with negative regions $C(t') < 0$, i.e., anticorrelations. We have discussed this point in a similar model in Sec. 1.4.5. Here, the period is roughly 4τ. We further recall that oscillatory solutions of first-order scalar differential equations are indeed a characteristic of delay differential equations (see Sec. 1.4.4).

Distributed delay Let us now consider the *exponentially distributed delay*, specifically

$$\gamma\dot{X} = -(\gamma/\tau)\int_{-\infty}^{t} e^{-(t-t')/\tau}X(t')\mathrm{d}t' + \sqrt{2D_0}\,\xi(t), \tag{5.47}$$

which corresponds to the case $n = 1$ of the LE (5.2). To better see the impact of the distributed delay, we do not consider colored noise here ($\mathcal{T}' = 0$). Figure 5.4 shows

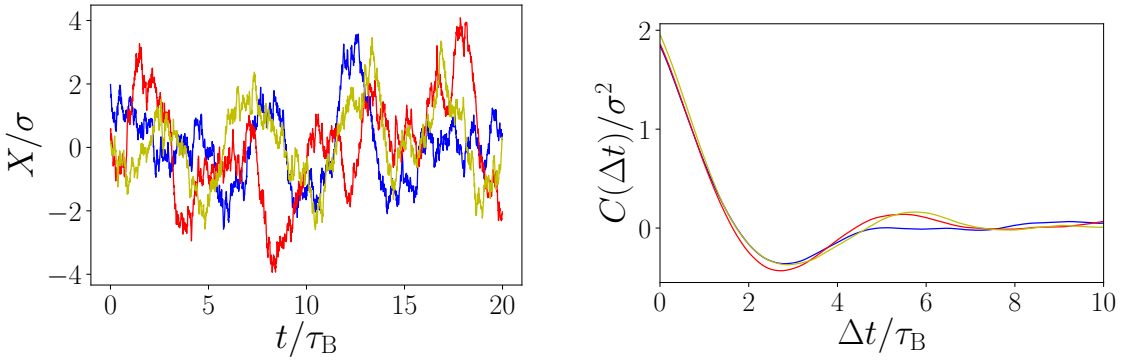

Figure 5.4: *Left panel: Three example trajectories of a Brownian particle subject to a linear feedback force involving distributed delay with an exponential memory kernel at $\tau = \tau_B$, in the steady state. The LE reads (5.47). Right panel: Corresponding temporal position-autocorrelation function from the three single numerically generated trajectories.*

some trajectories and the position-autocorrelations. Remarkably, the trajectories and correlations are found to be quite similar to the discrete delay in Fig. 5.3, despite the very different type of memory. To make both cases more comparable, the first moment of the memory kernel [given by the delay time in (5.46)] is the same. One likewise notices damped oscillatory behavior with a similar oscillation period and anticorrelations. However, the oscillatory component seems a bit less pronounced, and the decay is faster. Furthermore, the initial value $C(0) = \langle X^2 \rangle$, which gives the second moment of the steady-state PDF, is smaller, i.e., the density is *more located* in the case of distributed delay.

We will consider this type of memory in more detail in Part III of this thesis from a thermodynamic point of view. Let us now briefly consider the probability densities for discrete and distributed delay.

5.4.2 Probability densities in the presence of discrete and distributed delay

Above, we have compared the trajectories of LEs with discrete and distributed delay, corresponding to the cases $n = 1$ and $n \to \infty$. Let us now switch to the probabilistic perspective. We are particularly interested in the idea that keeping n finite readily implies an approximation for the PDF, which is the main objective of the remainder of Part II.

We recall that the dynamics is described as the first degree of freedom of the Markovian set of Langevin equations (5.1). This system can also be described by a self-sufficient set of $n + 1$ Fokker-Planck equations (5.27). By keeping n finite, Eq. (5.27) is a well-defined, closed FPE, which can be readily used to calculate the $(n + 1)$-time PDF by means of standard techniques for Markovian FPEs (at least numerically). By marginalization of the resulting PDFs, one would obtain an approximate PDF of the delayed process. On the level of the LE, approximating the delay process by keeping n finite is in fact a common method [32, 57, 293,

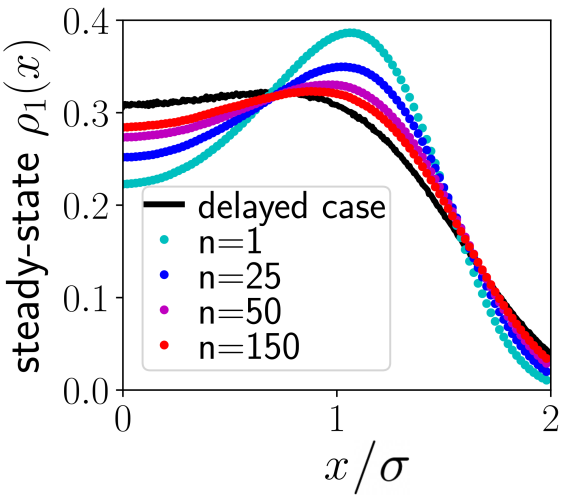

Figure 5.5: *Comparison of the one-time PDF for different values of n in the doublewell potential with linear delay force, preliminary numerical simulation results. At $n = 1$ (cyan symbols), the memory kernel (5.3) is exponentially decaying, while it is non-monotonic for all $n > 1$, with a maximum around τ. The black line plots the PDF of the system with discrete delay, which corresponds to the limit $n \to \infty$. Here, the parameters are set to $V_0 = 1.5\,k_\mathrm{B}\mathcal{T}$, $k = 1k_\mathrm{B}\mathcal{T}/\sigma^2$, $\tau = 1\,\tau_\mathrm{B}$, $\mathcal{T}' = 0.5\mathcal{T}$. Thanks to Simon Hermann for the numerical data for $n = 1, 25, 50, 150$.*

332]. However, a systematic investigation of this approximation (including an error estimate), has, to the best of our knowledge, not been performed yet.

Such an approximation has the advantage that one readily has a corresponding LE. In this sense, one can immediately "see" the consequences of this approximation on the stochastic level. This is very appealing from a conceptional point of view. However, while being closed, the resulting FPE (5.27) is still a set of multiple coupled partial differential equations, whose analytical solution is a very involved problem on its own. This is especially true for systems with nonlinear forces F.

As we are mostly focus on analytical approximations for the PDFs, we do not further proceed with this finite n idea in the present context. Let us still consider some exemplary, numerical results to get an idea of the impact of n on the probability densities. Figure 5.5 depicts some numerical results from this "approximation", i.e., the steady-state one-time PDF for different finite values of n, in the doublewell potential. These results indicate that the probability density converges, but it seems to converge relatively slowly. Our preliminary numerical investigations further suggest that the convergence is faster, if $\mathcal{T}' = \mathcal{T}$, i.e., if the temperature of the auxiliary variables matches the temperature of X. This is an interesting observation, which might be worth to further elucidate in future research.

As we have already mentioned, it is not obvious how to solve the set of $(n + 1)$ FPEs (5.27). It is noteworthy that this is because the coupling terms between the individual equations (e.g., the coupling between X_i and X_{i+1}) resemble non-reciprocal interactions, which we have introduced in Sec. 3.7.3. For this reason, there is for example no Boltzmann-like steady-state solution. We will come back to this observation in Chapter 9 and discuss its implications from a thermodynamic perspective in great detail. In Chapter 9, we will also consider the impact of changing n on thermodynamic and dynamical properties of the system, and reconsider the "discrete delay" limit $n \to \infty$.

In the following, we will discuss alternative approximation schemes that yield an approximate first member of the FP hierarchy, and even introduce an approach that operates on the level of the second member of the FP hierarchy. The focus is

on approaches that can be treated analytically.

6 | Force-linearization closure

In this Chapter, we turn to the problem of finding (approximate) probabilistic solutions of systems with delay. In particular, we discuss an approximation scheme for the steady-state one-time PDF, which we have introduced in Ref. [1], called *Force-linearization closure* (FLC).

Starting point The FLC applies to systems described by the delayed Langevin equation (1.42), recall

$$\dot{X}(t) = \gamma^{-1} F\left[X(t), X(t-\tau)\right] + \sqrt{2D_0}\,\xi(t).$$

Consistent with the example systems considered throughout this thesis, we assume that the total deterministic force F can be written as

$$F(x, x_\tau) = F_\mathrm{s}(x) + F_\mathrm{d}(x_\tau, x), \tag{6.1}$$

where the force $F_\mathrm{s}(x) = -\partial_x V_\mathrm{s}(x)$ stems from a static potential V_s, and F_d is a delayed force involving the system state at times t and $t-\tau$.

We further recall the corresponding FPE (2.15) for the one-time PDF ρ_1 [discussed in Sec. 2.2], which we here recast into the form

$$\partial_t \rho_1(x,t) = -\gamma^{-1}\partial_x\Big[\widehat{F}(x,t)\rho_1(x,t)\Big] + D_0\partial_{xx}\rho_1(x,t), \tag{6.2a}$$

with

$$\widehat{F}(x,t) = F_\mathrm{s}(x) + \int_\Omega F_\mathrm{d}(x,x_\tau)\rho_\mathrm{c}(x_\tau, t-\tau|x,t)\mathrm{d}x_\tau. \tag{6.2b}$$

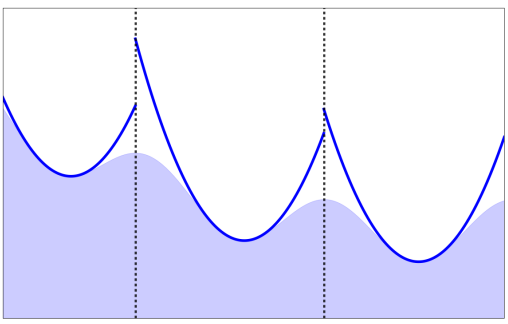

Figure 6.1: *Illustration of the main idea behind the Force-linearization closure, that is, piecewise approximating the static potential by parabolic functions in the higher members of the Fokker-Planck hierarchy. To this end, an effective static potential V_STAT is defined, see Eq. (6.3), and the spatial domain is chopped between its maxima.*

The FPE (6.2) is not self-sufficient as it involves ρ_c (which in turn depends on ρ_2). As we have discussed in Sec. 2.2, in fact, an infinite hierarchy of equations emerges, whose $(n+1)$st member depends on the n-time probability density $\rho_n(x,t;...;x_{n\tau},t-n\tau)$.

By finding suitable approximations for ρ_c, Eqs. (6.2) can be closed and the hierarchy truncated. The FLC represents such a closure scheme.

Main idea The FLC bases on the fact that the exact steady-state solution of the FPE hierarchy is known for systems where all deterministic forces are linear in x and x_τ [see Eqs. (4.19, 4.24)]. The main idea is to *linearize the deterministic forces* in all members of the infinite FPE hierarchy apart from the first one, in this way obtaining a closed approximate FPE for the one-time PDF ρ_1. Importantly, this linearization is carried out separately on suitable subdomains, as sketched in Fig. 6.1, which is crucial in situations where the potential landscape has multiple minima. In this way, we reduce the problem to a solvable, linear situation, keeping just enough aspects of the nonlinear original system to get an interesting result.

Having this main idea in mind, one would already expect –without further reasoning– that the FLC renders a good approximation, whenever the PDF is strongly centered around to the potential minima. This is, for instance, the case, when the potential barriers are high compared to thermal fluctuations. Indeed, our results presented at the end of this Chapter, confirm these expectations.

6.1 Details of the approximation

Here we present the details of the FLC, which is a multi-step procedure. To understand the results discussed in Secs. 6.3 and 6.4, it is, however, not necessary to be aware of all the mathematical details, and the Reader who is mainly interested in the general idea and the results may safely skip this Part and proceed immediately with Sec. 6.2.

We won't make further physical assumptions in the present Section. However, as will become clear towards the end of the Section, the expressions severely simplify for systems with natural boundary conditions and linear delay forces, which is the case in our example applications. The following text is closely related to Ref. [1].

6.1.1 Linearization of the deterministic forces

Slitting the space into subdomains We start by considering the history-dependent energy landscape corresponding to the deterministic force $F(x,x_\tau) = F_s(x) + F_d(x,x_\tau)$ in the LE (6.1). As a first step, we introduce a simple estimate for the total steady-state energy landscape. In particular, we assume at this point the system to be at rest, i.e., $X(t-\tau) \equiv X(t)$, such that the total (static) potential is given by

$$V_{\mathrm{STAT}}(x) = -\int^x F(x', x_\tau = x')\mathrm{d}x'. \tag{6.3}$$

We number each minimum $x_{0,i}$ of V_{STAT} with an integer $i \in \mathcal{I}$. To formulate the FLC, we split the spatial domain Ω into non-overlapping intervals Ω_i, such that each local maximum of V_{STAT} (as well as both boundaries of Ω) represents a bound of an interval, and every Ω_i contains exactly one minimum $x_{0,i}$. By this procedure, we obtain a splitting of Ω into subdomains Ω_i, whose union is again the entire spatial domain: $\Omega = \bigcap_{i\in\mathcal{I}} \Omega_i$.

Linearization We return to the actual deterministic force $F = F_{\mathrm{s}} + F_{\mathrm{d}}$, and perform for each i a Taylor expansion of F in both spatial variables x and x_τ. More specifically, we expand around the deviations from the enclosed minimum, i.e., $\Delta x_i = x - x_{0,i}$ and $\Delta x_{\tau,i} = x_\tau - x_{0,i}$. Neglecting all terms of quadratic orders $\mathcal{O}(\Delta x_i^2)$, $\mathcal{O}(\Delta x_i \Delta x_{\tau,i})$, $\mathcal{O}(\Delta x_{\tau,i}^2)$ or higher, we obtain

$$F^{\mathrm{lin},i}(x, x_\tau) = -\alpha_i \Delta x_i - \beta_i \Delta x_{\tau,i}, \tag{6.4}$$

where $-\alpha_i$ and $-\beta_i$ are the first-order derivatives of F with respect to x and x_τ, respectively, evaluated at the minimum $x_{0,i}$. [Note that the constant terms always vanish as we expand around the minima of V_{STAT} (6.3), thus, $F(x_{0,i}, x_{0,i}) = 0$.]

This procedure yields an approximation of the steady-state energy landscape composed of a sequence of quadratic polynomials, each subdomain Ω_i reaching from one local maximum to the following one (or to a bound of the entire spatial domain). An example is sketched in Fig. 6.1.

6.1.2 Analytical probabilistic solution for linearized forces

We now turn back to the full system which involves the original nonlinear deterministic forces and thermal noise, and apply the linearization procedure of Sec. 6.1.1 to the deterministic forces in the higher members of the Fokker-Planck hierarchy. The following steps are performed separately for each subdomain Ω_i. We first apply the linearization $F \approx F^{\mathrm{lin},i}$ [from Eq. (6.4)] to the force terms in all members of the infinite FPE hierarchy starting from the second, i.e., the equation for the two-time PDF $\rho_{2,\mathrm{ss}}(x, t; x_\tau, t-\tau)$.

Then, we utilize the linear solutions for the PDF, which we have discussed in Sec. 4.2.2. Assuming natural boundary conditions at every subdomain bound in the higher members, the FPEs for all $\rho_{\mathrm{n,ss}}$ can be solved by multivariate Gaussian distributions [77]. We have given the explicit results for $\rho_{1,\mathrm{ss}}^{\mathrm{lin}}$, $\rho_{2,\mathrm{ss}}^{\mathrm{lin}}$ and $\rho_{3,\mathrm{ss}}^{\mathrm{lin}}$ before in Eq. (4.19) and below. For each subdomain, the corresponding conditional PDF[1] reads

$$\rho_{\mathrm{c,ss}}^{\mathrm{lin},i}(x_\tau|x;\tau) = \sqrt{\frac{2K_i}{\pi}(d_{1,i})^2\left[1-(d_{2,i})^2\right]}\; e^{\frac{\Delta x_i{}^2}{2K_i}} e^{d_{1,i}\left[2d_{2,i}\,\Delta x_i\,\Delta x_\tau^i - \Delta x_i{}^2 - (\Delta x_\tau^i)^2\right]} \tag{6.5}$$

[1] Recall that $\rho_2(x; x_\tau) = \rho_{\mathrm{c}}(x_\tau|x)\rho_1(x)$.

with the coefficients

$$\Omega_i \equiv \quad \omega(\alpha^i, \beta^i) \quad = \sqrt{\alpha_i^2 - \beta_i^2}/\gamma \quad \in \mathbb{C}, \tag{6.6a}$$

$$K_i \equiv \quad K(\alpha_i, \beta_i, \tau) \quad = D_0 \frac{\gamma + (\beta_i/\omega_i)\sinh(\tau\omega_i)}{\alpha_i + \beta_i \cosh(\tau\omega_i)}, \tag{6.6b}$$

$$d_{1,i} \equiv \quad d_1(\alpha_i, \beta_i, \tau) \quad = \frac{\beta_i^2 K_i/2}{(\beta_i K_i)^2 - (D_0\gamma - \alpha_i K_i)^2}, \tag{6.6c}$$

$$d_{2,i} \equiv \quad d_2(\alpha_i, \beta_i, \tau) \quad = (D_0\gamma - \alpha_i K_i)/(\beta_i K_i), \tag{6.6d}$$

where $\Delta x_i = x - x_{0,i}$, with $x_{0,i}$ being is the enclosed minimum of V_{STAT} from Eq. (6.3). d_2 scales the "correlation" between Δx and Δx_τ as one can see from Eq. (6.5), and is the most important parameter [as we will see in Eq. (6.12) below]. Further, K is the variance of the respective one-time PDF $\rho_1(\Delta x)$, while ω scales how sensitive the variance K changes with τ. More details and the stability conditions of the solutions are given in Sec. 4.2.2.

With Eq. (6.5), the first member of the approximate FPE hierarchy accordingly reads

$$\gamma D_0 \, \partial_{xx} \rho_{1,\mathrm{ss}}^{\mathrm{FLC}}(x) = \partial_x \left[\widehat{F}^{\mathrm{FLC}}(x) \, \rho_{1,\mathrm{ss}}^{\mathrm{FLC}}(x) \right], \tag{6.7a}$$

with

$$\widehat{F}^{\mathrm{FLC}}(x) = \int_\Omega F(x, x_\tau) \rho_{\mathrm{c,ss}}^{\mathrm{lin},i}(x_\tau | x; \tau) \mathrm{d}x_\tau, \ \forall x \in \Omega_i, \tag{6.7b}$$

where, importantly, the deterministic force F is *not* linearized. Equations (6.5 - 6.7b) [together with the linearization rule (6.4)] form a closed set of equations. $\rho_{1,\mathrm{ss}}^{\mathrm{FLC}}$ from (6.7) represents the approximate steady-state one-time PDF we aimed for. Thus, the force linearization applied to the second and all higher members yielded a closure of the FPE hierarchy. This is the main idea of our approach.

Since $\widehat{F}^{\mathrm{FLC}}$ can be calculated separately [from Eq. (6.5)], the FLC formally converts the delayed, not self-sufficient equation for $\rho_{1,\mathrm{ss}}$ itself into a closed, *quasi*-Markovian FPE [Eq. (6.7a)]. The fact that the delay and the non-Markovianity are still somewhat present, becomes apparent by considering the special case of a linear force F. Then, our approximate FPE (6.7) is *identical* to the exact delayed FPE. This is a crucial difference as compared to earlier approaches known from the literature (discussed below, see Sec. 6.2) effectively yield Markovian descriptions.

Furthermore, for the sake of consistency, one can easily check that the usual Markovian FPE is recovered from Eq. (6.7) in the absence of delay (where $\widehat{F}^{\mathrm{FLC}} \to F$).

6.1.3 Vanishing steady-state probability current

The key equations of the FLC (6.5 - 6.7b) are valid for steady states of, in principle, arbitrary systems which can be meaningfully linearized according to the procedure described in Sec. 6.1.1. A particularly simple situation arises when the probability

current given by $J = [\widehat{F}^{\mathrm{FLC}}(x) - \gamma D_0 \partial_x]\rho_{1,\mathrm{ss}}^{\mathrm{FLC}}(x)$ vanishes[2]. If $J = 0$, the formal solution of Eq. (6.7a) takes the simple Boltzmann-like form [compare with Eq. (2.13)]

$$\rho_{1,\mathrm{ss}}^{\mathrm{FLC}}(x) = Z^{-1} e^{-V_{\mathrm{eff}}^{\mathrm{FLC}}(x)/(k_{\mathrm{B}}\mathcal{T})} \tag{6.8}$$

with the *effective static potential*

$$V_{\mathrm{eff}}^{\mathrm{FLC}}(x) = -\int_{\hat{x}}^{x} \mathrm{d}x' \widehat{F}^{\mathrm{FLC}}(x'), \tag{6.9}$$

where $\widehat{F}^{\mathrm{FLC}}$ is given in Eq. (6.7b), and $\hat{x} \in \Omega$ is arbitrary, but fixed. Here and in the following, we denote the normalization constant by Z.

We stress that the assumption $J = 0$ simplifies the analysis, but it is not a necessary condition. Also for non-zero currents, Eq. (6.7a) can be treated using standard techniques for (Markovian) Fokker-Planck equations [158, 173].

6.1.4 Specification to linear delay force

In some situations, the expressions further simplify severely. As a particularly important example in the present context, represent systems with vanishing probability current and linear delay force (1.44). The nonlinear models introduced in Sec. 1.5 fall into this class. Later, in Secs. 6.3 and 6.4, we will consider these nonlinear models as exemplary applications of the FLC.

The general expression for $\rho_{1,\mathrm{ss}}^{\mathrm{FLC}}$ from solution (6.8) is obtained as follows. The delay force is already linear and yields $\beta_i = -k, \forall i \in \mathcal{I}$, where β_i and k are the coefficients appearing in Eqs. (6.4) and (1.44), respectively. After performing several Gaussian integrals, Eq. (6.7b) yields

$$\widehat{F}^{\mathrm{FLC}} = -\partial_x V_{\mathrm{s}}(x) - k\Delta x_i + kx d_{2,i}\sqrt{2K_i d_{1,i}}\,\sqrt{1 - {d_{2,i}}^2}\, e^{\left[d_{1,i}{d_{2,i}}^2 - d_{1,i} + 1/(2K_i)\right]{\Delta x_i}^2}. \tag{6.10}$$

$\widehat{F}^{\mathrm{FLC}}$ can be further simplified by using the identity

$$d_{1,i}{d_{2,i}}^2 - d_{1,i} + 1/(2K_i) = 0, \tag{6.11}$$

where the quantities $d_{1,i}$ and $d_{2,i}$ are given in Eqs. (6.6c) and (6.6d). In this way, we obtain the piecewise defined effective potential

$$V_{\mathrm{eff}}^{\mathrm{FLC}}(x) = V_{\mathrm{s}}(x) + k/2(1 - |d_{2,i}|){\Delta x_i}^2, \tag{6.12}$$

for $x \in \Omega_i$. Together with (6.8), this yields the approximate PDF $\rho_{1,\mathrm{ss}}^{\mathrm{FLC}} = Z^{-1} e^{V_{\mathrm{eff}}^{\mathrm{FLC}}/(k_{\mathrm{B}}\mathcal{T})}$.

In the case of linear delay forces, the estimate for the steady state energy landscape further coincides with the static potential $V_{\mathrm{STAT}} = V_{\mathrm{s}}$ [since $F_{\mathrm{d}}(x, x) = 0$], such that $x_{0,i}$ and the bounds of Ω_i are readily determined by the extrema of V_{s}. Thus, for a given V_{s}, one only needs to calculate α^i by linearizing F_{s}, and therewith the coefficient $d_{2,i}$ [Eq. (6.6d)].

[2] This is, for example, automatically the case, when a system obeys natural boundary conditions.

6.2 Comparison to earlier approaches

As we have mentioned in Chapter 2.2, alternative schemes to obtain approximate one-time PDFs are known from the literature, particularly the small delay expansion and the perturbation theory (see Sec. 2.2.1). Let us show the resulting approximations for the one-time PDFs in the present case. Later, we will compare them with the FLC.

6.2.1 Small delay expansion

The small delay approximation on the level of the LE [82, 83] is essentially a Taylor expansion up to linear order of the total deterministic forces in powers of τ. It renders a time-local LE that approximately describes the dynamics, where τ appears as a parameter. We recall that this approach is, by construction, appropriate for very short delay times only. At the same time, this is a regime where the dynamics is qualitatively very similar to a Markovian one. Thus, we expect that many effects induced by the memory may not captured by this approximation.

For the case of linear delay force $F_\mathrm{d} = k(x_\tau - x)$ (which we will consider in Secs. 6.3 and 6.4), the small delay expansion gives rise to a Boltzmann-like one-time PDF, analogously to the FLC (6.8), with

$$\rho_{1,\mathrm{ss}}^{\mathrm{approx}}(x) = Z^{-1} e^{-V_{\mathrm{eff}}^{\mathrm{approx}}(x)/(k_\mathrm{B}\mathcal{T})}, \tag{6.13}$$

and the effective potential $V_{\mathrm{eff}}^{\mathrm{approx}} \equiv V_{\mathrm{eff}}^{\mathrm{small}\tau}$ given by [86]

$$V_{\mathrm{eff}}^{\mathrm{small}\tau}(x) = V_\mathrm{s}(x)/(1 - k\tau). \tag{6.14}$$

Thus, the small delay expansion renders an effective potential which is proportional to the static potential V_s. This further implies that also the approximate PDF has the same functional form as in the case $F_\mathrm{d} \equiv 0$ (given if k or $\tau = 0$). We will come back to this observation in short, and consider it from the perspective of effective temperatures [see Sec. 6.2.3 below].

6.2.2 Perturbation theory

Another approximation scheme introduced in earlier literature involves a first-order perturbation-theoretical ansatz for the density on level of the FPE. Here, the delay force F_d is regarded as a *small perturbation* to the non-delayed, Markovian dynamics, i.e., $|F_\mathrm{d}| \ll |F_\mathrm{s}|$. The resulting first-order equation has the same form as Eq. (6.2a), but the $\widehat{F}$-term [given in Eq. (6.2b)] contains the conditional probability density with respect to the unperturbed, i.e., non-delayed system with $F_\mathrm{d} \equiv 0$, which we will refer to as $\rho_c^{F_d \equiv 0}$. The latter follows a closed *Markovian* FPE, and consequentially, the combination of the resulting two FPEs is self-sufficient.

Thus, the FLC and the PT both render closed Fokker-Planck equations, where the conditional PDF $\rho_c(x_\tau|x)$ in the drift term is an approximate one. While in the FLC, it is approximated with the conditional PDF of the *linearized, delayed* system; the PT involves the one corresponding to the (nonlinear) *non-delayed* system, $\rho_\mathrm{c}^{F_d \equiv 0}$.

From a practical perspective, the PT is particularly appealing for systems where an analytical expression for $\rho_{\mathrm{c}}^{F_d\equiv 0}$ is available. However, such solutions are known, in fact, only for very few nonlinear static forces F_{s}. For this reason, the application of the PT approach usually requires additional approximations.

Short time propagator As suggested in Ref. [87], one can combine the perturbation-theoretical ansatz with the *short time propagator* (s.t.)

$$\rho_{\mathrm{c}}(x_\tau|x;\tau) = \frac{1}{\sqrt{4\pi D_0 \tau}} \exp\left[-\frac{\{x_\tau - x - F_{\mathrm{s}}(x)\tau/\gamma\}^2}{4D_0\tau}\right], \tag{6.15}$$

which is derived and discussed in Ref. [173] on p. 73 and prior. The usage of the short time propagator implies a first-order approximation in τ [86, 173].

For the case of a linear delay force $F_{\mathrm{d}} = k(x_\tau - x)$, the resulting one-time PDF is a Boltzmann distribution, like Eqs. (6.8, 6.13), with the effective potential

$$V_{\mathrm{eff}}^{\mathrm{PT}}(x) = (1 + k\tau)V_{\mathrm{s}}(x). \tag{6.16}$$

Interestingly, when the short time propagator is used for $\rho_{\mathrm{c,ss}}^{F_d\equiv 0}$, the PT approach yields a qualitatively very similar expression as the small delay expansion [compare Eqs. (6.14) and (6.16)]. According to both approximations, the only effect of the delay force, is a constant (yet different) factor within the exponent of $\rho_{\mathrm{1,ss}}^{\mathrm{approx}}$. In fact, the small delay expansion and the PT with short time propagator even become equivalent for small τk, since then these factors coincide. This can be seen from a Taylor expansion around $\tau k = 0$, which yields $1/(1 - k\tau) = (1 + k\tau) + \mathcal{O}([\tau k]^2)$.

Ornstein-Uhlenbeck approximation Alternatively to the short-time propagator, we propose to approximate $\rho_{\mathrm{c,ss}}^{F_d\equiv 0}$ by the conditional PDF from the corresponding unperturbed ($F_d \equiv 0$), linearized system, i.e., the corresponding Ornstein-Uhlenbeck (OU) process, given in Ref. [173] (p. 100). We have introduced this idea in [1]. The resulting effective potential from the PT approach reads

$$V_{\mathrm{eff}}^{\mathrm{PT}}(x) = V_{\mathrm{s}}(x) + (k/2)\Big[1 - e^{-(\alpha^i - k)\tau/\gamma}\Big](\Delta x_i)^2, \ \forall x \in \Omega_i. \tag{6.17}$$

As opposed to the result obtained with the short-time propagator [Eq. (6.16)], the effective potential in Eq. (6.17) has a different functional form than V_{s}. More specifically, the delay force now effectively adds to the static potential a fixed quadratic potential around the minima of V_{s}. Since the linear delay force [Eq. (1.45)] indeed stems from a quadratic confining potential (with history-dependent center), this effective potential appears to be more realistic than Eq. (6.16). However, the history-dependency of the position of the "trap" imposed by F_{d} is, of course, not captured by this Markovian approximation.

As one can see from Eq. (6.17), this PT-OU approach also requires a multi-step procedure, where one first needs to define suitable subdomains, just like in the FLC. Further, one might intuitively expect this approach to be cruder than the FLC, as it amounts to neglecting the delay as well as the nonlinearity of the forces, in the higher

members of the hierarchy. Indeed, this approach renders a poorer approximation compared to the FLC, as we will show in the following. Only in the regime of very small delay force magnitudes $|F_\mathrm{d}| \ll |F_\mathrm{s}|$, the results are comparable.

6.2.3 Effective temperatures

At this point, we briefly resume the discussion from Sec. 4.2.3, about of the notion of effective temperatures. In Sec. 4.2.3, we have discussed purely linear delay equations only.

Among the expressions above, the physically relevant quantities are the PDFs. Our notation and wording indicate the standpoint that the one-time PDF describes the diffusion of a particle in an "effective potential" at the (original) temperature $\mathcal{T}$. Thus, the feedback force alters the energy landscape but not the temperature. However, considering the results of the small delay expansion or the PT with short-time propagator [given in Eqs. (6.13,6.14,6.16)], the Boltzmann-distributed PDFs could just as well be interpreted as describing a particle in the (original) potential V_s at the *effective temperature* $\mathcal{T}(1-k\tau)$, or $\mathcal{T}/(1+k\tau)$, respectively. This offers a convenient view on the problem, and the concept of effective temperatures is indeed quite common in other contexts. According to both approximations, the effective temperature can either be reduced or increased due to the delay force, depending on the sign of k. We note that the effective temperature can also become negative for large magnitudes of $|k\tau|$, where the approximations break down.

However, this viewpoint loses meaning for the other approaches, where the effective potentials have a different functional form than V_s. This includes the FLC [see Eq. (6.12)] and the other version of the PT, with Ornstein-Uhlenbeck approximation (6.17). The fact that these other approaches yield indeed more accurate PDFs (as we show below), is an argument against the notion of an effective temperature in the present context.

Furthermore, as we have shown in Sec. 4.2.3, the concept of effective temperature is also not appropriate when one considers higher-time PDFs. In Sec. 4.2.3, we have explicitly shown this for the two-time PDF of the linear system.

6.3 Application to the periodic potential

We apply the different approximations to the periodic washboard potential (6.18) introduced in Sec. 1.5.2, which is given by a cosine function

$$V_\mathrm{s}(x) = -(V_0/2)\,\cos(x/\sigma), \tag{6.18}$$

with barriers of height V_0. We further consider a linear delay force (1.44), i.e.,

$$F_\mathrm{d}(x, x_\tau) = -k\,(x - x_\tau).$$

If $k \geq 0$, the force is always directed towards the delayed position counteracting a deviation between the instantaneous and delayed particle position, i.e., negative feedback, and can be associated with the delayed confining potential $V_\mathrm{d}(x, x_\tau) =$

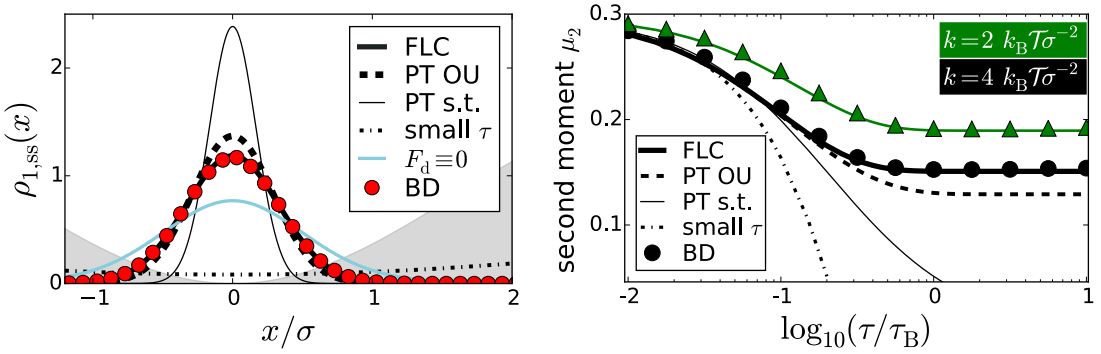

Figure 6.2: *Steady-state PDF $\rho_{1,\mathrm{ss}}$ within one well of the periodic washboard potential with linear delay force from several approximations. Symbols: numerical data (BD), thick solid lines: force-linearization closure (FLC), dashed lines: perturbation theory with Ornstein-Uhlenbeck approximation (PT OU), thin lines: PT with short time propagator (PT s.t.), dashed-dotted line: small delay expansion (small τ). Left panel: The PDF itself, at delay force strength $k = 8\,k_\mathrm{B}\mathcal{T}/\sigma^2$ and delay time $\tau = \tau_\mathrm{B}$. The light blue line additionally shows the PDF $\rho_{1,\mathrm{ss}}(x) \propto \exp[-V_\mathrm{s}(x)/(k_\mathrm{B}\mathcal{T})]$ for the case without delay force ($F_\mathrm{d} \equiv 0$). Right panel: The second moment μ_2 of the PDF. The green color refers to $k = 2\,k_\mathrm{B}\mathcal{T}/\sigma^2$, while the black color refer to $k = 4\,k_\mathrm{B}\mathcal{T}/\sigma^2$. The lines for PT and small τ are only displayed for the second parameter case for sake of clarity. The barrier height is set to $V_0 = 8\,k_\mathrm{B}\mathcal{T}$. These results are published in [1].*

$(k/2)(x - x_\tau)^2$. We recall that such quadratic feedback potentials are commonly used to model optical traps, see Sec. 1.4.1. Thus, we will focus here on $k > 0$ meaning nonnegative feedback strengths[3].

Approximations The approximations from the small delay expansion and the PT are readily given by the Boltzmann-distributions with effective potentials (6.14) and (6.16, 6.17), respectively. To apply the FLC, a more elaborate procedure is required. However, due to the linearity of the feedback (1.44), we can now make use of the results from Sec. 6.1.4, where we have already carried out the main steps.

The only remaining ingredient is splitting the domain by the procedure described in 6.1.1. As $F_\mathrm{d}(x, x) = 0\,\forall x \in \Omega$, V_STAT is identical to V_s. Hence, $x_{0,i}$ and the interval bounds of Ω_i are given by the extrema of V_s. Thus, we only need the minima of the washboard potential, which are located at $x_{0,i} = 2\pi\sigma i$, $i \in \mathbb{Z}$. We split the domain $\mathbb{R}$ in subdomains $\Omega_i = [(2i-1)\pi\sigma, (2i+1)\pi\sigma]$, such that each interval contains one minima $x_{0,i}$, with $i \in \mathcal{I} = \mathbb{Z}$. Now, linearizing the static potential yields $\alpha^i = V_0/(2\sigma^2) + k$ (and $\beta^i = -k$) for all $i \in \mathcal{I}$. Thus, on all subdomains Ω_i, the density is given by the Boltzmann distribution (6.8) with effective potential (6.12), where

$$d_{2,i} = \frac{\gamma\omega\cosh(\tau\omega) - [V_0/(2\sigma^2) + k]\sinh(\tau\omega)}{\gamma\omega - k\sinh(\tau\omega)} \tag{6.19}$$

with $\gamma\omega\sigma = \sqrt{(V_0/2\sigma)^2 + kV_0}$ and $\Delta x_i = x - x_{0,i}$.

[3] In later Chapters, we will also consider negative values of k.

6.3.1 Discussion of results

Figure 6.2 (left panel) displays the steady-state one-time PDF, which is a periodic function in this case. In particular, it shows one peak of the PDF sitting in one potential well. To be more precise, we here consider the "reduced PDF", i.e., the normalization is one within each well, see Sec. 1.5.2. Besides the different approximation results, also data generated by BD simulations of the delayed LE (1.42) are shown. One clearly sees that the FLC renders the best approximation of the PDF. Similar convincing results are found for all other considered values of τ and k in the tested range $\tau/\tau_\mathrm{B} \in [10^{-2}, 10]$ and $k/(\sigma^{-2} k_\mathrm{B}\mathcal{T}) \in [0, 32]$. To compare approximation and simulation results more systematically, we calculate the moments μ_n of the (reduced) PDFs within one potential well (for x values between two maxima of V_s). Due to the spatial inversion symmetry of the total potential around the enclosed minimum, all odd moments automatically vanish. We hence consider the second central moment $\mu_2 = \langle (X-\mu_1)^2 \rangle_\mathrm{ss}$ (with $\mu_1 = \langle X \rangle_\mathrm{ss}$ and $X/\sigma \in [-\pi, \pi]$). Figure 6.2 (right panel) shows μ_2 as a function of the delay time τ for two exemplary values of delay force strength k. We find that in the presence of the delay force, the variance of the distribution is reduced. Thus, the thermal fluctuations are suppressed by the delay force. This is not surprising, since the delay force arises from a quadratic potential which "traps" the particle. Within the considered parameter regime, the variance is seen to decrease with increasing τ, until at about $\tau_\mathrm{sat}^\mathrm{W} \approx 0.5\,\tau_\mathrm{B}$, where the variance appears to saturate. (Due to strongly increasing simulation times, we have, however, not investigated the behavior for much larger τ than shown here, but we will discuss our expectations below.)

On the level of the PDFs and the FPE, the saturation of μ_2 at large τ indicates that the delay-averaged drift $\widehat{F}$, and hence $\rho_\mathrm{c,ss}(x_\tau|x;\tau)$, are essentially constant for $\tau > \tau_\mathrm{sat}^\mathrm{W}$, in the considered range of delay times. We recall that we have already observed such a saturation of the PDFs for the linear delayed system, see Fig. 4.4 and below. The fact that the conditional probability does not change when the time difference τ is further increased, suggests that a relaxation mechanism within the valley takes place, where the history of the stochastic process becomes "washed out" on the level of ensemble averaged quantities. In this context it is interesting to note that, for the case without delay force ($k=0$), the relaxation time within a potential well can roughly be estimated to be around $\tau_\mathrm{ir} \approx V_\mathrm{s}/\gamma = 0.25\,\tau_\mathrm{B}$, see Sec. 1.6. The saturation of μ_2, thus, sets in *after* the density relaxation within a potential well. This provides a "hand having" argument to understand the observed saturation.

We, however, expect a different behavior when τ has an even much larger value (than in Fig. 6.2); in particular, when τ gets into the range of the mean escape times. Since in this parameter regime, the particle has a high probability to escape the well, transport between the potential valleys becomes important. We thus expect again a τ-dependency of $\rho_\mathrm{c,ss}$ and therewith of μ_2. We will discuss the escape processes later in Sec. 6.5. This behavior is not captured by the FLC which yields a saturated PDF for all $\tau \to \infty$.

Finally, we compare the FLC results with the small delay (Taylor) expansion [Eq. (6.13) and (6.14)] and from the PT approach in more detail. It is obvious that the small delay expansion fails completely for the considered delay times (un-

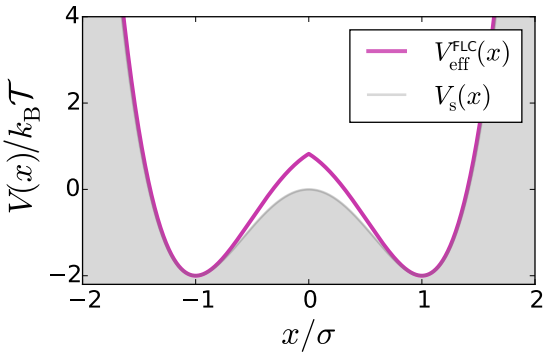

Figure 6.3: *Gray shaded: Bistable potential V_s (1.52), magenta line: the corresponding effective potential $V_\mathrm{eff}^\mathrm{FLC}$ from the FLC given in Eq. (6.12). The parameters are $V_0 = 8k_\mathrm{B}\mathcal{T}$, $k = 8\,k_\mathrm{B}\mathcal{T}/\sigma^2$ and $\tau = \tau_\mathrm{B}$.*

less $\tau \to 0$). Within the PT, we have either used the short time propagator [see Eq. (6.16)], or the corresponding OU approximation [see Eq. (6.17)]. As visible in both panels of Fig. 6.2, the PT generally overestimates the height of the density peak and underestimates its width. Furthermore, approximating the conditional PDF with the short time propagator yields worse results than using $\rho_\mathrm{c,ss}$ of the corresponding Ornstein-Uhlenbeck process. This observation matches our expectations, as the latter scheme seems to be the more sophisticated and accurate one. When the OU approximation is used, the PT at least reproduces the quantitative behavior of the function $\mu_2(\tau)$, see Fig. 6.2 (right panel).

Markovian limits If $\tau \to 0$ or $k \to 0$, the delay force vanishes and the system is Markovian. Thus, one expects the system to reach thermal equilibrium where the one-time PDF is a Boltzmann-distribution $\rho_1 \propto e^{-V_\mathrm{s}/(k_\mathrm{B}\mathcal{T})}$ (depicted as light blue line in the left panel of Fig. 6.2). Indeed, all approximations considered here correctly describe the limits, as can be seen from Fig. 6.2, right panel [and as is also evident from Eqs. (6.16, 6.14, 6.12)].

Conclusion Among the tested approaches, the FLC generates the best approximation of the steady-state one-time PDF in the periodic potential, especially in the regime of large τ or large k. On the one hand, the small delay expansion involves a truncated Taylor expansion in τ, making it plausible that this approach fails for large delay times. On the other hand, in this regime the "perturbation", i.e., $|F_\mathrm{d}|$ is not small compared to the force applied by the static potential, such that the basic assumption of the perturbation theory is not fulfilled. It is noteworthy that the PT-OU nevertheless yields quite reasonable results. The FLC is a reliable approximation for the intrawell dynamics even for large τk. This is because it renders an approximate, but still *non-Markovian*, description. We will discuss below, under which conditions the FLC breaks down.

6.4 Application to the bistable potential

As a second example, we consider the doublewell potential (1.52) introduced in Sec. 1.5, namely

$$V_\mathrm{s}(x) = V_0\,[(x/\sigma)^4 - 2\,(x/\sigma)^2],$$

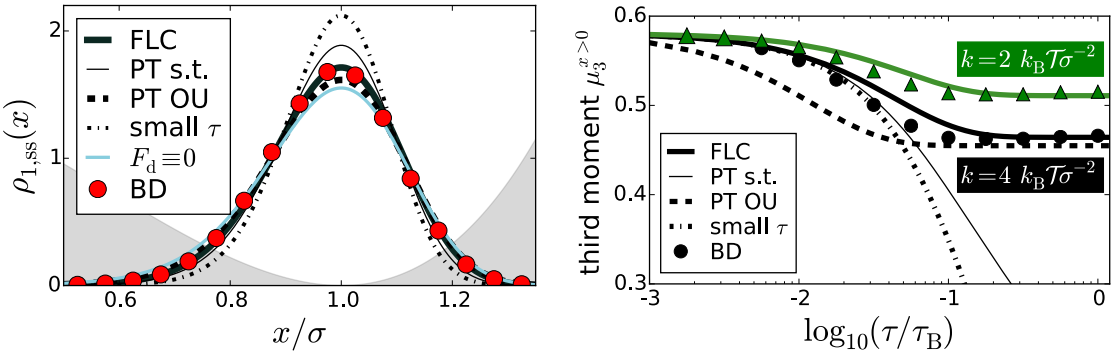

Figure 6.4: *Left: One peak of the bimodal PDF $\rho_{1,\mathrm{ss}}$ in the doublewell potential. The delay force amplitude and delay time are set to $k=18\,k_\mathrm{B}\mathcal{T}/\sigma^2$ and $\tau=0.1\,\tau_\mathrm{B}$, respectively. Right: Third moment $\mu_3^{x>0}$ of the PDF within the right well of the doublewell potential, i.e., of $\rho_{1,\mathrm{ss}}(x>0)$. The color code and parameters are as those in Fig. 6.2. These results are published in [1].*

with a barrier of height V_0 at $x/\sigma = 0$, and two minima at $x_{0,i} = i\sigma$ with $i \in \mathcal{I} = \{-1, 1\}$. We further reconsider the linear delay force (1.44). The application of the different approximations works analogously to the washboard potential described in the last section.

More specifically, for the FLC, we again have to employ the procedure described in Sec. 6.1.1, yielding $\alpha^i = 8\,V_0/\sigma^2 + k$ for both minima, and $\beta^i = -k$. For $x/\sigma \in [-\infty, 0]$ and $x_{0,i=-1} = -\sigma$, or $x/\sigma \in [0, \infty]$ and $x_{0,i=1} = \sigma$, the effective potential $V_\mathrm{eff}^\mathrm{FLC}$ is given by Eq. (6.12) with

$$d_{2,i} = \frac{\gamma\omega\cosh(\tau\omega) - [(8V_0/\sigma^2) + k]\sinh(\tau\omega)}{\gamma\omega - k\sinh(\tau\omega)} \tag{6.20}$$

and $\gamma\omega\sigma = \sqrt{(8V_0/\sigma)^2 + kV_0}$. The resulting effective potential is displayed in Fig. 6.3.

6.4.1 Discussion of results

As demonstrated in Fig. 6.4, the FLC again yields a very good approximation of the PDF, which is a bimodal distribution in this case. Contrary to Fig. 6.2, we consider for this example the third moment (because the second moment is here found to be barely affected by the delay force). More specifically, we consider the skew, i.e., the third central moment normalized with the standard deviation, of one peak of the bimodal PDF, which is defined as $\mu_3^{x>0} = \int_{x>0}[(x-\mu_1)/\sqrt{\mu_2}]^3\rho_{1,\mathrm{ss}}(x)\mathrm{d}x$. The comparison with the numerical data in Fig. 6.4 reveals that the FLC approach renders reasonably good predictions of $\mu_3^{x>0}$. We moreover find convincing results for all other tested parameters within the studied ranges of $\tau/\tau_\mathrm{B} \in [10^{-2}, 5]$ and

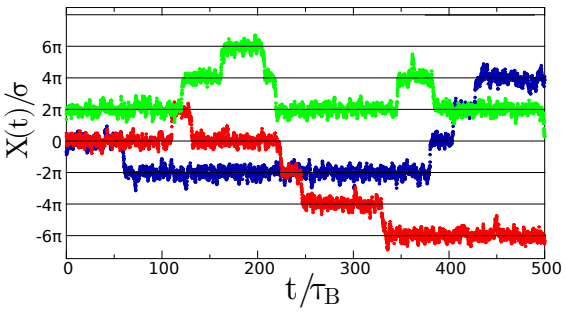

Figure 6.5: *Exemplary random trajectories in the washboard potential with $V_0 = 4\,k_\mathrm{B}\mathcal{T}$, in the absence of delay force ($k = 0$), from BD simulations. The horizontal lines are at the positions of the potential minima. One can see from these trajectories that the dynamics is characterized by intrawell fluctuations around the (current) potential minimum, on the one hand, and by random (interwell) jumps between the potential wells, on the other hand.*

$k/(\sigma^{-2}k_\mathrm{B}\mathcal{T}) \in [0, 18]$[4].

However, the FLC is not as accurate as it is in the case of the periodic potential (compare with Fig. 6.2). One reason might be the fact that the linearized potential "seen" by the higher-order PDFs is, by construction, symmetric with respect to the enclosed minimum (within each subdomain). In this sense, the doublewell potential, which is asymmetric around each minimum, is not as well approximated as the symmetric washboard potential. This manifests in a larger magnitude of the higher order Taylor terms which are neglected within the linearization procedure. We also note that the third moment is *per se* very sensitive against inaccuracies, since it involves cubic terms in the relative position. The deviations between the FLC and the exact result are hence expected to be larger.

Regarding the impact of increasing k and τ, we see from Fig. 6.4 (right panel) that the skew $\mu_3^{x>0}$ becomes smaller, which means that the probability distribution becomes more symmetric. This is not surprising since the pure delay potential is quadratic (i.e., symmetric) in the system state variable. Moreover, similar to the second moment in the washboard potential (see Fig. 6.2), $\mu_3^{x>0}$ approaches a saturation value within the considered parameter range. The corresponding delay time is about one order of magnitude smaller than in the washboard case: $\tau_\mathrm{sat} \approx 0.1\tau_\mathrm{B}$. This value of τ_sat is, as in the case of the washboard, significantly larger than that for the intrawell relaxation time τ_ir, which can be roughly estimated to be around $\tau_\mathrm{ir} \approx 1/64\tau_\mathrm{B}$ for the present case (see Sec. 1.6). Thus, the saturation of μ_3 can be explained by the same arguments as in the case of the washboard potential (Sec. 6.3).

Figure 6.4 also show the results according to the PT [Eqs. (6.16) and (6.17)], and from small delay expansion, see Eq. (6.14). Very similar to the case of the washboard, only the PT-OU is capable of reproducing the quantitative behavior (especially for large τ), and the FLC clearly provides the best approximation of the one-time PDF.

[4]This range spans from rather small delay forces that merely affect the shape of the energy landscape, to k values so high that the total potential has no second minimum due to the superimposed delayed potential. For all these qualitatively different cases, the agreement between FLC and numerical results is convincing.

Conclusion The FLC yields reasonable predictions for the one-time PDFs in quite large parameter regimes, and in both considered examples. Also the PT-OU approach is quite successful. In contrast, the PT with short time propagator is only appropriate for very small values of the delay time, where the system indeed behaves more or less like a Markovian one. The same thing can be said about the small delay expansion.

At the same time, the FLC and the PT-OU are more complicated to apply. They are not simply given by a single equation, but require a procedure with multiple steps (determination of V_{STAT}, finding the subdomains, then determining the parameters of the estimate). Thus, the more accurate approximations come at the "cost" of the more involved formalism. Among these two, the FLC is even superior as compared to the PT-OU, clearly yielding the best results.

A disadvantage of the FLC is, however, that it does not give an explicit estimate of the involved total error in the PDF. As opposed to this, the PT, for example, is a *first order* estimation in the "perturbation". Such a statement is not available for the FLC. But, due to the required secondary approximation in the PT approach, i.e., the Ornstein-Uhlenbeck approximation or the short-time propagator, this "error-estimation", is indeed also not readily available for the PT.

We finally add a comment concerning the limitations of these statements. First, we observe a bad performance of the FPE-based approaches, i.e., the PT and FLC, if k is negative, corresponding to a delay force that pushes the particle away from its delayed position (i.e., positive feedback). Interestingly, then, there even exists a parameter regime, where the small delay expansion outperforms the other approaches (for small delay times and negative k). An example will be given in the next Chapter (see Figs. 7.5, 7.6). Second, all of these approaches break down in parameter regimes, where the particle is very likely to escape its potential valley and delay induces interwell oscillations. The effectively Markovian descriptions given by the small delay expansion and the PT clearly fail as the delay force and delay time have large amplitudes; but also the FLC is expected to fail, because in these regimes the interplay of nonlinearities and memory matters. We will turn to the interwell jump dynamics in the next Section.

Thermodynamic application As we will explicitly show in Chapter 8, these approximations put as in the position to analytically calculate thermodynamic quantities in systems with discrete delay.

6.5 Estimation of escape times

6.5.1 Interwell dynamics

So far, we have focused on the intrawell dynamics, i.e., the fluctuations of the particle position X within one potential valley. However, an important feature of the studied systems is that the particle can also overcome the potential barriers and escape its valley, yielding *interwell dynamics*. The energy for such a jump is provided by the random thermal fluctuations. Due to the Gaussianity of the noise, an escape is, in

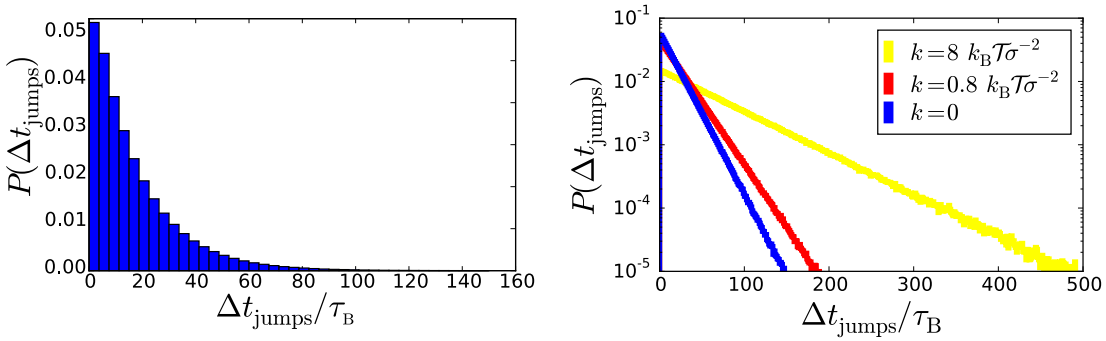

Figure 6.6: *Waiting time distribution in the doublewell potential, with ($k > 0$) and without ($k = 0$) linear delay force, for a short time delay $\tau = 0.1\tau_{\text{B}}$. The distributions are numerically (BD) generated (normalized) histograms of the times between sequential jumps Δt_{jumps}. Left panel: $k = 0$. Right panel: logarithmic plots with $k = 0$ shown again as blue histogram (we here only display the envelope for sake of clarity), red is the histogram for $k = 0.8 k_{\text{B}}\mathcal{T}/\sigma^2$, yellow is $k = 8 k_{\text{B}}\mathcal{T}/\sigma^2$. All distributions are found to decay exponentially. The barrier height is set to $V_0 = 4\, k_{\text{B}}\mathcal{T}$.*

fact, expected to happen eventually, no matter how high the barriers are compared to the thermal energy $k_{\text{B}}\mathcal{T}$ (which scales with the temperature $\mathcal{T}$). However, the expected *time* that passes until the jump occurs will, of course, depend on this very ratio $k_{\text{B}}\mathcal{T}/V_0$. Moreover, in the non-Markovian systems considered here (i.e., when $k\tau \neq 0$), the energy landscape is actually history-dependent, which may induce more intriguing jump dynamics than the one found in a Markovian system and dependency of the waiting times upon k and τ.

Markovian case For the Markovian case ($k\tau = 0$), the total time between two jumps Δt_{jump}, is described by the famous Kramers theory [254, 255], which we have briefly discussed in Sec. 1.6.1. This theory[5] predicts exponentially decaying waiting time probabilities, with a mean escape time given by Arrhenius formula [given by Eq. (1.58)]. We indeed find such an exponential decay for the considered systems, as displayed for an exemplary parameter setting in Fig. 6.6 (blue histograms).

Non-Markovian case Figure 6.6 further shows exponentially decaying distributions in non-Markovian cases where the delay force is present. Thus, the functional form is here identical to the Markovian case. However, the delay force yields a slower decay, corresponding to the particle staying longer in one well. As expected, this effect is enhanced by increasing k, i.e., by stiffening the feedback trap. The fact that the delay force does not change the functional form of the distributions motivates us to investigate the question: "*Is it possible to estimate the mean escape times in the non-Markovian models by similar means like in the Markovian case?*" To tackle this problem, we will utilize the FLC approach, which brings us into the position

[5] which involves several approximations, e.g., that thermal energy is low compared to barrier height, such that the jumps occur rather seldom

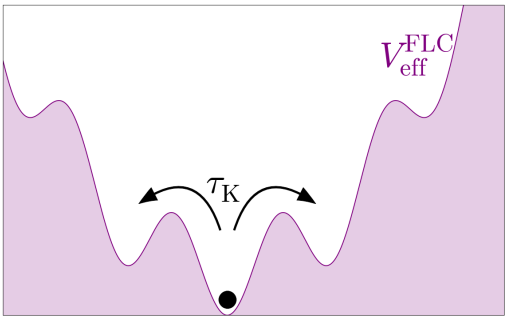

Figure 6.7: *Schematic illustration of the idea behind the Kramers-FLC estimate of the mean escape times τ_{K} in the washboard potential with linear delay force. Here, we pretend to particle to move in static (effective) potential, while it actually moves within a history-dependent energy landscape.*

to apply the Kramers theory[6]. The results of this idea are presented in the next paragraph.

As we will later show, the similarity of the waiting time distributions between Markovian and non-Markovian case is limited to the regime of short delay times. At the end of this Chapter, we will discuss situations, where the non-Markovian effects yield qualitatively different jump dynamics. In this regime, the Kramers-FLC approach must break down, since one of the main assumptions of Kramers theory is violated, i.e., that the jumps should be rather seldom. More importantly in the present context, here also the FLC is expected to fail as we have mentioned before. This is because the dynamics is dominated by the interplay of memory and the nonlinearities, which is neglected in this approximation.

6.5.2 The Kramers-FLC estimate

The escape time τ_{K} denotes the average time X spends in the vicinity of a potential minimum, until it leaves the valley by jumping to an adjacent one. We recall that the Kramers theory, which we have discussed in Sec. 1.6.1, was developed for Markovian systems characterized by a *static* potential landscape U, gives an approximation for τ_{K}. When the potential barriers ΔU are large compared to the thermal energy, i.e., $\gamma D_0 \ll \Delta U$, the Kramers theory provides an estimation for the escape rate r_{K} by the Arrhenius formula given in Eq. (1.58); recall $r_{\text{K}} = \sqrt{U''(x_{\min})|U''(x_{\max})|}/(2\pi\gamma)e^{-\Delta U/\gamma D_0}$. This estimate for r_{K} relies on a quasi-static approximation for ρ_1[7] For the washboard potential, the resulting mean escape time can be approximated by $2\tau_{\text{K}} = 1/r_{\text{K}}$ due to the two symmetric ways to leave a valley (see Ref. [255]). For the doublewell potential with the unique direction to exit each valley, $\tau_{\text{K}} = 1/r_{\text{K}}$.

Main idea For non-Markovian systems, the energy landscape is history-dependent, and there is no direct route to estimate the mean escape times by Kramers theory. However, in the FLC approach, the first member of the FPE

[6]By construction, escape times involve two or more locations, and multiple instances in time, and are therefore naturally connected to PDFs ρ_n with $n > 1$. Within the FLC, the higher order PDFs $\rho^{\text{FLC}}_{(\text{n}>2),\text{ss}}$ are multivariate Gaussian densities belonging to quadratic static potentials which have only one valley and, thus, no escape process. Therefore, there is no direct way to calculate the mean escape time on the basis of the FLC approach alone without Kramers theory.

[7]and does not involve ρ_2, or any higher order correlation between different times and spaces.

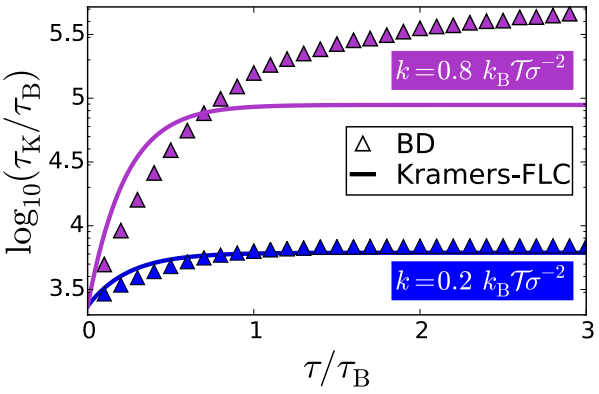

Figure 6.8: *Mean escape times τ_K in the washboard potential with $V_0 = 8\,k_B\mathcal{T}$, as a function of the delay time τ (in units of the Brownian timescale τ_B), in a logarithmic plot. Blue: delay force strength $k = 0.2\,k_B\mathcal{T}/\sigma^2$, Purple: $k = 0.8\,k_B\mathcal{T}/\sigma^2$, symbols: (BD) simulation results, lines: Kramers-FLC estimate as described in Sec. 6.5.2. These results are published in [1].*

hierarchy Eq. (6.7a) is formally identical to a Markovian (non-delayed) FPE for $\rho_{1,\mathrm{ss}}$ with a static effective potential $V_{\mathrm{eff}}^{\mathrm{FLC}}$ given by Eq. (6.9). Thus, we can in fact directly apply the Arrhenius formula (1.58) with $U \equiv V_{\mathrm{eff}}^{\mathrm{FLC}}$. For the linear delay force in our examples, one just needs to substitute $\Delta U = V_0 + (k/2)(x_{\max} - x_{\min})$ and $U''(x_{\mathrm{ex}}) = V_{\mathrm{s}}''(x_{\mathrm{ex}}) + k(1 - |d_{2,i}|)$. Taken altogether, our estimate of τ_K involves the following approximations: first, the FLC to obtain $V_{\mathrm{eff}}^{\mathrm{FLC}}$, second, the fact that we apply the Arrhenius formula to a non-Markovian system, third, all underlying simplifications made within the Kramers theory. The main idea behind this approach is illustrated in Fig. 6.7.

We note that one could employ the same strategy to get approximated escape times from the perturbation theory. We have indeed also tested this idea for the PT-OU, obtaining results very comparable to those from the FLC presented in the following.

6.5.3 Comparison with numerical results

We test the Kramers-FLC approximation by applying it to the example systems, comparing the predicted escape times to numerically obtained ones from BD simulation of the full model (1.42).

Periodic system Figure 6.8 displays exemplary results for τ_K in the (delayed) washboard potential. For $F_d \equiv 0$ (which corresponds to $\tau k = 0$), the numerical data and the Kramers-FLC estimate generally coincide. For $\tau k > 0$, the mean escape time τ_K generally increases by increasing k or τ. As one can see from Fig. 6.8, the FLC estimates are roughly in agreement with the actual escape rates. This is rather remarkable, due to the crudeness of applying the Kramers theory to this non-Markovian system. However, at large values of τ, they have a qualitatively different behavior: while the escape times resulting from the Kramers-FLC estimate saturate, the BD results continue to increase with τ. This difference becomes particularly prominent for large k, i.e., large delay force strengths.

We think that the saturation of the approximate results (and the discrepancy) can be explained as follows. As discussed in Secs. 6.3, 6.4, the conditional probability saturates at a finite value τ, which is related to the intrawell relaxation time τ_{ir}.

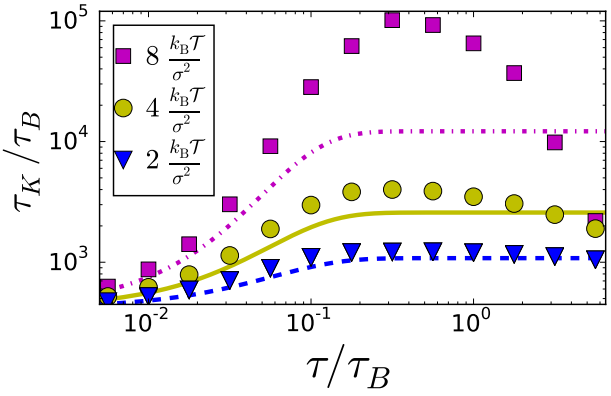

Figure 6.9: *Escape times τ_K vs. delay time τ in the doublewell potential with barrier height $V_0 = 8k_B\mathcal{T}$. Blue: $k = 2k_B\mathcal{T}/\sigma^2$, yellow: $k = 4k_B\mathcal{T}/\sigma^2$, magenta: $k = 8k_B\mathcal{T}/\sigma^2$. Lines plot the FLC-Kramers estimates, symbols show the Brownian Dynamics simulations result.*

Accordingly, also $V_{\text{eff}}^{\text{FLC}}$, and therewith τ_K from the Kramers-FLC approximation, saturate at this delay time. However, in the true non-Markovian system, the escape times do not only depend on $\rho_{1,\text{ss}}$, but also on higher n-time probabilities. Therefore, they do not necessarily saturate together with $\rho_{1,\text{ss}}$. Indeed, our numerical investigations reveal that τ_K does not saturate at all. Moreover, for larger τ, another, not yet taken into account timescale becomes increasingly important that is the jump duration time (see Sec. 1.6). If the delay time τ is in the range of the jump duration times, the temporal changes of the energy landscape and the dynamics of the system state variable X occur on similar timescales. The interplay between both motions then leads to new (interwell) dynamical behavior, which is not captured by the approximations. One example for such new dynamical behavior are the quasi-regular *delay-induced oscillations* of X between two valleys, which we discussed in Sec. 1.7 and reconsider below in Sec. 6.5.4. It hence becomes less justified to treat the delayed system as a quasi-Markovian one, whose stochastic (non-delayed) dynamics evolves in a static (effective) potential. Therefore, neither the FLC, PT, small τ, nor the application of Kramers theory are appropriate.

Bistable system Finally, we briefly look at the espace times in the doublewell potential. We find that the FLC estimated escape times again reach a saturation value (for the same reason as in the washboard potential explained above). However, here the BD results show a peak, and *decrease* for larger τ. For $k = 2k_B\mathcal{T}/\sigma^2$, the peak is at about $\tau \sim 0.2\tau_B$, for $k = 4k_B\mathcal{T}/\sigma^2$ it is at $\tau \sim 0.3\tau_B$, while the peak is around $\tau \sim 0.5\tau_B$ for $k = 8k_B\mathcal{T}/\sigma^2$. Interestingly, these delay times correspond quite accurately to the mean *jump duration times* within the respective potentials. We have determined the latter numerically, see Sec. 1.6 for more details. We find that in the regime of the peak delay-induced spontaneous oscillations occur.

6.5.4 Delay-induced oscillations

As discussed in Sec. 1.7, the delayed linear force F_d can induce more or less regular jumps between the wells in both static potentials, the washboard as well as the doublewell, if the delay time is large enough. As we have mentioned there, the delay-induced oscillations manifest in the autocorrelation function $C(\Delta t)$, yielding

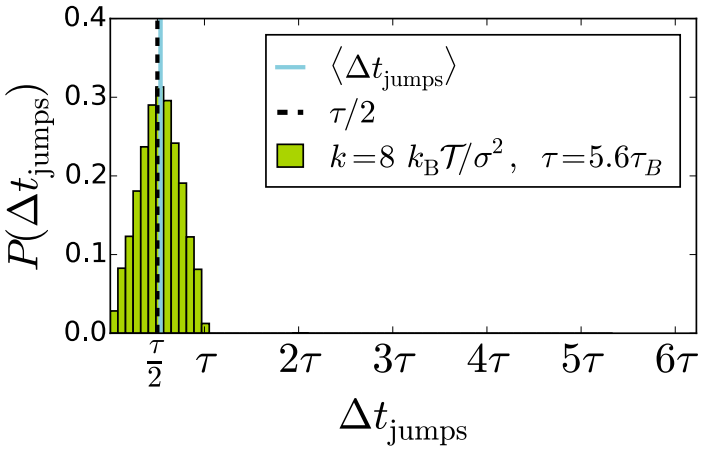

Figure 6.10: *Waiting time distribution in the delayed doublewell potential with* $k = 8\,k_{\text{B}}\mathcal{T}/\sigma^2$ *and* $\tau = 5.6\tau_{\text{B}}$*. The vertical purple line marks the mean waiting time* $\langle \Delta t_{\text{jumps}} \rangle$*, while the dashed black line marks* $\tau/2$*. The barrier height is set to* $V_0 = 8\,k_{\text{B}}\mathcal{T}$*. Due to the delay-induced jumps* $P(\Delta_{\text{jump}})$ *is strongly altered; compared with Fig. 6.6.*

hight peaks at $C(\tau)$, see Fig. 6.11 for an example.

Here we briefly consider their impact on the waiting time distributions. Let us first think about the expected effect on the mean value of the distributions. In the present case, the delay-induced oscillations have a mean period of about τ (in both static potentials) as $k > 0$. Because the mean escape times are usually much larger than τ, the occurrence of oscillations should reduce τ_{K} significantly. This is indeed visible in Fig. 6.9. Further, in addition to changing the mean value, these jumps manifest in a significantly altered functional form of the waiting time distributions (compare Figs. 6.10 and 6.6). In the presence of the oscillations, most waiting times lie within an interval $\Delta t_{\text{jump}} \in [0, \tau]$ around a single, yet broad peak at about $\tau/2$. The width of the peak illustrates the stochastic character of the oscillations, which start at random times, occasionally pause and randomly set in again.

With the frequent occurrence of delay-induced oscillations, the dynamics of the model deviates so strongly from a Markovian one that it is no surprise that the Kramers-FLC estimate completely breaks down.

One can understand the breakdown of all approaches so far from a very general perspective, when one thinks about the probability densities. In the parameter regime where jumps are very likely, the "memory of a jump", which is expressed in the non-Markovian ρ_2 (and higher PDFs), is what dominates the process. For example, it can trigger subsequent jumps, thus, oscillations. Since the higher PDFs correspond to the ones of Markovian systems in all approaches, the essence of the dynamics is not captured.

6.5.5 Side note: Normal diffusion despite non-Markovianity

As we have mentioned in Chapter 1, a common way to classify diffusion processes is via the mean square displacement $\langle [X(t) - X(0)]^2 \rangle$. In particular, when the latter grows linearly with time, the process is denoted normal diffusion, while nonlinear behavior is called anomalous diffusion. Let us briefly consider, in this regard, the washboard potential with delay force. Please note that the reason for not also considering the bistable potential is that the latter is confining.

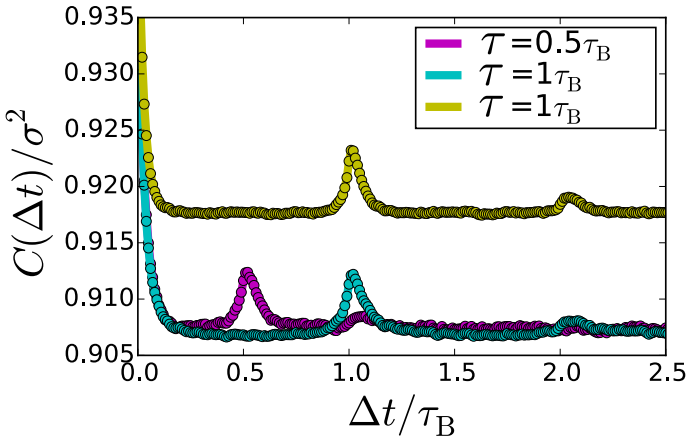

Figure 6.11: *Position-autocorrelation function in the doublewell with delay force. Cyan and magenta lines and symbols: $k = 8k_B\mathcal{T}/\sigma^2$, yellow: $k = 12k_B\mathcal{T}/\sigma^2$. The delay time is $\tau = \tau_B$ (yellow and cyan), or $\tau = 0.5\tau_B$ (magenta).*

In the presence of time delay, it is not immediately clear, which behavior to expect. However, the finite memory time (the memory kernel does not reach arbitrarily long into the past, but decays), one might expect normal diffusion at very large times. Indeed, Fig. 6.12 shows preliminary numerical results for a particle within the washboard potential with linear delay force, which indicate a linear growth of the MSD at large times. Note that this behavior seems to be irrespective of the occurring oscillations, as we have considered values of τ, where the delay-induced interwell oscillations are very prominent, e.g. $\tau = 4\tau_B$ shown by a green curve in Fig. 6.12, and values where they are not observed, e.g. $\tau = 2\tau_B$, blue curve. In the present case, the delay-induced oscillations are found to occur with high probability for about $\tau > 3\tau_B$.

Figure 6.12 further shows that the MSD is linear on short times (corresponding to the diffusion within the well before the particle "sees" the barriers), and saturates at intermediate times. This *plateau* roughly begins at the intrawell relaxation time, ends at the typical escape time out of the well (see Sec. 1.6 for a description of the timescales).

Please note that the kink at $t = \tau$ is a consequence of the initial condition, which was kept fixed along $[-\tau, 0]$, specifically $X(-\tau < t < 0) \equiv 0$. This yields an abrupt change in the diffusive behavior at $X(\tau)$, where the delayed positions suddenly start to be realization-dependent.

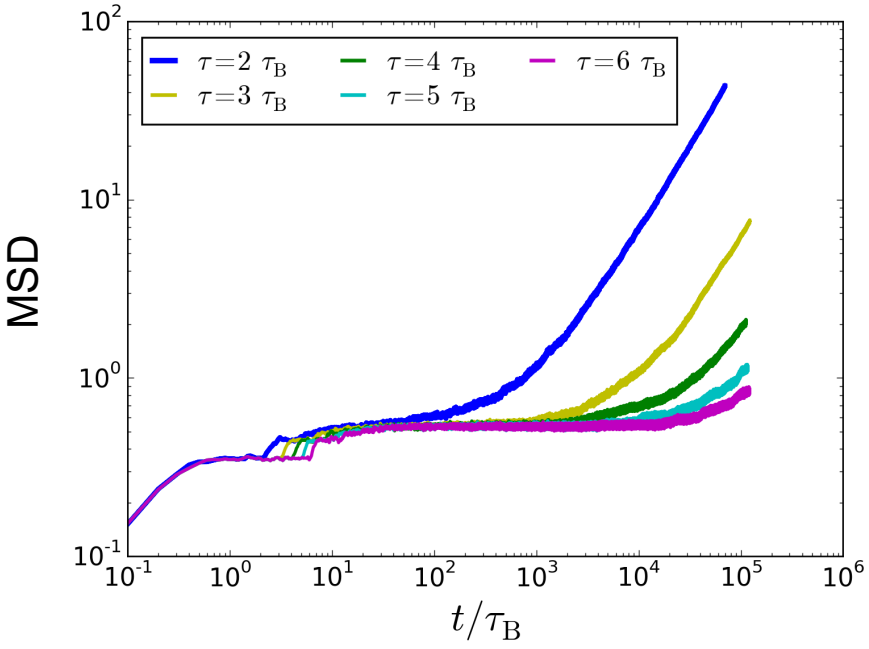

Figure 6.12: *Mean square displacement in the washboard potential with barrier height $V_0 = 1 k_B \mathcal{T}$, $k = 2 k_B \mathcal{T}/\sigma^2$, at different delay times. All considered cases seem to approach linear behavior (normal diffusion) for long times. Note that here $X(t') = 0$, $\forall\, t' \in [-\tau, 0]$. The MSD is given in units of σ^2.*

7 | Approximation for the two-time probability density

So far, we have discussed different approximations of the one-time probability density function, stemming from approaches based on the first member of the Fokker-Planck hierarchy. In particular, we have discussed the perturbation theory and the force-linearization closure (Chapter 6). While these approaches render quite accurate descriptions of the intrawell dynamics, they do not yield any information about the *two*-time PDF. The latter is, however, a very important quantity for the probabilistic description of the delayed dynamics. Many important features are only captured by two-time PDFs (or higher PDFs), which reflects the memory in this non-Markovian processes. As a matter of fact, only by studying them, non-Markovian steady-states can be qualitatively distinguished at all from thermal equilibria[1]. Moreover, we have seen that the FLC and PT estimates break down when escape events are likely, thus, when the interwell dynamics dominate. This is for example the case when the delay time is in the range of (or larger than) the jump duration times. Under such conditions, the interplay between the dynamics of the system state variable X and the time-dependent energy landscape causes oscillatory motion, whose description is clearly beyond the rather crude FLC-Kramers approximation described in Chapter 6. For this reason, an extension of the theory towards higher members of the FPE hierarchy would be worthwhile.

However, despite the clear benefits of a probabilistic treatment and the apparent limits of a study focusing on the first member of the FP hierarchy alone, no general approximation schemes for the two-time PDF or temporal position-correlation functions have been proposed in earlier literature (one exception being the two-state reduction employed in [2, 256], which we will discuss in Chapter 8).

In this spirit, we propose an approximation for the steady-state two-time PDF in the regime of large delay times, based on the new equation for the two-time PDF which we have derived in Chapter 5. We have introduced this approach in Ref. [3].

After reviewing the main idea, we will apply this approximation to our bistable example system with linear delay force. We will focus on the regime where the

[1] As an illustrative example, consider the following. Let $\rho_1^{\mathrm{nM}}(x)$ be the non-Markovian one-time PDF of an arbitrary nonequilibrium steady state. This PDF can in general not be distinguished from a PDF $\rho_1^{\mathrm{M}}(x)$ of a Markovian (equilibrium) process in a (fictive) potential defined by $U_{\mathrm{ficitve}}(x) \equiv k_{\mathrm{B}}\mathcal{T} \ln \rho_1^{\mathrm{nM}}(x)$, which would equally yield a distribution $\rho_1^{\mathrm{M}} = Z^{-1} e^{-U_{\mathrm{fictive}}/k_{\mathrm{B}}\mathcal{T}} = \rho_1^{\mathrm{nM}}$. Furthermore, the probability current vanishes, just like in an equilibrium systems.

system exhibits the delay-induced particle oscillations. As the delay-induced oscillations manifest in the two-time PDF, this system is ideal to test our approximation, and, at the same time, study non-Markovian effects on the level of the two-time PDF.

Main idea We reconsider the previously derived Eq. (5.44), recall

$$\begin{aligned}\partial_t\rho_2(x,t;x_\tau,t-\tau) =& -\partial_x\left[\gamma^{-1}F(x,x_\tau)\rho_2(x,t;x_\tau,t-\tau)\right] + D_0\partial_x^2\rho_2(x,t;x_\tau,t-\tau)\\ &+\partial_{x_\tau}\left[\langle\dot{X}(t-\tau)|x=X(t),x_\tau=X(t-\tau)\rangle\rho_2(x,t;x_\tau,t-\tau)\right].\end{aligned}$$

By neglecting statistical dependence between the (instantaneous and future) system state and the displacement, we obtain a closed equation that can be solved analytically in the steady state (where $\partial_t\rho=0$). As we will show in Sec. 7.1, it yields a surprisingly good approximation of the non-Markovian dynamics.

Technical details As a first step, we assume that the statistical dependence between the system state $X(t)$ and the delayed displacement $\mathrm{d}X(t-\tau)$ can be neglected, as well as the dependence between system state and instantaneous displacement. This amounts to the ansatz

$$\langle\dot{X}(t-\tau)|X(t),X(t-\tau)\rangle \approx \langle\dot{X}(t-\tau)\rangle. \tag{7.1}$$

With this simplification, Eq. (5.44), simplifies to

$$\begin{aligned}-\langle\dot{X}\rangle\,\partial_{x_\tau}\rho_2(x,t;x_\tau,t-\tau) =& -\partial_x\left[\gamma^{-1}F(x,x_\tau)\rho_2(x,t;x_\tau,t-\tau)\right]\\ &+D_0\,\partial_x^2\rho_2(x,t;x_\tau,t-\tau),\end{aligned} \tag{7.2}$$

which involves the *particle current* $\langle\dot{X}\rangle$.

In order to find an ansatz for the particle current, one can proceed by utilizing the framework [267] (developed for Markovian systems), which we have mentioned in Sec. 2.1.3. In particular, we can employ the relation $\langle\dot{X}\rangle = -\int_\Omega J(x,t)\mathrm{d}x$ with the probability current J. Recall that the latter is given by the FPE $\partial_t\rho_1 = -\partial_x J$. For the considered delayed systems, it reads

$$J(x,t) = \int_\Omega \gamma^{-1}F(x,x_\tau)\rho_2(x,x_\tau)\mathrm{d}x_\tau - D_0\partial_x\rho_1(x). \tag{7.3}$$

In combination, these simplifications give rise to the approximate, closed FPE (7.2) with $\langle\dot{X}\rangle = \iint_\Omega F(x,x_\tau)\rho_2(x,x_\tau)\mathrm{d}x_\tau\mathrm{d}x$.

Vanishing probability current When $\langle\dot{X}\rangle = 0$ in the steady-state, the approximate equation (7.2) automatically simplifies, without making further assumptions. This is indeed true for a common type of system, in particular, for any system that obeys a symmetry w.r.t. coordinate inversions $x\to -x$, like both of our example systems. In these cases (7.2) has the formal steady-state solution

$$\rho_2(x,t;x_\tau,t-\tau) = Z^{-1}h(x_\tau)e^{(\gamma D_0)^{-1}\int F(x,x_\tau)\,\mathrm{d}x}. \tag{7.4}$$

Z is again the normalization constant, and h is a function of x_τ which is not determined by (7.2) alone. However, h is determined by the steady-state constraint $\partial_t\rho_1 = 0$, which implies $\rho_1(x) = \rho_1(x_\tau) \Rightarrow \int_\Omega \rho_2(x,x_\tau)\mathrm{d}x = \int_\Omega \rho_2(x,x_\tau)\mathrm{d}x_\tau$. After specifying the delay force, which we do next, we can give a concrete example.

Linear delay force We again consider systems with *linear* delay force, and (inversion-symmetric) nonlinear, nondelayed part of the force $F_\mathrm{s}-kx$, i.e., $F(x,x_\tau) = F_\mathrm{s}(x)-kx+kx_\tau$, as in our two example systems (see Sec. 1). Then, we immediately have $h = e^{\int(\gamma D_0)^{-1}F_\mathrm{s}(x_\tau)\mathrm{d}x_\tau}$, yielding the approximate two-time PDF

$$\rho_2^{\mathrm{approx}}(x,t;x_\tau,t-\tau) = Z^{-1}\exp\frac{\int[F_\mathrm{s}(x)-kx]\,\mathrm{d}x\,+kxx_\tau\,+\int[F_\mathrm{s}(x_\tau)-kx_\tau]\,\mathrm{d}x_\tau}{\gamma D_0}, \tag{7.5}$$

which fulfills both, the approximate FPE (7.2) and the steady-state constraint.

A peculiar feature of (7.5) is that it does not explicitly involve τ. However, we have already seen before that the PDFs indeed saturate for large τ values above the intrawell relaxation time, (see Sec. 4.2.2 for linear, and in Sec. 6.4 and 6.3 for the nonlinear example systems).

The approximation discussed here is indeed designed to capture the regime of large τ values (above τ_ir). We suspect that we can understand the independence of (7.5) on τ in this sense. We further recall that the approximation (7.5) is *per se* expected to break down for small values of τ, because then the statistical dependence between $\dot{X}(t-\tau)$ and $X(t)$ are non-negligible.

7.1 Application to the bistable delayed system

We test the approximation (7.5) by applying it to our bistable example system, which we have already considered in Sec. 6.4 in the context of the approximations for the one-time PDF. In particular, we again consider the LE (1.53), recall

$$\gamma\dot{X}(t) = (4V_0/\sigma)[-(X(t)/\sigma)^3 + X(t)\sigma] - k[X(t)-X(t-\tau)] + \sqrt{2D_0}\,\xi(t),$$

which involves the doublewell potential $V = V_0[(x/\sigma)^4 - 2(x/\sigma)^2]$ from Eq. (1.52), with minima at $x \pm \sigma$ introduced in Sec. 1.5 and the linear delayed force (1.44), stemming from an optical tweezers potential.

From Eq. (7.5), readily follows the approximate steady-state two-time PDF

$$\rho_2(x,t;x_\tau,t-\tau) = Z^{-1}\exp\frac{-\frac{V_0}{\sigma^2}(\frac{x^4}{\sigma^2}-2x^2) - \frac{k}{2}x^2 - \frac{V_0}{\sigma^2}(\frac{x_\tau^4}{\sigma^2}-2x_\tau^2) - \frac{k}{2}x_\tau^2 + kxx_\tau}{\gamma D_0}. \tag{7.6}$$

As we have discussed already in Secs. 1.7 and 6.5.4, the bistable system (1.53) exhibits delay-induced oscillations between the two potential minima, when the delay time is larger than the typical jump duration time, and when $|k|$ has the same order of magnitude as V_0. Because these jumps manifest in the two-time PDF, as will be explained below, we consider a parameter regime, where they are pronounced (in particular, we fix $V_0 = 3.5\gamma D_0$ and $\tau = 10\tau_\mathrm{B}$) and only vary k, particularly focusing on whether the approximation can capture the delay-induced jumps. We test the results from a quantitative point of view by comparing with direct BD simulations[2] of Eq. (1.53).

[2]with $>$ 2000 realizations, $\Delta t = 10^{-4}$, and cut off $> 10^6$ time steps to reach steady-state conditions

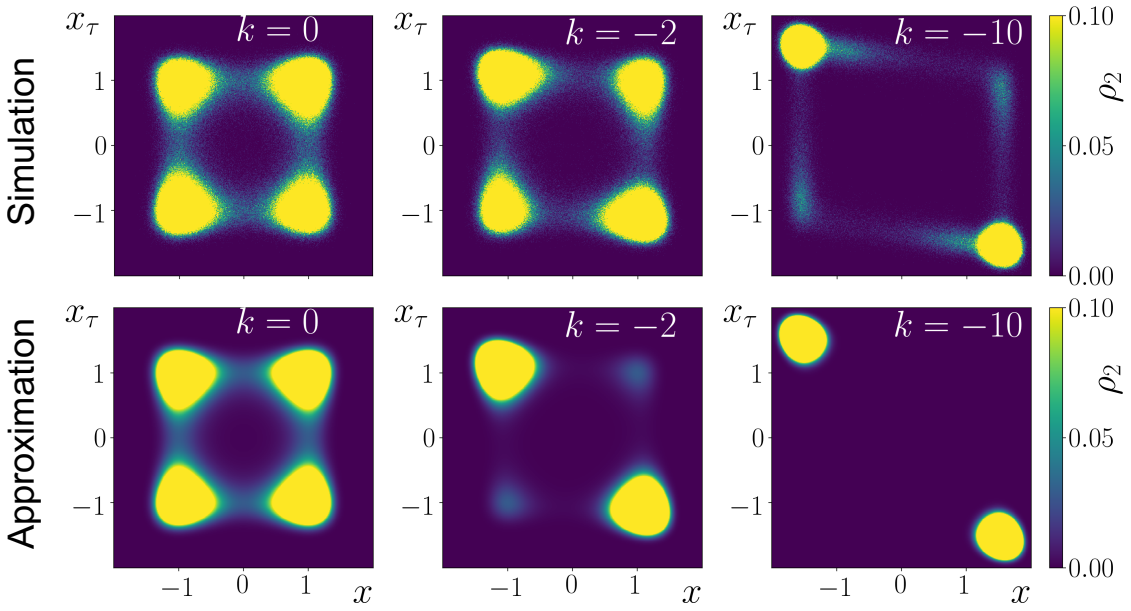

Figure 7.1: *Two-time PDF in the doublewell potential with a linear delay force of increasing (negative) magnitude k (positive feedback). Upper panels: numerical results from LE (1.53), lower panels: approximated ρ_2 from Eq. (7.6). Other parameters $V_0 = 3.5\gamma D_0$, $\tau = 10\tau_\mathrm{B}$. The feedback strength k is given in units of $(\gamma D_0/\sigma^2)$, while x is given in units of σ. This figure stems from [3].*

Markovian limit First, we consider the system in the absence of the delay force ($k = 0$), where the dynamics is Markovian and the system reaches thermal equilibrium. We expect the probability to find X to be the highest around the potential minima $x/\sigma = \pm 1$. As thermal jumps are very likely on the timescale of τ, the chance of finding the system in one well should be quasi independent of the question whether it was on the same or on the other side τ ago, meaning that the probabilities to find $X(t)$ and $X(t-\tau)$ are independent from each other. We hence expect the joint two-time PDF[3] $\rho_2(x, x_\tau)$ to have pronounced peaks of identical heights around $\{x, x_\tau\}/\sigma \approx \{\pm 1, \pm 1\}$ and $\{\pm 1, \mp 1\}$. This expectation is confirmed by numerical simulations, and is correctly described by our approximation (7.6), as plotted in Fig. 7.1 (left plots, upper and lower panel, respectively). We now consider the impact of $k \neq 0$ on the long time behavior, where the system is in a NESS.

Delay-induced oscillations For negative $k < 0$, i. e, when the linear feedback force is directed away from the delayed system state (positive feedback), the induced oscillations are known to have a mean period of roughly 2τ (see Sec. 1.7). This means that the jumps *reduce* the joint probability to find the system in the same well at times $t - \tau$ and t. We hence expect the delay-induced oscillations to lower the two-time PDF around $\{x, x_\tau\}/\sigma \approx \{\pm 1, \pm 1\}$, and increase it around $\{\pm 1, \mp 1\}$. This effect is indeed seen in both, the numerically obtained PDF and the approximation from (7.6), see Fig. 7.1. Interestingly, by changing the sign of k, the mean period

[3]Due to the inversion symmetry of the external potentials, we further always expect that $\rho_2(x, x_\tau)$ is symmetric w.r.t. $\{x, x_\tau\} \leftrightarrow -\{x, x_\tau\}$.

of the oscillations can be switched from 2τ to τ (see Sec. 1.7). Thus, we expect the delay-induced oscillations to have, in fact, the opposite effect for positive k, i.e., enhance the PDF at $\{\pm 1, \pm 1\}$, and reduce it around $\{\pm 1, \mp 1\}$. Again, this behavior is confirmed by simulations, and as well captured nicely by the proposed approximation, see Fig. 7.2, where we have considered negative feedback, $k > 0$.

The comparison between numerical results and the approximate ρ_2 from (7.6) in Figs. 7.1 and 7.2 includes a broad range of k values. Remarkably, we find that the proposed approximation yields very good results in all cases considered. In particular, it captures the main characteristics of the two-time PDF described above, for negative and for positive k. Even quantitatively, the agreement is very reasonable. We have further considered higher barrier heights ($V_0 = 5\,\gamma D_0$), smaller delays ($\tau = 5\,\tau_{\mathrm{B}}$), as well as higher feedback strength (up to $|k| = 20\,\gamma D_0/\sigma^2$) and found as well very good agreement (not shown here). On the contrary, the approximation breaks down if the delay time is very short, see for example Fig. 7.3, where $\tau = 0.1\,\tau_{\mathrm{B}}$ (left panel). We suspect that in this regime, the statistical dependence between $\dot{X}(t-\tau)$ and $X(t)$ is not negligible (hence, the assumption (7.1) is not justified).

Further interesting features of the two-time PDF We note that besides the features discussed above, the numerically obtained two-time PDFs in Figs. 7.1 and 7.2 show very interesting further details. An example is the broken symmetry $\rho_2(y,t;y',t-\tau) \neq \rho_2(y',t;y,t-\tau)$, which is not captured by the approximation (7.6). We have not further studied these features, but it indeed would be interesting to see how these details are connected to dynamical and thermodynamic properties of the non-equilibrium steady-state. (Equilibrium states are always symmetric.) This might be an interesting perspective for future research.

7.1.1 Comparison with approaches for one-time probability density

The accuracy of the approximate ρ_2 described above is indeed remarkable, given that the underlying assumption appears to be rather crude. Further, the approximation result has a simple closed form (7.6). This indicates that an approximation scheme on the level of the second member of the FP hierarchy is somewhat superior as compared to an approach on the level of the first member. To further support this argument, we compare the prediction of approximation (7.6) to the approaches for the one-time PDF which we have discussed in the previous Chapter (see Chapter 6).

To this end, we marginalize ρ_2 from Eq. (7.6), i.e., $\rho_1(x) = \int_\Omega \rho_2(x, x_\tau)\mathrm{d}x_\tau$ and compare the resulting approximate one-time PDF with the corresponding PDF from the small delay expansion, and the force-linearization closure (FLC), given in Sec. 6.4. As we have discussed in detail in Secs. 6.4, and 6.5, the FLC and the small τ, are both not justified in the regime where the delay-induced oscillations dominate the dynamics. We nevertheless use them to compare with, due to the lack of other approximation schemes for this nonlinear system.

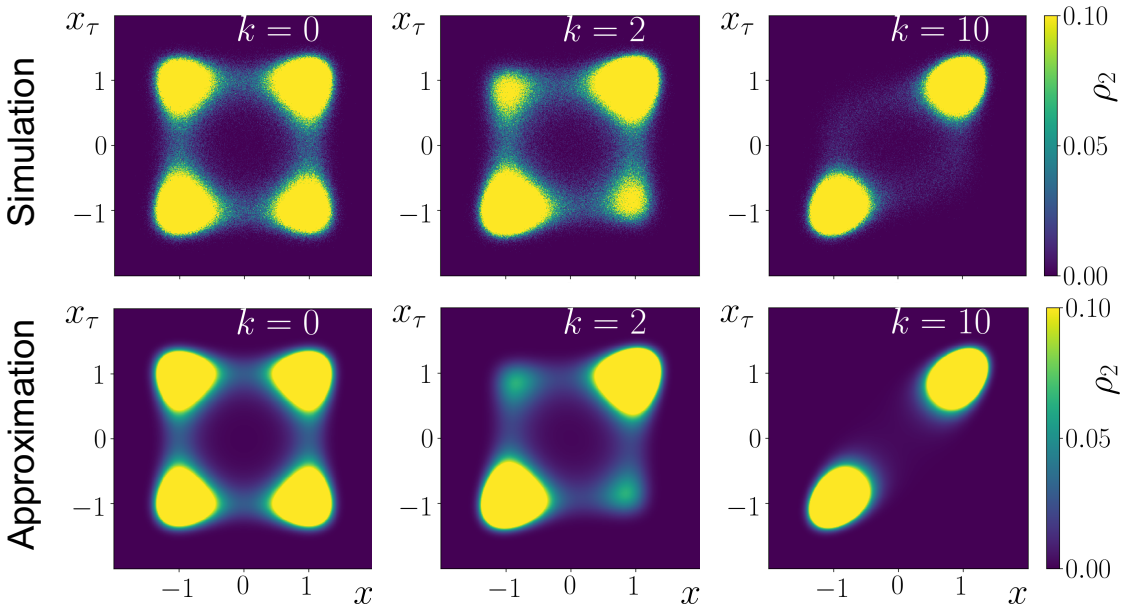

Figure 7.2: *Two-time PDF in the doublewell potential with linear delay force of increasing (positive) magnitude k (negative feedback). Upper panels: numerical results from LE (1.53), lower panels: approximated ρ_2 from Eq. (7.6). Other parameters $V_0 = 3.5\gamma D_0$, $\tau = 10\tau_\mathrm{B}$. The feedback strength k is given in units of $(\gamma D_0/\sigma^2)$. This figure stems from [3].*

Consistency check To make sure the consistency of the results, we briefly consider the Markovian limit $k \to 0$, where the system is expected to reach thermal equilibrium. Indeed, all approximations are found to correctly reproduce the Boltzmann-distribution $\rho_1 \propto e^{-V_s/k_\mathrm{B}\mathcal{T}}$, as displayed in Fig. 7.1.1.

Large τ regime Let us return to the case with feedback. In the regime of large delay times, the novel approximation yields the best results for ρ_1, for all considered values of the feedback strength k, as confirmed by Figure 7.5. In this regime, delay-induced oscillations dominate, such that the underlying assumption of the FLC is violated. The small delay expansion, on the other hand, is *per se* not justified, due to the large delay time $\tau = 10\,\tau_\mathrm{B}$.

Short delay times By construction, the approximate two-time PDF (7.6) is not reliable, if the delay time is very short, as the underlying assumption is violated. For the moderate $|k|/V_0$ and small τ values considered in Fig. 7.6 (a), the corresponding one-time PDF is found to still give a reasonable approximation. In this regime, also the other approaches from Chapter 6, e.g, the FLC, and the small delay expansion are justified (as well as the perturbation theory, which is not considered here. The two-time PDF approximation performs worse than to the FLC, but still better then the small delay expansion. Interestingly, for the large, negative k/V_0, the situation is reversed and only the small delay expansion yields a somewhat reasonable approximation, see Fig. 7.6 (b).

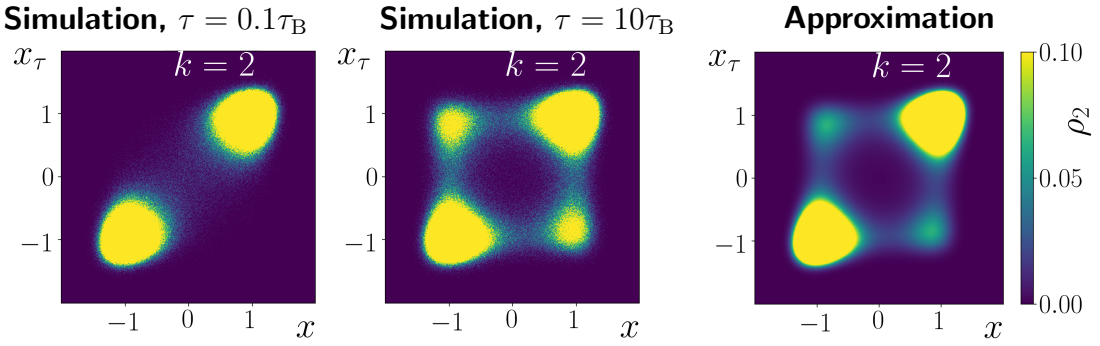

Figure 7.3: *Two-time PDF in the doublewell potential with linear delay force $kX(t-\tau)$, with positive feedback strength $k = 2(\gamma D_0/\sigma^2)$, and different time delays τ. Left: Simulation results with $\tau = 0.1\tau_B$, Center: Simulation results with $\tau = 10\tau_B$, Right: Approximation. Other parameters $V_0 = 3.5\gamma D_0$. The approximation matches better to the results at large values of τ, as expected.*

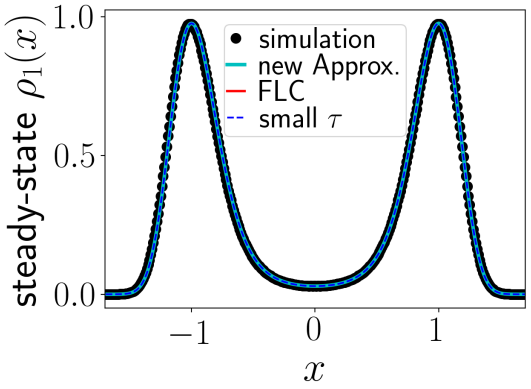

Figure 7.4: *Consistency check: The small delay expansion (small τ), the force-linearization closure (FLC), and the new approximation (7.6), all reproduce the correct equilibrium one-time PDF ρ_1 in the Markovian limit $k \to 0$. The potential barrier height is $V_0 = 3.5\,\gamma D_0$, x is in units of σ, ρ_1 is in units of $1/\sigma$.*

7.2 Concluding remarks

By neglecting statistical dependencies between the system state and the displacement, we have obtained a closed equation for the two-time PDF that can be solved analytically. To test the approximation, we have applied it to our bistable example system involving a doublewell potential and a linear delay force, which we have also used in Chapter 6 to test the approaches with base on the first FP hierarchy member. Here we have focused on the regime where the bistable model exhibits delay-induced particle oscillations between the two minima.

The described results demonstrate that the proposed scheme yields a good approximation of the actual[4] PDFs. This is remarkable given the apparent crudeness of the underlying assumptions. In particular, the approximation captures the main characteristics of the two-time PDF, in both situations, where the delay force is directed towards, and where it points away from the delayed position. To the best of our knowledge, this is the first approximation for the two-time probability density of continuous stochastic systems with time delay.

Projecting the obtained two-time PDF on one time argument, we can further

[4]which are obtained with numerical simulations of the full model

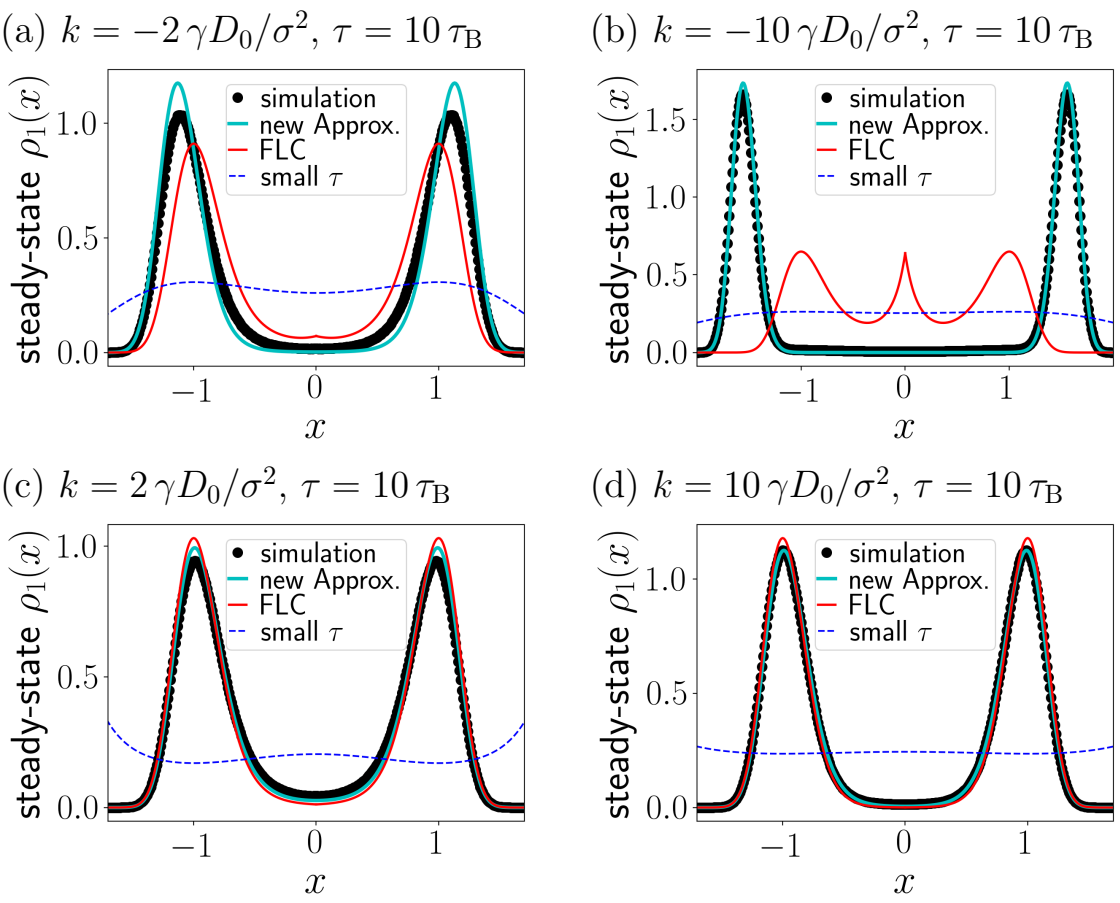

Figure 7.5: *Comparison of different approximations for the one-time PDF in the doublewell potential with linear delay force, at $V_0 = 3.5\,\gamma D_0$, $\tau = 10\,\tau_B$, and at various values of k: (a,b) positive feedback with (a) $k = -2\gamma D_0/\sigma^2$, (b) $k = -10\gamma D_0/\sigma^2$. (c,d) negative feedback, with (c) $k = 2\gamma D_0/\sigma^2$, (d) $k = 10\gamma D_0/\sigma^2$. blue dashed lines: small delay expansion (which is only justified for very short τ), cyan lines: the new approximation $\int_\Omega \rho_2(x, x_\tau)\mathrm{d}x_\tau$ with ρ_2 from Eq. (7.6), red lines: Force-linearization closure (FLC) discussed in Chapter 6. black disk: Brownian dynamics simulations. These figures stem from [3].*

compare the results to the approximations for the one-time PDF from Chapter 6. We find that the new approximation outperforms the other approaches in the regime of large delay times, demonstrating that a framework on the level of the second FP member, is somewhat superior compared to approaches on the first member.

In the following, we turn to the thermodynamic description of systems with time delay.

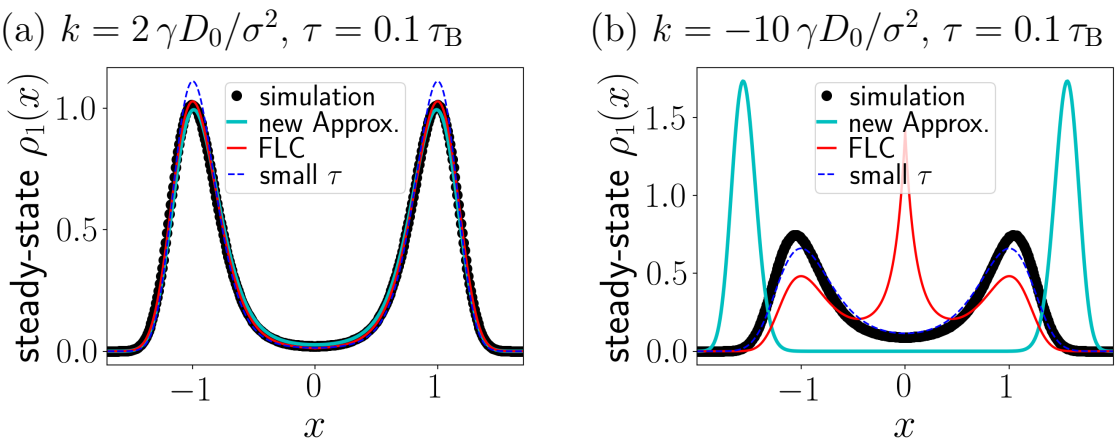

Figure 7.6: *Comparison of different approximations for the one-time PDF in the doublewell potential with linear delay force at $V_0 = 3.5\,\gamma D_0$, for two exemplary values of k: (a) negative feedback $k = 2\gamma D_0/\sigma^2$, (b) positive feedback $k = -10\gamma D_0/\sigma^2$. In contrast to Fig. 7.5 where $\tau = 10\,\tau_\mathrm{B}$, here a shorter delay time of $\tau = 0.1\,\tau_\mathrm{B}$ is considered. These figures stem from [3].*

Part III

Thermodynamic notions for systems with time delay

"Wenn sich die hier zu behandelnde [Brownsche] Bewegung samt den für sie zu erwartenden Gesetzmäßigkeiten wirklich beobachten läßt, so ist die klassische Thermodynamik schon für mikroskopisch unterscheidbare Räume nicht mehr als genau gültig anzusehen"

– Albert Einstein, 1905 [178]

After having discussed the probabilistic description in Part II, we now turn to another cornerstone of statistical physics: thermodynamics. As we deal with non-Markovian, nonequilibrium systems, a promising theoretical framework is stochastic thermodynamics. In Chapter 3, we have introduced the key concepts of this theory, and formulated the main research questions addressed in this thesis.
In the following, we will first apply stochastic thermodynamics to investigate the mean heat rate and medium entropy production for nonlinear systems with time delay. In particular, we will clarify whether energy extraction from the single heat bath is possible, address the heat flow in the limit of vanishing delay times, and discuss the consequence of the overdamped limit on thermodynamic properties. We aim to go as far as possible with analytical arguments, but also present numerical results.
Furthermore, we will discuss the notion of total entropy production, and calculate information flows in systems subject to time-delayed feedback. To this end, we will step away from the case of discrete delay, and consider systems with distributed delay and colored noise. To obtain an entropic description, we will employ a strategy related to the Markovian embedding introduced in Part II.

8 | The heat flow induced by a discrete delay

Finding thermodynamic notions for non-Markovian systems is a major problem. As a first step in this direction, we consider the *heat rate* $\dot{Q} = \langle \delta q/\mathrm{d}t \rangle_{\mathrm{ss}}$ of nonlinear systems with discrete delay. This quantity can be calculated via well-established concepts from stochastic thermodynamics, in particular, from the framework of stochastic energetics [124]. The heat flow and related medium entropy production are key thermodynamic quantities and nontrivial part of the total entropy production, which already provide important physical insight. Based on this consideration, we can already address several fundamental questions. First, we aim to clarify: *Can the delayed feedback induce a net energy flow from the bath to the colloid?* Here we recall the connection to the "Maxwell demon"-type of devices mentioned in Chapter 3. Next we want to find out: *Is the overdamped limit problematic in regard to the thermodynamic description?* Indeed, in the dynamical and probabilistic treatment, this limit was not problematic at all (see Chapter 6). However, from a thermodynamic perspective the situation might be a different. Specifically, the overdamped limit, which amounts to a neglect of the *inertia term* in the LE[1], implies that the kinetic energy is not described. In turn, the kinetic energy would yield additional contribution in the energy balance (see Chapter 3), thus, possibly affecting the calculated heat flow. Last, we also ask ourselves whether *thermodynamic quantities reflect delay-induced dynamical behavior?* As an example for the latter, we will again consider the delay-induced oscillations, which we have mentioned in Chapters 6, and 7. Indeed, a previous, numerical study [266] has indicated that the medium entropy production has a maximum around coherence resonance. However, in this study only a single parameter regime was considered, no analytical arguments are used to support the finding, and no physical explanation is provided. Therefore we believe that it is worthwhile to revisit this observation and investigate the heat and medium entropy production around coherence resonance more thoroughly.

We have published in Ref. [2] the main results presented in the following.

Previous results for linear systems For linear delay systems, it has already been shown that in the long-time limit a nonequilibrium steady state with finite heat production [99] is approached (in the absence of time-dependent forces), as the

[1]We recall that lists of all abbreviations and symbols are given on pages VIII and IX, respectively.

delay pushes the system out of equilibrium. However, this previous result fails to correctly describe the limit of vanishing delay time, and the question whether energy extraction from the single heat bath is possible, was not answered. Another issue is that the approach in [99] is limited to linear systems only.

8.1 Main idea

For nonlinear cases, the linear response function method employed in [99] cannot be applied. In the following, we present an alternative strategy to get analytical results. Focusing on polynomial energy landscapes brings us into the position to find some analytical statements for nonlinear, time-delayed systems. In particular, by using both, the FPE and LE, we can derive several identities linking different noise-position correlations as well as position-autocorrelations (with time difference 0 or τ) to positional moments. By means of these identities, we can then derive an equation for the mean heat rate that solely depends on positional moments. In combination with the approaches presented in Part II, this equation can be utilized to obtain closed approximate expressions for the heat, which become exact in the Markovian limits.

8.1.1 Polynomial energy landscapes

In this Chapter, we assume that the conservative forces in the delay LE (1.42) can be expressed as a *polynomial*, as it is the case for our bistable example, and further holds for a wide class of nonlinear potentials (even non-polynomial ones due to Taylor expansions). This assumption simplifies our analytical treatment, as will become clear below. The feedback control force is again chosen to be linear (as in the case of optical tweezers feedback, see Sec. 1.4.1). Hence, we here consider the overdamped LE

$$\gamma \mathrm{d}X(t) = -\sum_{i=1}^{m} a_i X(t)^i \mathrm{d}t - bX(t-\tau)\mathrm{d}t + \gamma\sqrt{2D_0}\,\xi(t)\,\mathrm{d}t, \tag{8.1}$$

with $a_i, b \in \mathbb{R}$. The non-delayed part of the deterministic force $-\sum_{i=1}^{m} a_i X(t)^i$ might stem from a static potential $V_\mathrm{s}(x) = \sum_{i=1}^{m}(a_i/i+1)x^{i+1}$, and is hence a conservative force[2].

Following the stochastic energetics framework by Sekimoto described in Chapter 3 [specifically, in Sec. 3.2, see Eqs. (3.3 - 3.1)], the fluctuating work δw done by the *nonconservative* delayed feedback force to the particle, the change of its internal energy $\mathrm{d}u$, which is given by the potential energy change, and the heat δq flowing to the reservoir during the infinitesimal time $\mathrm{d}t$ are given by [99]

$$\delta w(t) = -bX(t-\tau) \circ \mathrm{d}X(t), \tag{8.2}$$

$$\mathrm{d}u(t) = V_\mathrm{s}'\left[X(t)\right] \circ \mathrm{d}X(t), \tag{8.3}$$

$$\delta q(t) = \gamma\big[\mathrm{d}X(t)/\mathrm{d}t - \sqrt{2D_0}\,\xi(t)\big] \circ \mathrm{d}X(t) = \delta w(t) - \mathrm{d}u(t). \tag{8.4}$$

[2]Note that we here include the non-delayed part $-kX(t)$ of the "optical tweezers" force $-k[X(t) - X(t-\tau)]$ into the a_1-term of the polynomial sum. In this case, $b = -k$.

Plugging the LE (8.1) into Eqs. (8.2, 8.3) results in the following expressions for the NESS ensemble averages[3]

$$\dot{W} \equiv \left\langle \frac{\delta w}{\mathrm{d}t} \right\rangle_{\mathrm{ss}} = \frac{b}{\gamma}\Big\{ \sum_{i=1}^{m} a_i C_i(\tau) + b \left\langle X^2 \right\rangle_{\mathrm{ss}} - \gamma\sqrt{2D_0}\, \langle X(t-\tau)\xi(t)\rangle_{\mathrm{ss}} \Big\}, \tag{8.5}$$

$$\begin{aligned} \dot{U} \equiv \left\langle \frac{\mathrm{d}u}{\mathrm{d}t} \right\rangle_{\mathrm{ss}} = \frac{1}{\gamma} \sum_{i=1}^{m} \Big\{ & a_i \gamma \sqrt{2D_0} \left\langle X^i \xi \right\rangle_{\mathrm{ss}} - a_i b\, C_i(\tau) - a_i^2 \left\langle X^{2i} \right\rangle_{\mathrm{ss}} \\ & - \sum_{j>i}^{m} 2 a_i a_j \left\langle X^{i+j} \right\rangle_{\mathrm{ss}} \Big\}, \end{aligned} \tag{8.6}$$

$$\dot{Q} \equiv \left\langle \frac{\delta q}{\mathrm{d}t} \right\rangle_{\mathrm{ss}} = \dot{W} - \dot{U}. \tag{8.7}$$

Because of the polynomial from of the static potential, the mean values of the energy flows linearly depend on various positional and position–noise cross-correlations. Further, due to the time delay, the thermodynamic quantities depend on the temporal position-autocorrelation functions at time difference τ,

$$C_i(\tau) = \left\langle X(t)^i X(t-\tau) \right\rangle_{\mathrm{ss}}, \tag{8.8}$$

mirroring the non-Markovian nature of (8.1).

8.2 Mean heat rate & Medium entropy production

In the following, we derive exact expressions for (8.6, 8.5) which only involve positional moments at one time. To this end, we employ the FPE and the LE. Specifically, by projecting the FPE onto positional moments and subsequently inserting the LE, we obtain relations between different correlation functions, as was previously suggested in [84]. In a second step, these expressions can then be evaluated for example on the basis of known approximation schemes, which we have discussed in Part II of this thesis.

Derivation of relations between correlation functions First we derive useful relations between temporal position-autocorrelation functions. To this end, we project the FPE (2.15) onto its moments by multiplying the equation with x^{n+1}, with an arbitrary nonnegative integer n, and integrating over the spatial domain $x \in (-\infty, \infty)$. After several partial integrations, (where the boundary terms vanish due to the natural boundary conditions), yield

$$\gamma \frac{\mathrm{d}}{\mathrm{d}t} \frac{\langle X(t)^{n+1} \rangle}{(n+1)} = \gamma D_0\, n \left\langle X(t)^{n-1} \right\rangle - \sum_{i=1}^{m} a_i \left\langle X(t)^{n+i} \right\rangle - b \left\langle X(t)^n X(t-\tau) \right\rangle. \tag{8.9}$$

[3] Recall that capital letters denote ensemble averages.

In a NESS, the left side vanishes and (8.9) reduces to a relation between different moments and the temporal position-autocorrelation function [84]

$$b\,C_n(\tau) = \gamma D_0\, n \left\langle X^{n-1} \right\rangle_{\rm ss} - \sum_{i=1}^{m} a_i \left\langle X^{n+i} \right\rangle_{\rm ss}, \tag{8.10}$$

valid $\forall n \geq 1$. We will use this relation to substitute all the $C_i(\tau)$-terms in the mean work rate and internal energy from (8.6, 8.5).

Further, by plugging in the LE (8.1), one can deduce directly

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\frac{\langle X(t)^{n+1}\rangle}{(n+1)} = \left\langle X(t)^n \frac{\mathrm{d}X(t)}{\mathrm{d}t} \right\rangle &\overset{\text{LE}}{=} -\sum_{i=1}^{m} \frac{a_i}{\gamma} \left\langle X(t)^{i+n} \right\rangle + \sqrt{2D_0}\, \langle X(t)^n \xi(t) \rangle \\ &\quad - \frac{b}{\gamma} \langle X(t)^n X(t-\tau) \rangle . \end{aligned} \tag{8.11}$$

A comparison of Eqs. (8.11) and (8.9) readily provides for all $n \geq 1$ the surprisingly simple and generic relation

$$\langle X(t)^n \xi(t) \rangle = n \sqrt{D_0/2}\, \left\langle X(t)^{n-1} \right\rangle, \tag{8.12}$$

which, in fact, generally holds for any (nonlinear) force. This can be shown analogously, see Ref. [84]. Using Eq. (8.12), the instantaneous noise–position cross-correlations in (8.6) can be replaced by positional moments.

Finally, to evaluate $\langle X(t-\tau)\xi(t) \rangle$ at $\tau \neq 0$, we combine Eq. (8.12) with the causality argument (see Secs. 1.3.5 and 3.4.1). Because no physical quantity shall be influenced by future noise, statistical independence follows, and $\langle X(t)\xi(t') \rangle = \langle X(t) \rangle \langle \xi(t') \rangle \equiv 0$ must hold $\forall t' > t$. Hence

$$\langle X(t-\tau)\xi(t) \rangle = \sqrt{\frac{D_0}{2}}\delta_\tau = \begin{cases} \sqrt{D_0/2}, & \tau = 0 \\ 0, & \tau > 0 \end{cases}. \tag{8.13}$$

With these relations, we can now consider the mean rates of the energy flows. Substituting Eqs. (8.10) and (8.12) at $n = 1, 2, .., m$, into Eq. (8.6) yields

$$\dot{U} = 0. \tag{8.14}$$

This is expected, since the net internal energy should be a conserved quantity in a NESS [125]. Thus $\dot{W} \equiv \dot{Q}$. Substituting (8.10) at $n = 1, 2, .., m$ into Eq. (8.5) and using (8.13) further yields

$$\begin{aligned} \dot{Q} = \dot{W} &= \sum_{i=1}^{m} \Big\{ i a_i D_0 \left\langle X^{i-1} \right\rangle_{\rm ss} - \sum_{j=1}^{m} \frac{a_i a_j}{\gamma} \left\langle X^{i+j} \right\rangle_{\rm ss} \Big\} + \frac{b^2}{\gamma} \left\langle X^2 \right\rangle_{\rm ss} - b D_0 \delta_\tau \\ &= \dot{S}_{\rm m}(\gamma D_0 / k_{\rm B}). \end{aligned} \tag{8.15}$$

Please note that this simplification step is possible due to the polynomial form of the deterministic forces.

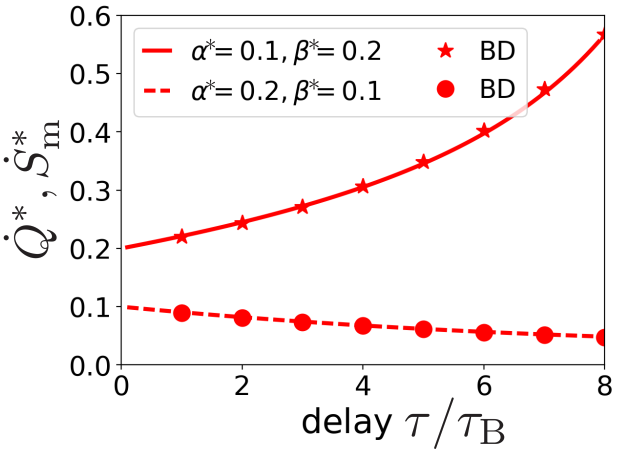

Figure 8.1: *Mean heat rate* $\dot{Q}^* = \dot{Q}\,(\tau_B/(k_B\mathcal{T}))$ *and medium entropy production rate* $\dot{S}_m^* = \dot{S}_m\,(\tau_B/k_B)$ *plotted against the delay time* τ/τ_B*, in the linear system with time delay at different* $\alpha^* = \alpha\,\tau_B/\gamma$ *and* $\beta^* = \beta\,\tau_B/\gamma$*. Solid line and stars:* $\alpha^* = 0.1$*,* $\beta^* = 0.2$*, dashed line and disks:* $\alpha^* = 0.2$*,* $\beta^* = 0.1$ *. The lines show the exact solution (8.17), the symbols stem from simulations (BD). The parameters are identical to FIG. 3 of Ref. [99], demonstrating the consistency of the results at* $\tau > 0$*.*

Equation (8.15) is an exact expression only involving positional moments, i.e., *one*-time ensemble averages over X^n. $\dot{Q}$ and $\dot{S}_m$ can therefore be computed directly from the steady-state one-time PDF, and hence, on the basis of the various approximations discussed in Part II.

The δ_τ-term suggests discontinuous behavior of $\dot{Q}$ at $\tau \to 0$. But in order to study this limit properly, one also has to investigate the behavior of the PDF to clarify whether the moments behave continuously. Obviously, this can be readily done for linear systems, where the exact solutions are known. In fact, we can as well find exact statements for nonlinear systems in the limits. We will discuss both in the following, starting with the linear case. As in the limit $\tau \to 0$ Markovianity is recovered, one expects that the system equilibrates (because of the absence of external driving), and therefore $\dot{Q} = \dot{S}_m = 0$. The same is expected when $b \to 0$, which we also will discuss below.

8.2.1 Linear systems

Let us first consider the purely linear systems with LE (1.49), where the force has the form

$$F = -\alpha x - \beta x_\tau. \tag{8.16}$$

Thus, comparing to notation here and there, $a_1 = \alpha$, $b = \beta$. Recall that for this case, the exact steady-state PDFs are known, as we have discussed in Sec. 4.2. Therefore, we can immediately write down a closed expression for the mean heat rate from Eq. (8.15), which reads

$$\frac{\dot{Q}}{D_0} = \frac{\gamma \dot{S}_m}{k_B} = \alpha - \beta\delta_\tau + (\alpha^2 - \beta^2)\frac{1 + \beta \sinh\left(\sqrt{\alpha^2-\beta^2}\,\tau/\gamma\right)\Big/\sqrt{|\alpha^2-\beta^2|}}{\alpha + \beta\cosh\left(\sqrt{\alpha^2-\beta^2}\,\tau/\gamma\right)}. \tag{8.17}$$

Equation (8.17) is consistent with Ref. [99], where only the linear case is considered (and a different approach is used via linear response functions), apart

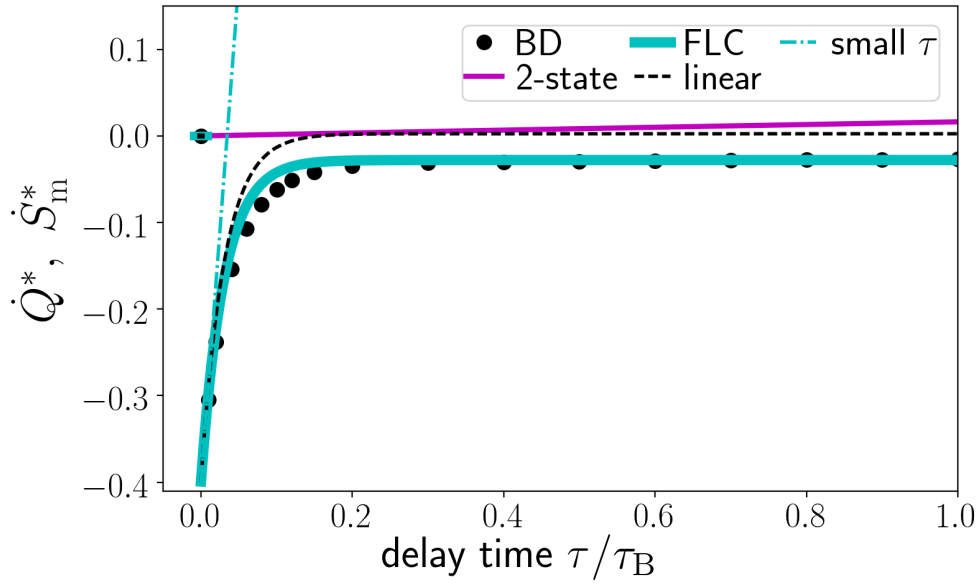

Figure 8.2: *Mean heat rate $\dot{Q}^* = \dot{Q}\,(\tau_\mathrm{B}/(k_\mathrm{B}\mathcal{T}))$ and medium entropy production rate $\dot{S}^*_\mathrm{m} = \dot{S}_\mathrm{m}\,\sigma^2/(k_\mathrm{B}D_0)$ vs. delay time (scaled with the Brownian time $\tau_\mathrm{B} = \sigma^2/D_0$), in the delayed bistable system at $V_0 = 4k_\mathrm{B}\mathcal{T}$, $k = 0.4k_\mathrm{B}\mathcal{T}$. The results stem from simulations (BD), and from several approximations: force-linearization closure (FLC), small delay small expansion (small τ), and the 2-state reduction. The dashed line shows $\dot{Q}^*$ and $\dot{S}^*_\mathrm{m}$ of the corresponding linearized system at $\alpha = -(8V_0 + k)/\sigma$ and $\beta = -k/\sigma$ from Eq. (8.17). The same results are again shown in Fig. 8.3 below, in a logarithmic plot.*

from the *additional* δ_τ-term. Consequentially, only Eq. (8.17) correctly predicts $\dot{S}_\mathrm{m}(\tau \to 0) = 0$. We will discuss this point in detail below.

Let us now take a closer look to the mean heat rate $\dot{Q}$ and the medium entropy production $\dot{S}_\mathrm{m}$ (which are proportional to each other) from Eq. (8.17). Figure 8.1 depicts the results for two exemplary parameter settings. One immediately sees that the heat rate is a nonzero quantity for $|\beta|\tau > 0$, proving the true non-equilibrium nature of this steady state. Here, $\dot{Q} > 0$ implying heating of the surrounding fluid. However, according to (8.17), the heat rate can also be negative for certain β, τ values. An example is given in Fig. 8.3 (dashed line[4]). At this point, we can answer one of our initially posed research questions about the thermodynamic behavior induced by time-delayed feedback (stated in the Introduction and in Sec. 3.6.1), that is, whether a delay force can be used to extract energy from the surrounding heat bath (like a Szilárd engine). Indeed, a linear delay force is capable of performing this task, already in this fully linear system.

Moreover, according to Eq. (8.17), $\dot{Q}$ continuously approaches zero as $\beta \to 0$, and nullifies at $\beta = 0$, i.e., when the steady state is the thermal equilibrium. On the contrary, the limit $\tau \to 0$ has a different characteristic. Remarkably, the mean heat rate has an apparent offset at $\tau = 0$, where it *discontinuously* drops to zero. This can be seen in Fig. 8.1, as well as in the following Figs. 8.3 and 8.4 (black dashed

[4]The other results shown in this Figure will be discussed below.

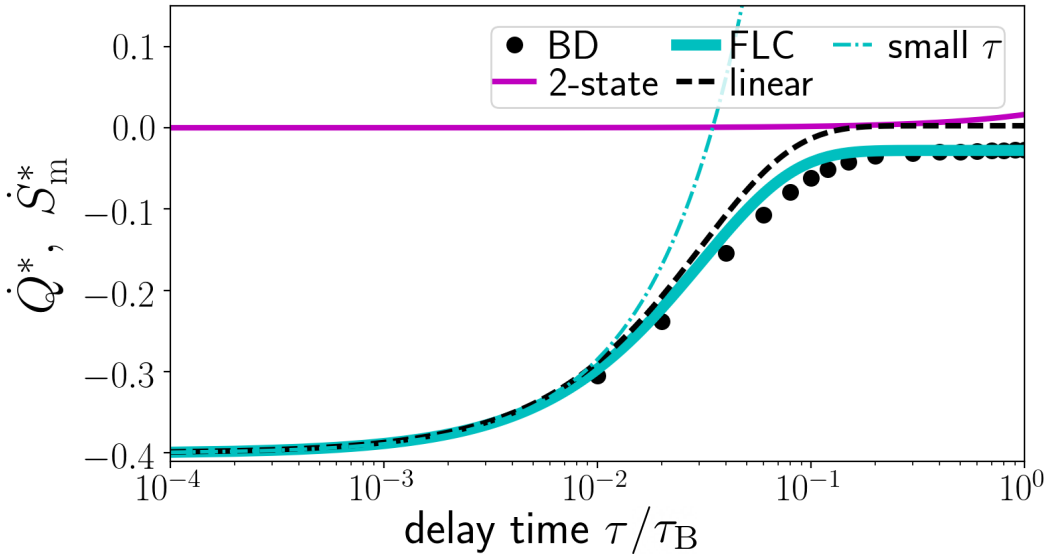

Figure 8.3: *Mean heat rate* $\dot{Q}^* = \dot{Q}\,(\tau_{\mathrm{B}}/(k_{\mathrm{B}}\mathcal{T}))$ *and medium entropy production rate* $\dot{S}^*_{\mathrm{m}} = \dot{S}_{\mathrm{m}}\,\sigma^2/(k_{\mathrm{B}}D_0)$ *vs. delay time as in Fig. 8.2. Note the logarithmic x-axis here, which makes it easier to see the behavior at very small* τ *values.*

lines), where $\dot{Q}$ is plotted as a function of τ (for two different parameter settings). In fact, for some parameters, $|\dot{Q}|$ even *grows*, when starting from large τ and decreasing τ towards zero, (until a saturation value is reached giving the apparent offset).

The discontinuity and the growth are odd features of the heat flow. We will investigate them in detail in the following. But before, let us clarify the limit behavior in *nonlinear* cases, because the situation turns out to be analogous there.

8.2.2 Markovian limits in nonlinear systems

Now we turn to the nonlinear cases, where we have no general analytic solutions. By utilizing the results presented in Part II, we can nevertheless obtain some strict statements.

Main idea For nonlinear systems, no exact solutions for the one-time PDF in the presence of delay are known. However, we can nevertheless analytically address the question how the PDFs behave in the Markovian limits, since we have approximative PDFs which become *exact*. This will give us exact results from Eq. (8.15) in the limits. In particular, the small delay expansion (which we have introduced in Sec. 6.2.1), is obtained by a first-order Taylor expansion around $\tau = 0$ in the LE. By construction, it becomes exact[5] in the limit $\tau \to 0$. Further, the perturbation theory (see Sec. 6.2.2), where the delay force is treated as a small perturbation to

[5] As we have noted in Sec. 6.2.1, this is in fact a delicate point, due to the question whether the Taylor expansion of the stochastic quantity has the same properties as in deterministic cases.

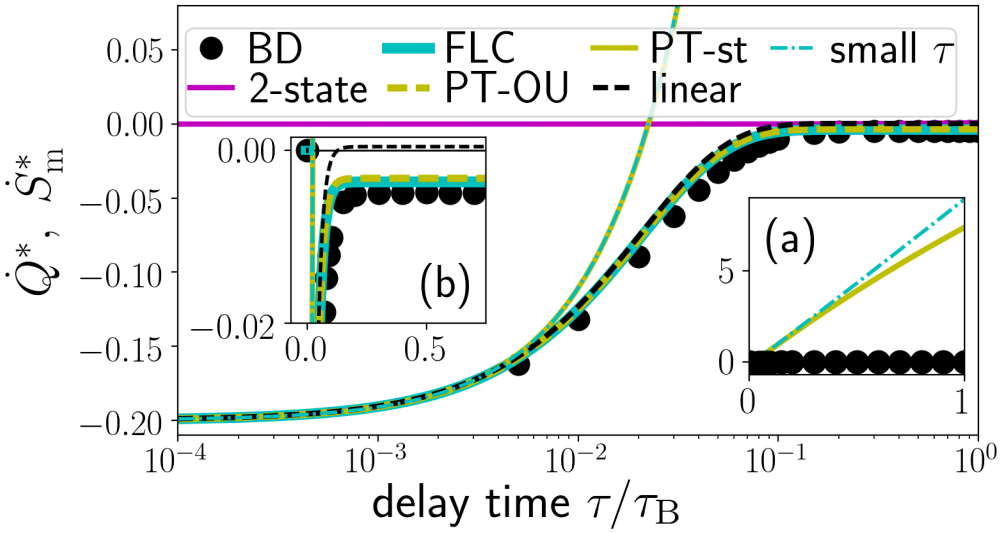

Figure 8.4: *Mean heat rate $\dot{Q}^* = \dot{Q}\,\tau_{\rm B}/(k_{\rm B}\mathcal{T})$ and medium entropy production (EP) rate $\dot{S}^*_{\rm m} = \dot{S}_{\rm m}\,\sigma^2/(k_{\rm B}D_0)$ in the bistable system with different parameters similar to Figs. 8.3 and 8.2, but here at $V_0 = 6k_{\rm B}\mathcal{T}$, $k = 0.2k_{\rm B}\mathcal{T}$. This plot is further complemented with the results from the perturbation theory, once with short time propagator (PT-st), one with Ornstein-Uhlenbeck approximation (PT-OU), shown as dashed and solid yellow lines, respectively. (a) zoom out, (b) magnification. These results are also published in Ref. [2].*

the Markovian dynamics, becomes exact in the limit $b \to 0$ (where this perturbation vanishes). We refer the interested Reader to Chapter 6, where detailed descriptions of the approximation schemes and their limitations are provided.

Limit of vanishing feedback strength We first consider the Markovian limit of vanishing feedback strength, i.e., $b \to 0$. By construction, the one-time steady-state PDFs from the PT (and also from the FLC) become exact at $b \to 0$. They converge in a *continuous* manner to the equilibrium (Boltzmann) distribution $\rho_{\rm ss}^{b=0}$, as can be easily seen from the PT-st approximation from Eq. (6.16). This reads

$$\begin{aligned}\rho_{\rm ss}^{\rm PT}(x) &= Z\exp\left\{-\frac{[1-(b\tau/\gamma)]\,[V_{\rm s}(x)+(b/2)x^2]}{\gamma D_0}\right\}\\ &\xrightarrow{b\to 0}\ \rho_{\rm ss}^{b=0}(x) = Z^{-1}\exp\left[-\frac{V_{\rm s}(x)}{\gamma D_0}\right],\end{aligned} \tag{8.18}$$

with the static potential giving the polynomial force $V_{\rm s}'(x) = \sum_{i=1}^{m} a_i x^i$ and normalization constant Z. After performing a partial integration step and plugging in the natural boundary conditions, the $(i-1)^{st}$ positional moment can therewith be

expressed as

$$\begin{aligned}\left\langle X^{i-1}\right\rangle_{\mathrm{ss}} &= \int_{-\infty}^{\infty} Z^{-1}\exp\left[-\frac{V_{\mathrm{s}}(x)}{\gamma D_0}\right]x^{i-1}\mathrm{d}x \\ &= \left[\rho_{\mathrm{ss}}^{b=0}(x)\,\frac{x^i}{i}\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty}\frac{F_{\mathrm{con}}(x)}{\gamma D_0}Z\exp\left[-\frac{V_{\mathrm{s}}(x)}{\gamma D_0}\right]\frac{x^i}{i}\mathrm{d}x \\ &= \sum_{j=1}^{m}\frac{a_j}{i\gamma D_0}\int_{-\infty}^{\infty}x^j\rho_{\mathrm{ss}}^{b=0}(x)x^i\mathrm{d}x = \sum_{j=1}^{m}\frac{a_j}{i\gamma D_0}\left\langle X^{j+i}\right\rangle_{\mathrm{ss}}, \quad \forall i\geq 1, \end{aligned} \tag{8.19}$$

from which immediately follows

$$\Rightarrow \sum_{i=1}^{m} i\,a_i D_0\left\langle X^{i-1}\right\rangle_{\mathrm{ss}} = \sum_{i=1}^{m}\sum_{j=1}^{m}\frac{a_j a_i}{\gamma}\left\langle X^{j+i}\right\rangle_{\mathrm{ss}}. \tag{8.20}$$

Plugging the last expression into Eq. (8.15), we finally conclude that $\lim_{b\to 0}\dot{Q}=0$ and $\lim_{b\to 0}\dot{S}_{\mathrm{m}}=0$ in a *continuous* manner. This is in agreement with the linear case, see Eq. (8.17).

8.2.3 Limit of vanishing delay time

The other relevant limit is that of vanishing delay time, $\tau\to 0$. The one-time PDF [1] from the small delay expansion [also see Eq. (6.14)]

$$\rho_{\mathrm{ss}}^{s\tau}(x) = Z^{-1}\exp\left[-\frac{V_{\mathrm{s}}(x)+(b/2)x^2}{(\gamma+b\tau)D_0}\right] \overset{\tau\to 0}{\longrightarrow} \rho_{\mathrm{ss}}^{\tau=0}(x) = Z\exp\left[-\frac{V_{\mathrm{s}}(x)+(b/2)x^2}{\gamma D_0}\right] \tag{8.21}$$

clearly becomes exact for $\tau\to 0$, where it *continuously* converges to the Boltzmann distribution $\rho_{\mathrm{ss}}^{\tau=0}$. Performing steps analogously to Eq. (8.20), we obtain for all $i\geq 1$

$$\begin{aligned}\left\langle X^{i-1}\right\rangle_{\mathrm{ss}} &= \left[Z\rho_{\mathrm{ss}}^{\tau=0}(x)\,(x^i/i)\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty}\frac{F_{\mathrm{con}}(x)-bx}{i\gamma D_0}x^i\rho_{\mathrm{ss}}^{\tau=0}(x)\mathrm{d}x \\ &= \sum_{j=1}^{m}\frac{a_j}{i\gamma D_0}\left\langle X^{j+i}\right\rangle_{\mathrm{ss}} + \frac{b}{i\gamma D_0}\left\langle X^{i+1}\right\rangle_{\mathrm{ss}}, \end{aligned} \tag{8.22}$$

as $\tau\to 0$, which results in the identities

$$\sum_{i=1}^{m} i\,a_i D_0\left\langle X^{i-1}\right\rangle_{\mathrm{ss}} = \sum_{i=1}^{m}\sum_{j=1}^{m}\frac{a_j a_i}{\gamma}\left\langle X^{j+i}\right\rangle_{\mathrm{ss}} + \sum_{i=1}^{m}\frac{b a_i}{\gamma}\left\langle X^{i+1}\right\rangle_{\mathrm{ss}}, \tag{8.23}$$

$$\left\langle X^0\right\rangle_{\mathrm{ss}}\equiv 1 = \frac{b}{\gamma D_0}\left\langle X^2\right\rangle_{\mathrm{ss}} + \sum_{j=1}^{m}\frac{a_j}{\gamma D_0}\left\langle X^{j+1}\right\rangle_{\mathrm{ss}}. \tag{8.24}$$

Plugging both into Eq. (8.15) in an iterative manner, yields the *discontinuous* limit

$$\lim_{\tau\to 0}\dot{S}_{\mathrm{m}} = \lim_{\tau\to 0}\frac{k_{\mathrm{B}}b}{\gamma}(1-\delta_\tau) = \begin{cases}0, & \tau=0\\ k_{\mathrm{B}}b/\gamma, & \tau>0\end{cases}. \tag{8.25}$$

This ubiquitous jump-discontinuity of $\dot{Q}$ and $\dot{S}_{\mathrm{m}}$ at $\tau \to 0$ is a central theoretical result. Remarkably, it is independent of the details of the potential landscape. This apparent offset indicates an abrupt qualitative change of the thermodynamics when non-Markovianity sets in. It arises due to the discontinuity of the noise–position cross correlations (8.13) at the onset of causal relationship. Explicit examples are shown Figs. 8.3, 8.4, which we will describe in Sec. 8.3.

We recall that the apparent offset has already been observed and discussed in the context of linear systems [93, 99, 116, 152]. In [99] it has been considered as a consequence of inconsistent usage of Itô and Stratonovich calculus. However, as shown here, it is not a mathematical error but also arises within consistently applied Stratonovich calculus. The underlying reason is the interplay of white noise and delay below the short (ballistic) relaxation timescale [158], which becomes relevant as $\tau \to 0$. Let us further differentiate between physical effects which should be measurable in experiments, and consequences of the model. To better understand the impact of the overdamped limit, we now consider the full Langevin equation.

Ballistic motion To further elucidate the behavior at $\tau \to 0$, we also briefly consider the limit in the *underdamped* case, where the LE

$$m\ddot{X}(t) = -\gamma\dot{X}(t) + F_{\mathrm{con}}[X(t)] - bX(t-\tau) + \gamma\sqrt{2D_0}\,\xi(t) \quad \text{(underdamped LE)}, \tag{8.26}$$

involves the inertial term with mass m, yielding ballistic motion below the velocity-relaxation timescale, m/γ. For this system, the heat can, in principle, be calculated using the same expression as in the overdamped case [Eqs. (8.4, 8.2)]. For linear systems, it has been shown by [93] via a Green's functions method, that $\dot{Q}$ smoothly decays to zero with τ in the underdamped case. For nonlinear systems, this strategy cannot be applied and the calculations are much more involved.

To obtain a statement for nonlinear systems, we should consider the correlation $\langle X(t-\tau)\xi(t)\rangle$ appearing in Eq. (8.5), which, in our framework, "causes" the discontinuity from a mathematical point of view [compare Eq. (8.13, 8.25)].

Derivation of noise–position (8.32) and noise–velocity cross-correlations for underdamped dynamics We aim to calculate the correlations $\langle X(t-\tau)\xi(t)\rangle$ and $\langle \dot{X}(t-\tau)\xi(t)\rangle$ for underdamped motion with time-delayed feedback. To this end, we first again use the causality argument (i.e., the noise cannot influence the past position or velocity, see Sec. 1.3.5). This implies that both correlations must vanish at $\tau > 0$. Second, we rewrite the correlations at $\tau = 0$ as follows, starting from a formal integration over the LE (8.26)

$$m\dot{X}(t) = -\gamma X(t) + \sum_i a_i \int_0^t X(s)^i \mathrm{d}s - b\int_0^t X(s-\tau)\mathrm{d}s + \gamma\sqrt{2D_0}\int_0^t \xi(s)\mathrm{d}s + \mathcal{C}, \tag{8.27}$$

with constant $\mathcal{C}$ accounting for initial conditions. From this equation, we readily find

$$\begin{aligned} m\langle\dot{X}(t)\xi(t)\rangle =& -\gamma\langle X(t)\xi(t)\rangle + \sum_i a_i \int_0^t \left\langle X(s)^i\xi(t)\right\rangle \mathrm{d}s - b\int_0^t \langle X(s)\xi(t)\rangle\,\mathrm{d}s \\ &+\gamma\sqrt{2D_0}\int_0^t \langle\xi(s)\xi(t)\rangle\,\mathrm{d}s \\ =& -\gamma\langle X(t)\xi(t)\rangle + \sum_i a_i \int_0^t \left\langle X(t)^i\xi(t)\right\rangle \delta_{s,t}\mathrm{d}s - b\int_0^t \langle X(t)\xi(t)\rangle\,\delta_{s,t}\mathrm{d}s \\ &+\gamma\sqrt{D_0/2}. \end{aligned} \tag{8.28}$$

In the last step, we have simplified the integrals of types $\int_0^t \langle X(s)^i\xi(t)\rangle\,\mathrm{d}s$ by using the causality argument which implies that the integrands are zero for all $s < t$. Third, we use the identity $X(t) = \int_0^t \dot{X}(s)\mathrm{d}s + X(0)$, and find

$$\langle X(t)\xi(t)\rangle = \int_0^t \left\langle \dot{X}(s)\xi(t)\right\rangle \mathrm{d}s = \int_0^t \langle\dot{X}(t)\xi(t)\rangle\delta_{s,t}\,\mathrm{d}s. \tag{8.29}$$

Fourth, via a proof by contradiction, we establish that Eqs. (8.28, 8.29) only allows for *finite* solutions. Starting with the assumption $\langle\dot{X}(t)\xi(t)\rangle \to \infty$, Eq. (8.29) implies a finite value of $\langle X(t)\xi(t)\rangle$. This, then again yields a finite value of $\langle\dot{X}(t)\xi(t)\rangle$ from (8.28), thus, a contradiction. Using this finiteness, we finally obtain for $\tau \geq 0$

$$\langle X(t-\tau)\xi(t)\rangle = 0, \tag{8.30}$$

$$\langle\dot{X}(t-\tau)\xi(t)\rangle = (\gamma/m)\sqrt{D_0/2}\,\delta_\tau. \tag{8.31}$$

Thus, we find that in contrast to the overdamped case, the position–noise cross-correlation is continuous at $\tau = 0$,

$$\langle X(t-\tau)\xi(t)\rangle = 0,\ \forall\tau \geq 0 \tag{8.32}$$

(not only for linear systems), indicating that $\dot{Q}$ behaves continuously. Therefore, the discontinuity is absent in the underdamped description. This is in agreement with previous results for underdamped systems, e.g., see Fig. 1 in [93].

8.2.4 Discussion of the behavior for small delay times

Here we discuss three interesting observations that follow from the previous considerations: The discontinuous limit, the sign of the heat rate and the growth of its absolute value when τ is reduced.

Meaning of discontinuous limit from a broader perspective Above we have shown that the heat flow is discontinuous at $\tau \to 0$ in the overdamped limit, while it behaves continuously in the underdamped equation. The discontinuity is thus a consequence of the overdamped limit [recall (8.32) vs. (8.13)].

From the viewpoint of theoretical frameworks of stochastic modeling, this is very interesting, because the overdamped limit is an extremely common and well-established assumption, which normally does not yield unphysical behavior, especially not on the large timescales. Here, we consider the long-time behavior $t \to \infty$, and find that it nevertheless breaks down. The reason is that we – although focusing on steady-state behavior – indeed implicitly address the ballistic timescale, when we let τ go to zero.

We would like to remind the Reader that in models of feedback-controlled systems, the assumption of (vanishingly) small τ, is very common. While it often renders nice results for the dynamical quantities, the thermodynamic considerations reveal that it is indeed somewhat problematic. This is an example for a general observation we make: the thermodynamic quantities are more "sensitive" then the typically considered dynamical quantities[6].

Reversed heat flow Turning back to our analytical formula for the limit in the overdamped case, i.e., Eq. (8.25), we further find that at small, positive τ, $\dot{S}_{\rm m}$ and $\dot{Q}$ have negative values. This is true if $b < 0$, which is always the case in an optical tweezers setup (the optical trap is attractive, i.e., pulls the particle towards the delayed position). The negative sign of $\dot{Q}$ indicates a steady heat flow from the bath to the particle. This corresponds to energy *extraction* from the heat bath, or in other words, feedback cooling of the surrounding fluid. This is a delay-induced phenomenon, which would be impossible in the Markovian counterpart of this system due to the second law, and is one of the goals of "Maxwell demon"-types of devices. We will explain this point in more detail in the next Chapter (see Sec. 9.4.2). The fact that $\dot{S}_{\rm m} < 0$ (the bath constantly loses entropy) underlines that further entropic terms must contribute to the (nonnegative) total entropy production (as discussed for underdamped dynamics in [93, 116, 152]). In the regime of reversed heat flow, the amount of medium entropy loss thus provides a lower bound to the *entropic cost* of the feedback. We will explicitly discuss the entropy balance, as well as the information flow which is a needed to generate such a negative heat flow, in Chapter 9.

Growth upon reduction of delay time The (negative) heat flow is quite pronounced in the regime of small τ, and *increases* in absolute value, when reducing τ (compare with Fig. 8.3). Notably, this is also true for underdamped linear systems, where $\dot{Q}$ only decays for τ *below* the velocity-relaxation timescale (compare with Fig. 1 in [93]). Analytical evaluation of the heat rate for a nonlinear example system, which we will provide below, reveals the same behavior.

Before proceeding with this concrete examples, we here propose an explanation of the phenomenon that $|\dot{Q}|$ increases with τ.

When τ is so small that

$$X(t-\tau) \approx X(t) - \tau \dot{X}(t-\tau),$$

[6]The underlying mathematical reason is that the thermodynamic quantities naturally involve higher order correlations.

the delayed feedback force $F_\mathrm{d} = -bX(t-\tau)$ gets a contribution proportional to the strongly fluctuating *velocity*. This changes the induced steady-state heat flow $\dot{Q} = \langle F_\mathrm{d} \circ \dot{X}(t)\rangle$ significantly. For $b < 0$, $F_\mathrm{d} \sim \tau b\dot{X}(t-\tau)$ has a *friction*-like contribution. As is well-known for underdamped systems with velocity-dependent (delayed [93] or non-delayed [63, 231]) feedback, such a control yields medium entropy reduction due to *entropy pumping*. Please note that a velocity-dependent feedback $F_\mathrm{d} \sim b\dot{X}(t)$ drives a system out of equilibrium [63, 231] even *without* delay, contrary to a (non-delayed) position-dependent control. Entropy pumping is for example utilized to realize "molecular refrigerators", see [63, 93, 231].

The reason is that the additional friction-like force reduces the thermal fluctuations of the particle, inducing an energy transfer from the bath to the particle, hence, a heat flow. Reversely, the particle fluctuations are enhanced for $b > 0$, yielding a positive heat flow. This means, an additional "entropy pumping" contribution to the heat flow arises for small τ, explaining the enlarged heat flow at small delay times. Please note that this effect is independent of the question whether the overdamped limit is used, or not.

When τ is even smaller and gets below the velocity-relaxation time m/γ, both systems behave differently. For underdamped dynamics, the velocities are correlated, implying that the delay-induced heat flow $\dot{Q} \sim \tau\, b\langle\dot{X}(t-\tau)\dot{X}(t)\rangle \sim \tau\, e^{-\gamma\tau/m}$ eventually decays smoothly to zero (resulting in a maximum of $|\dot{Q}|$ around $\tau = m/\gamma$, see Fig. 1 in [93]). For overdamped dynamics, the velocities are, in contrast, uncorrelated. This is because the white noise directly acts on it $\dot{X} \sim \xi$, giving rise to the nontrivial limit $\tau\,\delta(\tau)$ of type "$0 \times \infty$". This yields a finite value at $\tau \to 0$, as we know from Eq. (8.25).

Hence, in an experimental setup, a position-dependent feedback is expected to induce heat flow, unless the (typically unavoidable) delay is well-below the velocity-relaxation timescale.

8.3 Application to the bistable potential

Now, we reconsider the bistable example system involving the doublewell potential with minima at $\pm\sigma$ and potential barrier height V_0, and the delayed optical trap potential $(k/2)[(x/\sigma) - (x_\tau/\sigma)]^2$, which we have already discussed in previous Chapters, see e. g. Secs. 1.5, 6.4, 6.5.2, and 7.1.

The potentials yield a polynomial deterministic force $F = -a_3x^3 - a_1x - bx_\tau$ with coefficients

$$b = -k\sigma^{-2} \tag{8.33}$$

$$a_1 = (-4V_0 + k)\sigma^{-2} \tag{8.34}$$

$$a_3 = 4V_0\sigma^{-4}. \tag{8.35}$$

For this example system, the mean heat rate from Eq. (8.15) involves the even moments up to 6^{th} order.

As in Part II, we will consider both: the nontrivial intrawell dynamics within the asymmetric potential wells, and the noise- and delay-induced positional oscillations

between the wells. Again, we focus on the regime where the deterministic counterpart of the system has no delay-induced oscillations, i.e., $0 < k < 4V_0$ (see Sec. 1.7). Thus, the oscillations are delay- and noise-induced.

We will also use BD simulations (see Appendix A.1), to validate the analytical results. As a first step, we check the consistency between simulations and theory, and, at the same time, confirm the analytical expression (8.15). Computing the NESS heat rate, both, from the stochastic definition along each simulated trajectory (8.4) and from the numerically obtained PDF (8.15), we indeed find perfect agreement.

In the following, we compare the BD results to those obtained from our analytical approach, i.e., Eq. (8.15) combined with established approximations. Since the approximations are known to perform best when the particle is likely to stay around a potential minimum (see Chapter 6), this approach seems appropriate in the low thermal energy regime ($\gamma D_0 \ll V_0$). At the end, we will introduce a complementary approach for larger noise levels.

8.3.1 Low thermal energy – intrawell dynamics

Figure 8.3 shows $\dot{Q} = \dot{W}$ and $\dot{S}_{\mathrm{m}}$, as functions of the delay time τ, for an exemplary parameter setting in the low noise regime. The simulation results confirm the discontinuous τ-limit with apparent offset given by Eq. (8.25), and the predicted reversed heat flow. Similar results as in Fig. 8.3 but for a different parameter setting are also displayed in Fig. 8.4, which additionally shows the PT approximations. The PT-OU predictions are quantitatively very similar to the ones obtained by the FLC, while the PT-st yields very similar results to the small delay expansion. We will therefore not mention the PT approximations explicitly in the following, but rather discuss only the FLC and the small delay expansion.

While the small delay expansion fails outside the Markovian limit, our analytical approach with FLC makes quantitatively correct predictions for all delay times considered in Fig. 8.3. The thermodynamic quantities grow with τ, until they approach constant values.

Saturation Interestingly, this saturation occurs around the intrawell relaxation timescale, which can roughly be estimated by $\tau_{\mathrm{ir}} \approx 1/8\, k_{\mathrm{B}}\mathcal{T}/V_0$ in the present case (see Sec. 1.6). Note that this estimation relies on arguments for Markovian systems and is hence rather a crude estimate. The underlying reason for this saturation is the same as of the PDF's saturation, which we have discussed in Sec. 6.4. Figure 8.3 also displays the results from the corresponding linear system, $F = -\alpha x - \beta x_\tau$. The parameters are $\alpha = -(8V_0 + k)/\sigma$ and $\beta = -k/\sigma$, as obtained by a second order Taylor expansion of $V_{\mathrm{s}} + V_{\mathrm{d}}$ around a (deterministically) stable fix point $(x, x_\tau) = \pm(\sigma, \sigma)$. Interestingly, $\dot{Q}$ is not only equivalent to the nonlinear case for very small τ [as already expected from (8.25)], but also saturates on the same timescale.

Reversed heat flow and break down of approximations The agreement between BD and FLC persists at larger barrier heights V_0 and at smaller k. However,

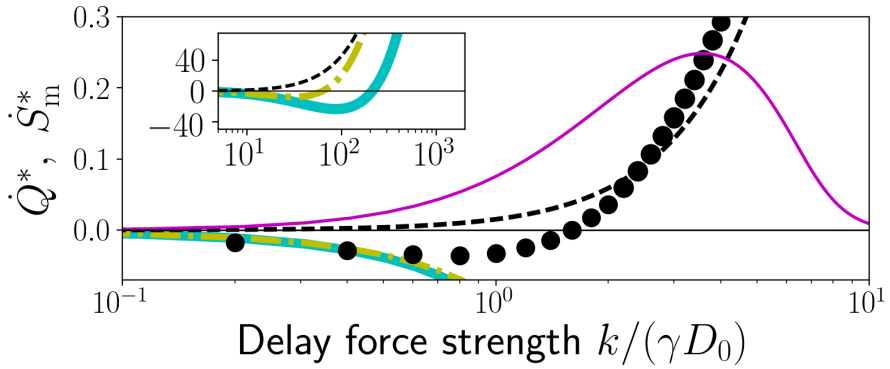

Figure 8.5: *Scaled NESS mean heat rate $\dot{Q}^* = \dot{Q}\,(\tau_{\mathrm{B}}/(k_{\mathrm{B}}\mathcal{T}))$ and medium entropy production rate $\dot{S}^*_{\mathrm{m}} = \dot{S}_{\mathrm{m}}\,\sigma^2/(k_{\mathrm{B}}D_0)$ vs. delay force strength k over thermal energy γD_0; at $V_0 = 4\gamma D_0$, $\tau = \sigma^2/D_0$. Inset: zoom out. Color code as in Fig. 8.3. These results are also published in Ref. [2].*

for larger $k/(\gamma D_0)$, also the FLC approach breaks down, as can be seen in Fig. 8.5. The inset shows that the FLC nevertheless captures the qualitative behavior (but quantitatively, it is really far off). Interestingly, upon increase of k (i.e., the laser intensity in case of optical tweezers [220]), the direction of the mean heat flow can be changed. Thus, in the nonlinear case, one may switch from feedback cooling to heating of the surrounding fluid, and both can be induced by employing *negative feedback*. We recall that this is the type of feedback realizable with an optical trap, where always $k > 0$. This contrasts the linear system (dashed line in Fig. 8.5). There, we could not find a parameter regime where the direction of the heat flow can be revered upon increase of positive k. Instead, the reversal seems to occur when k changes its sign, or upon increase of τ.

The current reversal implies the existence of a certain finite delay force strength k, where the mean heat flow is zero, despite the presence of time-delayed driving. For example, in Fig. 8.5, the BD simulations indicate vanishing heat around $k \approx 1.7\gamma D_0$. Upon increase of τ, the linear system also shows a point of zero heat rate, despite the presence of the delay force, see 8.4 (b) for an example. Recall that $\delta W = 0$ further implies $\delta W = 0$, while $\mathrm{d}U = 0$ is readily given in the steady state. Thus, *all* energy (ex)changes are zero, as well as the medium EP. Based on this consideration, we cannot see any trace of nonequilibrium here, but we do not think that the system is in thermal equilibrium[7]. We will clarify this point in the next Chapter, where we will reconsider the vanishing heat flow from an entropic perspective. An interesting objective of future work would be to see how the fluctuations of the heat and other thermodynamic quantities behave at this point.

8.3.2 High thermal energy – interwell dynamics

When the thermal energy is sufficiently high, jump processes between the potential wells dominate the dynamics. Then, the previously discussed approximations

[7] Contrary to $k = 0$, where the heat flow nullifies as well and thermal equilibrium is approached.

by construction all break down. The reason is that the "memory of a jump" expressed in the non-Markovian ρ_2 (and higher PDFs), is then what dominates the process and, e.g., triggers subsequent jumps; and precisely these higher PDFs are the (over)simplified quantities in all approaches. In particular, the FLC effectively assumes an harmonic potential well out of which no jumps can occur, while the PT and small delay expansion render Markovian systems which don't exhibit spontaneous oscillations, as the latter are a delay-induced phenomenon. This includes situations, where the interplay of noise and delay leads to the spontaneous oscillations of X. However, we can still treat this regime analytically via an alternative strategy, the *2-state reduction*, which bases on the discretized approach for multistable systems [256] (which we have mentioned earlier in Chapter 2).

To motivate this ansatz in the present context, let us first look at the waiting time distributions.

Waiting time distributions The delay-induced oscillations manifest in the waiting time distributions, as we have already mentioned in Chapter 6, and exemplary shown in Figs. 6.6, 6.10 there. Here, we revisit this idea and consider in more detail waiting time distributions for different thermal energies. Let us first remind the Reader that in the Markovian case, for example at $k = 0$, the distributions are exponential decays whose exponent is given by the Arrhenius formula (see Sec. 1.6.1, specifically the blue histograms in Fig. 6.6). Here, in Fig. 8.6, the corresponding waiting time distributions according to Kramers theory are plotted by blue dashed lines.

Figure 8.6 further shows numerically obtained distributions before, at and after coherence resonance. For very small thermal energy levels below CR (upper left panel), the jumps are very rare with extremely long waiting times before a jump occurs. Some jumps are followed by a quick subsequent jump back to the original valley which occurs within the time interval $[0, \tau]$, because in this time, the delay force drags the particle back. These subsequent jumps yield a small peak within $[0, \tau]$ (note the different y-axis-scales, which make this peak appear higher). We observe that these jumps typically occur two or three times in a row and then randomly stop again. When the thermal energy is increased, the number of escape events grows and also delay-induced *oscillations*, i.e., a sequence of regular jumps with period $\sim \tau$, are getting more likely. These oscillations manifest in a high jump probability around $\tau/2$. This value is marked by a gray vertical line in Fig. 8.6 as a guide to the eye. At thermal energy in the range of CR (upper right and lower left panel), a pronounced peak around $\tau/2$ is shown, caused by the delay-induced oscillations. For even higher thermal energies, there is still a pronounced peak around $\tau/2$, but the exponential background gets stronger. This is because, for high $\gamma D_0/V_0$, the effect of the delayed force on the dynamics becomes less important, as compared to the thermal interwell motion. In the lower left panel an exemplary value slightly above CR is depicted. Finally, for even higher thermal energies, the waiting time distribution is dominated by the exponential thermal decay, and the distributions is very similar to the case of a Markovian Kramers escape problem (lower right panel). The latter is depicted by blue lines in Fig. 8.6.

Quasistatic approximation As a crude estimation of the waiting time distributions in the non-Markovian case, one can use the following hand waving argument based on a quasistatic approximation. The quasistatic approximation is expected to be justified when the delay time is small compared to the intrawell relaxation time, $\tau \gg \tau_{\rm ir}$, and $\gamma D_0 < V_0$.

The true system has a history-dependent energy landscape with dynamical barrier. However, when the particle is in the same valley as it was τ times ago, which we denote case 1, the effective energy landscape can roughly be estimated by fixing x_τ to the minima around which x can be found which yields the quasistatic potential $V_{\rm qs}(x) = V_{\rm s}(x) + V_{\rm d}[x, x_\tau = \sigma\,{\rm sign}(x)]$. This is approximately the situation for all detected jumps with $\Delta t_{\rm jump} > \tau$. The red lines in Fig. 8.6 plot the corresponding Kramers estimate, which will be given below.

On the contrary, in case 2, that is, when the particle has jumped within the last τ times, the quasistatic potential reads $V_{\rm qs}(x) = V_{\rm s}(x) + V_{\rm d}[x, x_\tau = -\sigma\,{\rm sign}(x)]$. This is approximately the situation for all detected jumps with $\Delta t_{\rm jump} < \tau$ (orange lines in Fig. 8.6).

The transition rates in cases 1 and 2 can be estimated by the two Kramers rates associated with $V_{\rm qs}$, given by (1.58) with $r_{\rm K,j} = r_{\rm K}(x_{\rm min,j})$ for escapes out of the two minima, $x_{\rm min,1} = \pm\sigma$ and $x_{\rm min,2} = \mp(\sigma/2)[1 + \sqrt{1-(k/V_0)}] \approx \mp\sigma$, over the barrier at $x_{\rm max} = \mp(\sigma/2)[1 - \sqrt{1-(k/V_0)}] \approx 0$. In the last step we have used twice $\sqrt{1-(k/V_0)} \approx 1$ for $k \ll V_0$. The corresponding second derivatives can be calculated exactly as $V''_{\rm qs}(x_{\rm max}) = -(4V_0 - k)/\sigma^2$ and $V''_{\rm qs}(x_{{\rm min,j}\in\{1,2\}}) = (8V_0 + k)/\sigma^2$, while the potential barrier heights read $\Delta V_{\rm qs,1} = V_0 + k/2$, and $\Delta V_{\rm qs,2} = V_0 - 3k/2$. This finally yields the transition rates

$$r_{{\rm K},j\in\{1,2\}} = \sqrt{(8V_0 + k)(4V_0 - k)}/\left(2\pi\gamma\,\sigma^2\right)\,\exp\left(-\Delta V_{\rm qs,j}/\gamma D_0\right), \tag{8.36}$$

One can see in Figure 8.6 that the Kramers distributions from the quasistatic approximations indeed roughly capture the characteristics of the true jump distributions, at least in the regime $\Delta t_{\rm jump} > \tau$. Of course, the characteristic shape of the distributions in the interval $[0, \tau]$ is not described by this exponential ansatz, but the decay rates are in fact, surprisingly accurate.

2-state reduction We have seen that the waiting time distributions can roughly be approximated by two quasistatic situations, where we have fixed the delayed position to be either in the same, or in the opposite potential valley. Inspired by this finding we go one step further, and ask: *Is it possible to describe the interwell dynamics by only taking into account in which minima $X(t)$ and $X(t-\tau)$ are? And is this description sufficient to capture the characteristics from a thermodynamic perspective?* In the regime, where the dynamics is dominated by the jumps between the valleys, it seems somewhat natural to take the intrawell fluctuations out of the description.

This leads us to the state discretization proposed by Tsimring and Pikovsky [256], which is a type of coarse-graining of the system state space. In particular, the X-dynamics is reduced to switching processes between *two discrete* states $s = \pm\sigma$, corresponding to the two potential valleys. There are two different transition rates

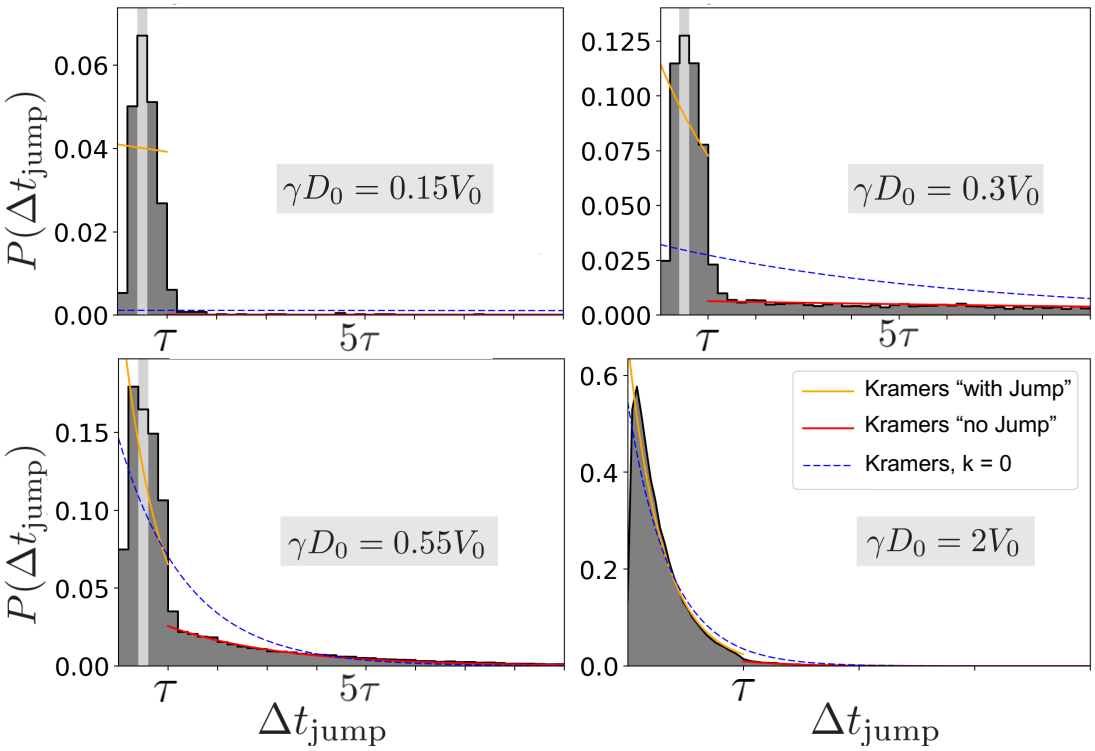

Figure 8.6: *Waiting time distribution for different ratios of thermal energy over barrier height $\gamma D_0/V_0$, below and above coherence resonance (CR). The value $\tau/2$ is marked by a vertical gray line as a guide to the eye. Upper left: Small thermal energy below CR. Upper right: Thermal energy in the range of CR, pronounced peak around $\tau/2$. Lower left: Thermal energy slightly above CR. Lower Right: Very high thermal energy. The distributions are numerically (BD) generated (normalized) histograms of the times between sequential jumps Δt_{jumps} in the doublewell potential with linear delay force at $k = 0.6V_0/\sigma$ and $\tau = 5\tau_{\text{B}}$. The orange, blue and red lines indicate estimates based on Kramers theory, see text for details.*

corresponding to the two possible scenarios that the delayed and instantaneous state have the same, or opposite sign, respectively, which are just given by $r_{\text{K},\{1,2\}}$ from the quasistatic approximation (8.36).

This 2-state reduction brings two simplifications to the calculation of the heat rate from Eq. (8.5). First, various relevant temporal position-autocorrelation functions are significantly simplified, since the intrawell dynamics is omitted. In particular, (for $\tau > 0$)

$$\left\langle X(t)^2 \right\rangle \to \left\langle s(t)s(t) \right\rangle = \left\langle \sigma^2 \right\rangle = \sigma^2 \tag{8.37}$$

$$\left\langle X^i(t)X(t-\tau) \right\rangle \to \left\langle s^i(t)s(t-\tau) \right\rangle = \begin{cases} \sigma^i \left\langle s(t-\tau) \right\rangle = 0, & \forall \text{ even } i \\ \sigma^{i-1} \left\langle s(t)s(t-\tau) \right\rangle = C_1(\tau)\, \sigma^{i-1}, & \forall \text{ odd } i \end{cases}. \tag{8.38}$$

Thus, all relevant correlations can be expressed via $C_1(\tau)$. Consequentially, Eq. (8.5)

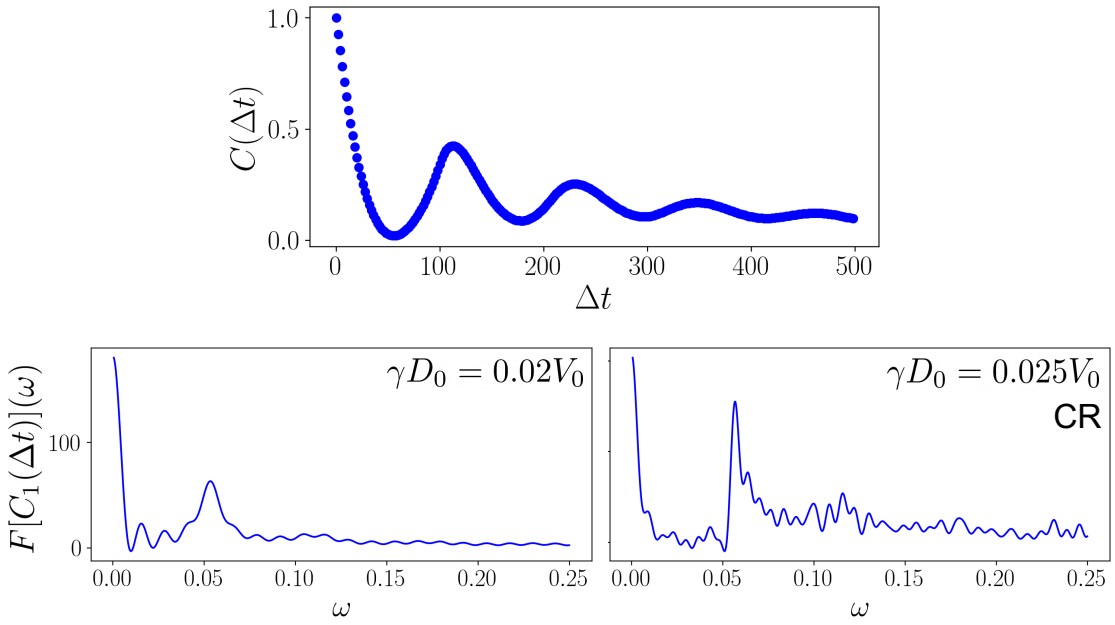

Figure 8.7: *Detecting CR in the temporal position-autocorrelation function* $C_1(\Delta t) = \langle X(t)X(t-\Delta t)\rangle$ *from numerical results. Upper plot:* $C_1(\Delta t)$ *itself (at CR), which shows peaks around multiples of* τ*. Lower plot: The power spectrum, i.e., the Fourier transformation of* $F[C_1(\Delta t)](\omega)$*, left: at a value of* $\gamma D_0/V_0$ *before, and right: at CR. The position of the main peak around* $\omega = 0.057$ *is close to* $2\pi/\tau$*. The height of this peak is used as order parameter to detect coherence resonance. Here, the parameters are* $V_0 = 4$*,* $k = 0.1$*, and* $\tau = 100$*. This corresponds to the parameters* $k = 0.4V_0$*,* $\tau = 25\gamma\sigma^2/V_0$*, which are also considered in the following Figures 8.8 and 8.9.*

simplifies to

$$\begin{aligned}\dot{W} &= (b/\gamma)\left(a_1 + \sigma^2 a_3\right) C_1(\tau) + \left(b^2/\gamma\right)\sigma^2 \\ &= k^2/\left(\gamma\sigma^2\right)\left[1 - C_1(\tau)/\sigma^2\right] = \dot{Q} = \dot{S}_{\mathrm{m}}\gamma D_0/k_{\mathrm{B}},\end{aligned} \tag{8.39}$$

where we have used that the mean change of internal energy must vanish in the NESS.

Second, the full position-autocorrelation function $C_1(\Delta t)$ can now be calculated analytically, via

$$C_1(\Delta t) = \frac{\left(\sqrt{r_{\mathrm{K},1}} + \sqrt{r_{\mathrm{K},2}}\right) e^{-2\sqrt{r_{\mathrm{K},1} r_{\mathrm{K},2}}\,\Delta t} + \left(\sqrt{r_{\mathrm{K},1}} - \sqrt{r_{\mathrm{K},2}}\right) e^{-2\sqrt{r_{\mathrm{K},1} r_{\mathrm{K},2}}\,(\tau - \Delta t)}}{\sqrt{r_{\mathrm{K},1}} + \sqrt{r_{\mathrm{K},2}} + \left(\sqrt{r_{\mathrm{K},1}} - \sqrt{r_{\mathrm{K},2}}\right) e^{-2\sqrt{r_{\mathrm{K},1} r_{\mathrm{K},2}}\,\tau}}, \tag{8.40}$$

which was derived in [256]. We have plotted $C_1(\Delta t)$ for three exemplary values in Fig. 8.8. Thus, from Eqs. (8.40, 8.39), the mean heat rate and medium EP rate can readily be evaluated. Corresponding results are plotted in Fig. 8.9 by magenta lines.

Coherence resonance As we have already mentioned in Sec. 1.7, the delay-induced oscillations manifest in the position-autocorrelation function. The latter

can be computed by numerical simulations, and from the 2-state reduction (8.39). The delay-induced oscillations have a period of about τ. Thus, the more regular and likely the oscillations are, the more pronounced are the peaks around τ, 2τ, 3τ,... in $C_1(\Delta t)$. This shows up as a significant peak around a frequency of $\omega = 2\pi/\tau$ in the power spectrum, i.e., in the Fourier-transformation of $C_1(\Delta t)$. As suggested in [256], we use the peak's height as an order parameter to detect CR. Figure 8.7 shows a numerically obtained correlation function and the corresponding power spectrum, while Fig. 8.8 shows power spectra from the 2-state reduction. Figure 8.8 further displays the order parameter from both, simulation and approximation, which give very similar results, confirming the idea of employing a 2-state reduction. The order parameter has a maximum at $0.36\gamma D_0/V_0$ indicating CR.

Heat and medium entropy production The numerical data in Fig. 8.9 reveals that, when the thermal energy is small compared to V_0, $\dot{Q}$ increases linearly with γD_0, as it also does for a linear delayed system [see Eq. (8.17)]. In this regime, approximating the doublewell by its linearized version (dashed line) even renders the correct slope. By contrast, the 2-state model yields $\dot{Q} = 0$. This is expected, since the intrawell dynamics is completely neglected. At larger $\gamma D_0/V_0$, the slope of $\dot{Q}$ abruptly changes and nonlinear behavior sets in. This occurs at the onset of the delay-induced oscillations, as reflected by a sudden increase of the CR order parameter from both, the 2-state reduction (8.40), and from numerical BD simulations (see Fig. 8.8). The actual value of $\gamma D_0/V_0$ where the nonlinear behavior sets in, is predicted more or less correctly by the 2-state reduction.

By further increasing $\gamma D_0/V_0$, one enters the regime where delay-induced oscillations dominate the dynamics. At a certain finite thermal energy of about $0.36\gamma D_0/V_0$ (indicated by cyan vertical lines in Fig. 8.9 and Fig. 8.8, and plotted with a cyan line in Fig. 8.8 a.), the peak in the power spectrum is highest, resulting in a maximum of the order parameter. In this range of thermal energies, the delay-induced oscillations with mean period $\approx \tau$ are most pronounced.

We find that in this range of $\gamma D_0/V_0$, the 2-state reduction renders a qualitatively good approximation. In particular, it predicts accurately that a region of steep slope is followed by a lower slope of $\dot{Q}$ accompanied by a maximum of $\dot{S}_\mathrm{m}$, and that the latter lies in the regime of CR. However, it yields an overestimation (note the different Y-axis scales).

Generality To test the generality of the discussed behavior and of the approximations, we have tested several parameter settings. Figures 8.10 and 8.11 explicitly show the results for the case $k = 1\,V_0$, $\tau = 5\,\gamma\sigma^2/V_0$. We again observe the coincidence of CR order parameter maximum and maximum in the medium EP. The plots in Fig. 8.10 are complemented by the results for the corresponding linearized system. Like in Fig. 8.9, the linear system has a very comparable heat rate in the regime of small $\gamma D_0/V_0$, where the particle only "sees" its approximately quadratic surrounding (dashed cyan lines). This time, we have also considered the regime of *high* $\gamma D_0/V_0$, where the particle does not "see" the potential barrier (blue dashed lines). Here we also find that the heat rate is similar to a particle in the correspond-

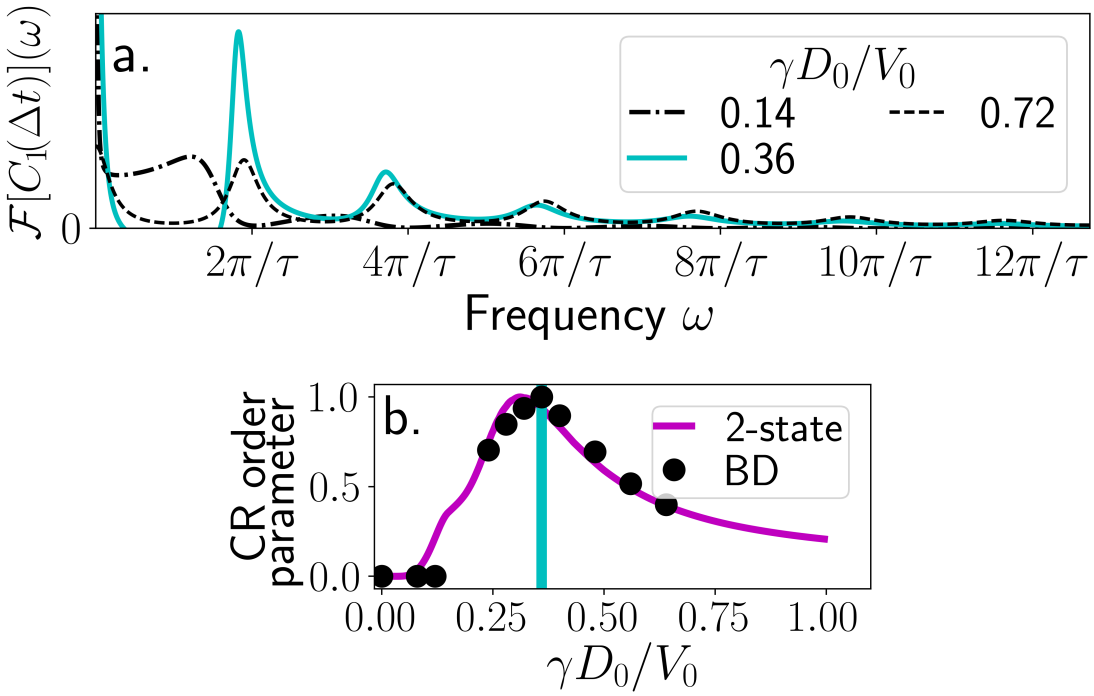

Figure 8.8: *Detecting coherence resonance (CR) in the temporal position-autocorrelation function in the bistable system with $k = 0.4V_0$, $\tau = 25\gamma\sigma^2/V_0$. a. Analytical power spectrum (from 2-state model) given by the Fourier transform of $C_1(\Delta t) = \langle X(t)X(t-\Delta t)\rangle_{\rm ss}$ [Eq. (8.40)], at three thermal energies: below (0.14), at (0.36) and above (0.72) the value of CR. Note that Fig. 8.7 shows corresponding results from numerical simulation, which are naturally a bit noisy. b. Order parameter: Normalized main peak at $\omega \approx 2\pi/\tau$ of the power spectrum of C_1, from BD (symbols) and from the 2-state model (magenta line). The vertical line at $0.36\gamma D_0/V_0$ indicates the maximum around which the system is coherence-resonant. These results are also published in Ref. [2].*

ing quadratic trap. The agreement with the numerical results for the doublewell potential is remarkable, especially given the fact that the fourth order potential is not well-describable by a quadratic trap. For this reason, we do not expect this agreement to generally hold. It is nevertheless an indication for the fact that a linearized system captures the dynamics quite well, except for the delay-induced oscillations, which dominate around CR.

For the second parameter case, Fig. 8.11 shows the escape times estimated by the Kramers rates in the corresponding system without delay ($\tau = 0$). Interestingly, these escape times are comparable to the delay time, $\tau_{\rm K} \approx \tau$, at the very ratio of thermal energy and barrier height, where CR is observed (compare left and right panel in Figure 8.11). This justifies the idea that the pronounced peak in the power spectrum signals a *resonant* response of the system to the delay.

We have tested several parameter settings confirming that the $\dot{S}_{\rm m}$ is indeed maximal at CR conditions as predicted by the 2-state model. Quantitatively, we always observe an overestimation of the heat rate by a factor of about three, as can be seen in Figs. 8.9 and 8.10. This is somewhat surprising, since the phase-space reduction

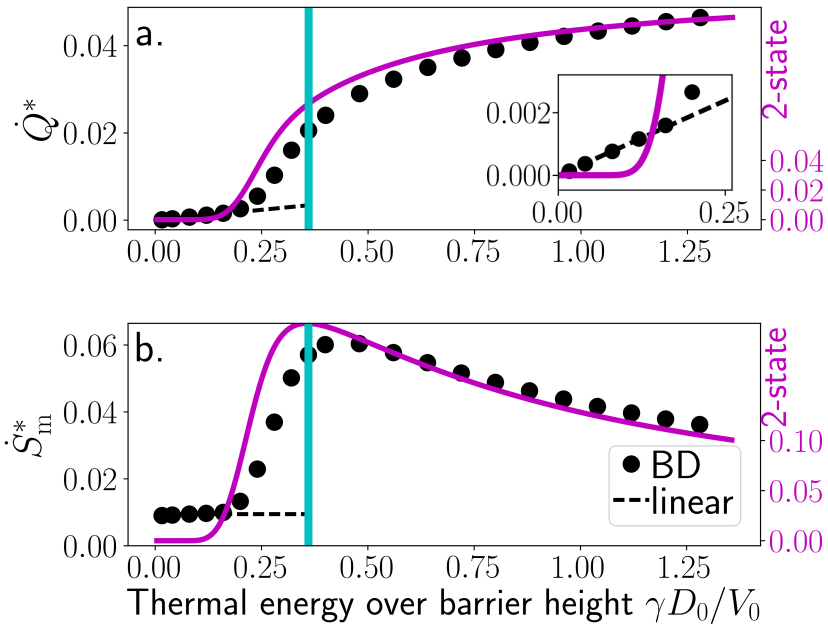

Figure 8.9: *Influence of the ratio of thermal energy γD_0 and barrier height V_0 on mean heat and medium entropy production, in the delayed bistable system at $k = 0.4\,V_0$, $\tau = 25\,\gamma\sigma^2/V_0$. a. Heat rate $\dot{Q}^* = \dot{Q}\,(\gamma\sigma^2/V_0^2)$, the inset gives a magnification, b. Medium EP rate $\dot{S}_{\rm m}^* = \dot{S}_{\rm m}\,(\gamma\sigma^2/V_0 k_{\rm B})$. Magenta lines and y-axes display the results from the 2-state reduction. Vertical cyan lines: Maximum of CR order parameter (see Fig. 8.8). The symbols are the (BD) simulations, the dashed lines correspond the the linearized system. These results are also published in Ref. [2].*

and coarse-graining inherent to the discretization is rather expected to yield an *underestimation* of the heat and medium EP. This was explicitly shown in Ref. [126], where, however, a different kind of coarse-graining was considered. A possible explanation lies in the involved quasistatic approximation, or it could be due to the details of this very coarse-graining, but the precise reason is unclear. This would be an interesting subject of future research.

Explanation for the observed maximum of entropy production at CR The behavior of $\dot{Q}$ and $\dot{S}_{\rm m}$ can be understood on the basis of the 2-state model, specifically Eq. (8.39), which neglects the intrawell fluctuations. If the particle rests in one potential valley, this yields $C_1(\tau) = 1$. Further, when the delay-induced oscillations occur, one would still expect $C_1(\tau) = 1$ for perfect oscillations, because they have mean period τ. However, due to the stochastic nature of their period (see, e.g., Fig. 8.6) and because they set in and pause randomly, sometimes after an odd number of jumps, $C_1(\tau)$ is on average lowered with the occurrence of oscillatory events. When the timescales of noise-induced escapes and oscillatory motion become comparable, the particle dynamics responds resonantly to the noise. A slight increase of $\gamma D_0/V_0$ then significantly increases the number of occurring oscillation periods, yielding a strong reduction of $C_1(\tau)$. Thus, $\dot{Q}$ steeply increases and $\dot{S}_{\rm m}$ reaches high values [Eq. (8.39)]. At the CR maximum (vertical cyan lines in Figs. 8.9, 8.8), this

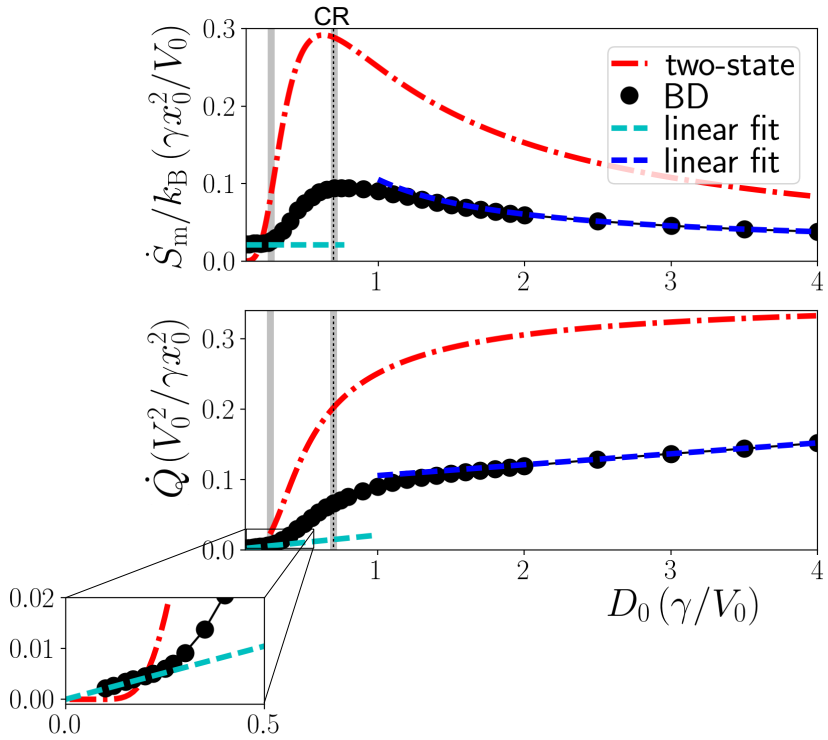

Figure 8.10: *Mean heat rate and medium EP vs. thermal energy over barrier height $\gamma D_0/V_0$, for other parameters than in Fig. 8.9, here $k = V_0$, $\tau = 5\,\gamma\sigma^2/V_0$. Again, the maximum of the medium EP is around coherence resonance (CR), which is at $\gamma D_0 \approx 0.7 V_0$ (dashed vertical lines). The 2-state model (red dashed-dotted line) correctly predicts the behavior of heat and medium EP, but overestimates them by a factor of three, as in the other parameter case. Moreover, the linear system yields again a good approximation for very small $\gamma D_0/V_0$ (cyan dashed line), where the intrawell dynamics dominate, as well as for the high $\gamma D_0/V_0$, where the particle does not see the barrier anymore (blue dashed line).*

effect saturates, resulting in a reduced slope of $\dot{Q}$ and a *maximum* of $\dot{S}_\mathrm{m}$ at CR.

For even higher γD_0, the superimposing noise generates irregular, low correlated motion with $C_1(\tau) \to 0$, hence $\dot{S}_\mathrm{m} \to 0$ in the 2-state model. In contrast, $\dot{S}_\mathrm{m}$ of the full system approaches a (nonzero) constant (not shown here). The breakdown of the 2-state approximation in this limit is indeed expected since the state discretization then becomes meaningless.

In the next Chapter, we will again present analytical results for the mean heat flow in non-Markovian systems. There, we will focus on linear systems with distributed delay and colored noise. But let us first briefly consider the distributions of the heat, work and internal energy fluctuations in the presence of delay, which we have, so far, not discussed.

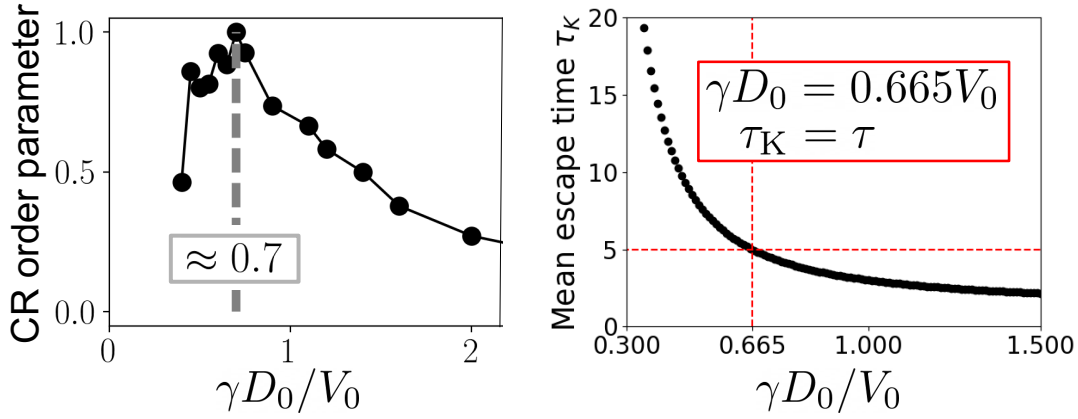

Figure 8.11: *Left panel: Coherence resonance (CR) order parameter, given by the (normalized) height of the first peak in the power spectrum of the temporal position-autocorrelation function, as explicitly shown in 8.7, where a different parameter setting is considered. The peak is highest around $\gamma D_0 \approx 0.7 D_0$, indicating highly regular oscillations with mean period τ. Right panel: Kramers escape times, in the absence of delay force. In the $\gamma D_0/D_0$ regime of CR, the escape times are comparable to the delay time. Thus, one can say that the noisy system has a resonant response to the delay force. Parameters as in Fig. 8.10.*

8.4 Preliminary numerical results for fluctuation of heat, work and internal energy

So far, we have considered the ensemble averaged rates of the energy flows associated with the delay force. However, as these energy flows are fluctuation quantities, they are not only characterized by their mean, but by an entire probability distribution, quantifying the chances of measuring a particular value along a random trajectory. A systematic study of the distributions in the presence of time-delayed feedback force is an objective of future work and beyond the scope of this thesis. But let us here nevertheless show some preliminary numerical results to give an impression of how they look like. As a nonlinear example system, we again consider the particle in the doublewell potential with and without the delayed feedback force.

First, we consider the equilibrium approached in the absence of delayed feedback, shown in the left panel of Fig. 8.12. All mean values are zero in this case, as the system is in thermal equilibrium (ensemble averages are plotted with vertical lines). Please note that we here consider parameters, where escape events are seldom. Furthermore, as there are no external nonconservative forces acting on the particle, there is no work performed on it, thus, w is strictly zero along each fluctuating trajectory. In contrast, the internal energy is a fluctuating quantity, owing to the fact that the particle can have a different internal (potential) energy e depending on its position within the static doublewell potential. The particle can increase its potential energy due to a kick from the surrounding heat bath, associated with an energy flow from the bath to the particle, i.e., a negative heat flow. On the other contrary, when the particle slides down the static potential loosing its internal

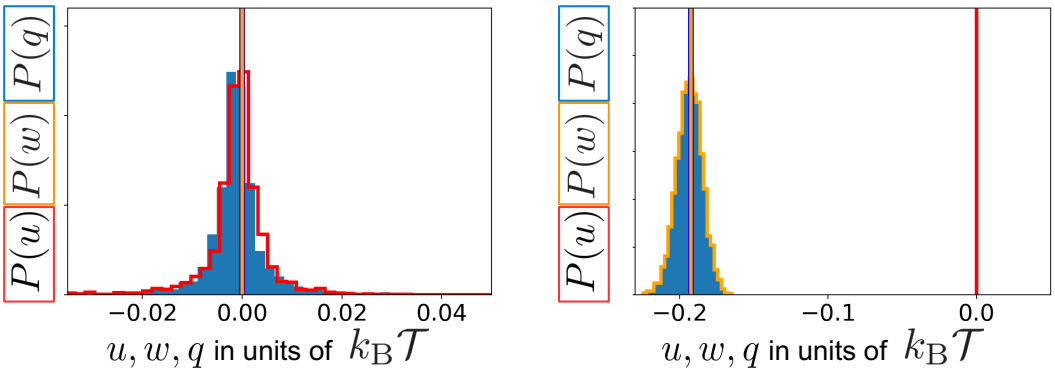

Figure 8.12: *Steady-state probability density of heat q (blue), work w (orange) and internal energy e (red) along fluctuating trajectories of length $\Delta t = 200\tau_B$ from numerical simulations with 5×10^4 realizations, in different situations. The averages are marked with vertical lines in the respective color. In steady states, one always has $\langle e \rangle = 0$. Left: Particle in doublewell potential (with barrier height $V_0 = 0.1 k_B \mathcal{T}$) in the absence of delayed feedback force. The system is in an equilibrium, thus $w \equiv 0$, and $\langle q \rangle = \langle w \rangle = 0$. Right: In the presence of delayed feedback (with $k = 0.1 k_B \mathcal{T}$, $\tau = 0.1\tau_B$), without external static potential, thus $e \equiv 0$. The system is in a NESS characterized by $\langle q \rangle = \langle w \rangle \neq 0$.*

(potential) energy, it automatically transfers energy to the bath via friction, i.e, a positive heat flow. Thus, the distributions of internal energy and the heat are identical mirror images from each other, as can be seen in Fig. 8.12 (left panel). $P(q)$ resembles the equilibrium distribution in a linear system, which was analytically and numerically investigated in [137], see Fig. 3.1. However, the bistable potential considered here, renders a slightly more peaked distribution than the second order Bessel function observed in the linear case [137].

In the right panel of Fig. 8.12, the system in the absence of the bistable static potential is shown, but with feedback force. Here, the linear delayed feedback force induces nonzero mean heat and work, with distributions that are symmetric with respect to their mean value. The distributions appear to be Gaussian-like, yet they are non-Gaussian. We could not find literature results for the distributions of the stochastic energies induced by delay forces to compare with. In Markovian systems, a linear external force (for example, a constant drag of the particle by constantly moving the harmonic trap), yields a Gaussian heat distribution. This was also shown in [137].

Finally, when combining the nonlinear static forces from the bistable potential, with the linear delayed feedback force, the distributions of all fluctuating energies are non-symmetric with a pronounced skewness, as shown in Fig. 8.13. The question whether the mean heat and work are positive or negative, depends on the precise values of k and τ, as we have discussed earlier in this Chapter, see for example Fig. 8.5. At first sight, we could not detect apparent differences in the distributions between cases with $\dot{Q} > 0$ and $\dot{Q} < 0$.

In the next Chapter, we will come back to fluctuations of thermodynamic quan-

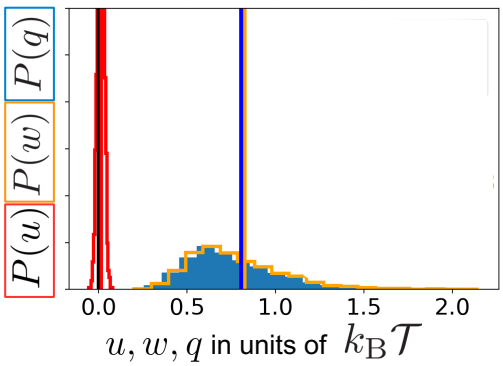

Figure 8.13: *Steady-state probability distributions of heat q (blue), work w (orange) and internal energy e (red), like in Fig. 8.12. Here, the doublewell potential and the delayed feedback are both present, thus, heat, work and internal energy are all fluctuating quantities. As always in a NESS: $\langle e \rangle = 0$ and $\langle q \rangle = \langle w \rangle \neq 0$. Parameters: $V_0 = k = 2k_{\mathrm{B}}\mathcal{T}$, and $\tau = \tau_{\mathrm{B}}$.*

tities, specifically, we will discuss the distribution of the total entropy for the fully linear system with feedback forces involving distributed delay. We will show that the entropy fulfills well-known fluctuation theorems, which impose universal symmetry conditions on the distributions. Finding corresponding statements for the distributions of stochastic work, heat and internal energy in the non-Markovian case considered here is an open problem, which has, to the best of our knowledge, not been tackled to date.

8.5 Concluding remarks

In this Chapter, we have investigated the steady-state heat flow and medium entropy production induced by time delay, in linear and nonlinear systems. We have seen that the presence of a delay force is almost always accompanied by net heat flow, except for rare parameter settings. The associated entropy production implies that the system operates far from equilibrium. The heat flow can be negative corresponding to energy extraction of the heat bath. We have further shown that the overdamped limit can be problematic, if τ gets below the velocity-relaxation timescale, and linked the behavior at small τ to the phenomenon of entropy pumping. Moreover, we found that the medium entropy production due to the feedback is maximal, when the delay-induced oscillations are coherence-resonant.

Thus, by investigating the mean heat flow and medium EP, we could already answer several of the initially posed questions. Such an energetic investigation is an important first step towards an understanding of thermodynamic notions in delayed systems. In the next Chapter, we will discuss the complete entropy balance. While this turns out the be a much harder task, it will, at the same time, provide a more profound physical picture.

9 | Entropy, Information & Energy flows

"[W]e must return to Maxwell's original concern about the second law and try to address the basic problems of statistical mechanics, such as the emergence of the macroscopic world and the subjectivity of entropy, in the light of a general physical theory of information."

– Juan Parrondo, 2015 [131]

In the last Chapter, we have studied one component of the steady-state entropy balance, the medium entropy production and associated heat flow. While this investigation has already provided interesting insights, open questions remain. First, we observed that the mean heat rate can vanish at some specific, rare parameter values, despite the presence of time-delayed feedback. On the basis of the energetic consideration alone, we cannot *clarify whether these points correspond to equilibrium.*
Second, we have seen that the medium EP can be negative. We recall that this would not be possible in a Markovian system with position-dependent driving, where the medium EP is the only entropic contribution and must be non-negative, as stated by the second law. The negative medium EP observed in the case of time-delayed position-dependent driving, thus, indicates *missing entropic terms.* While the amount of negative medium EP provides a lower bound to these missing entropy contributions, we could, so far, not get any further physical insights about what these terms look like and how they arise. For this purpose, we shall investigate the whole entropy balance.

Due to the presence of delay in our equations, we run into a major problem of stochastic thermodynamics, that is, the notion of entropy production of non-Markovian systems, which is connected to the role of hidden degrees of freedom and coarse-graining. We have introduced this problem in Chapter 3. We recall the technical problem associated with the acausality of the backward process, and we recall the different attempts to formulate an effective thermodynamic description, which yet comprises ambiguities. We will not focus on this debate, but rather go one step back, and reconsider the first step of the whole problem.

Main goal and general idea In this last Chapter 9, we will investigate a specific way to address the problem. The embedding technique introduced in Part II, yielded a *Markovian* network description of the delayed process. Let us recall the equations (5.1),

$$\gamma \dot{X}_0(t) = \quad -kX_n(t) + F_0[X_0] \quad + \sqrt{2k_{\mathrm{B}}\mathcal{T}_0\gamma}\,\xi_0(t) \tag{9.1a}$$

$$\gamma' \dot{X}_j(t) = \; (n/\tau)[X_{j-1}(t) - X_j(t)] \; + \sqrt{2k_{\mathrm{B}}\mathcal{T}'\gamma'}\,\xi_j(t), \tag{9.1b}$$

with $j \in \{1, 2, ..., n\}$. This representation allows application of the standard formulae of stochastic thermodynamics. In this Chapter, we will investigate whether the thermodynamic picture obtained by studying the total EP of the network (9.1) (i.e., of the "super-system"), instead of the delay LE, *is meaningful and interesting.*

We will obtain entropic contributions of the $X_{j>0}$, which are so far only abstract auxiliary variables. On this abstract level, these entropy terms can be considered *thermodynamic cost of the memory* [93, 116, 117, 152]. However, the physical meaning of these terms is subject of an ongoing debate. To give an example, we will suggest an interpretation for the $X_{j>0}$ for a specific realization of a feedback controller. From the viewpoint of effective thermodynamics, our investigation represents a first step, which the effective thermodynamics could be build upon.

As mentioned above, one view is to consider the $X_{j>0}$ variables as degrees of freedom of the controller (i.e., the controller is now included in the model, contrary to the considerations in the rest of this thesis). At this point, one should stress that our approach is different from studies, where the memory of a feedback-controller is realized with a *tape* [130, 131, 289, 292, 333–335]. In particular, in the interpretation presented here, all degrees of freedom are "physical", meaning that the forces they perform are associated with energy cost. To further differentiate from earlier studies, we note that the description employed here is a space-continuous, stochastic one, which contrasts earlier bipartite-models basing on Master equations, like [292, 336, 337].

There is yet another way to tell the following story. As a second application of the investigation presented in this Chapter serve systems of type (9.1), where all d.o.f. X_j are physical subsystems. In such a situation, the total EP calculated here has a clear meaning, it is the true thermodynamic dissipation and irreversibility measure.

The network (9.1) has unidirectional coupling, which is a specific type of *non-reciprocal* interactions. Such interactions are very uncommon in statistical physics and thermodynamics, as mechanical systems describable by Hamiltonians are always reciprocally coupled (usually the underlying equations of motion fulfill Newton's laws). Thus, a significant part of this Chapter is dedicated to the discussion of NR interactions, which have not been studied in earlier literature from a thermodynamic perspective. We will show that non-reciprocity has various interesting implications.

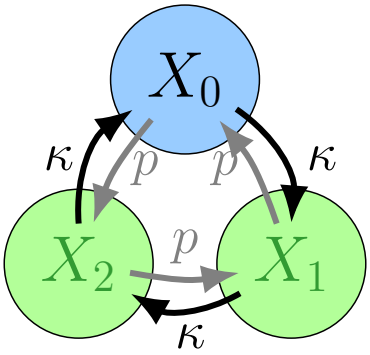

Figure 9.1: *All-to-all coupled network of three nodes. For $\kappa = p$, the network corresponds to a model for three colloids reciprocally coupled by springs of stiffness κ. For $p \to 0$, the network is unidirectionally coupled.*

Contrary to the rest of this thesis, we will here not only consider feedback with discrete delay, and LEs with white noise, but we will also consider more general cases with colored noise and memory kernels, which can be obtained, e.g., by keeping n finite in the description (see Sec. 5.1). The colored noise will be linked to *measurement errors* [289, 338, 339] of the controller, which we have also not considered so far. Stepping away from the concrete example of an optical tweezers feedback controller, distributed delay is indeed the more realistic scenario for many delayed processes, e.g., in biological contexts [9, 57]. Furthermore, we will specify all deterministic forces to be *linear*, i.e., $F \propto X_0$. This is because we are interested in the main idea, and further aim to discuss the concepts based on exact results. However, extensions to the nonlinear models considered earlier in this thesis would in principle be possible (at least numerically).

9.1 Emergence of non-monotonic memory

An important property of the time-delayed systems is that the memory is sharply peaked around a particular value of the past, i.e., $t-\tau$. This implies a non-monotonic memory kernel. Stepping away from the (idealized) case of a delta-peaked kernel, i.e, discrete delay, a non-monotonic kernel with a pronounced maximum seems a reasonable model.

Here we will investigate the connection between topological properties of the Markovian network representation and the non-monotonicity of the kernel in the non-Markovian projected equation, by using a simple example network with three d.o.f.. In Sec. 5.1, we have already seen that the Markovian embedding of the time-delayed process yielded NR interactions. However, let us investigate here, whether we really need the non-reciprocity, or if we can obtain a similar kernel with a symmetrically coupled network.

Let us consider as an illustrative example an all-to-all coupled network of three nodes X_j with $j \in \{0, 1, 2\}$, described by

$$\begin{aligned}
\dot{X}_0 &= -(p+\kappa)X_0 + pX_1 - \kappa X_2 + \xi_0, \\
\dot{X}_1 &= -\kappa X_0 - (p+\kappa)X_1 + pX_2 + \xi_1, \\
\dot{X}_2 &= +pX_0 - \kappa X_1 - (p+\kappa)X_2 + \xi_2,
\end{aligned} \tag{9.2}$$

as schematically shown in Fig 9.1. This network can also be written as

$$\dot{X}_j = -(p+\kappa)X_j + pX_{j+1} - \kappa X_{j-1} + \xi_j, \tag{9.3}$$

with the periodic index $j \in \{0, 1, 2\}$.

Projection Analogously to the projection in Chapter 5, we project this set of equations onto X_0. To this end, we perform a Laplace-transformation and iteratively solve the algebraic equations for X_0. In this way, we immediately obtain

$$\begin{aligned} s\hat{X}_0 =& p\left[\frac{p^2}{(s+p+\kappa)^2 - p\kappa} + \frac{\kappa(s+p+\kappa)}{(s+p+\kappa)^2 - p\kappa}\right]\hat{X}_0 \\ &+ \kappa\left[\frac{\kappa^2}{(s+p+\kappa)^2 - p\kappa} + \frac{p(s+p+\kappa)}{(s+p+\kappa)^2 - p\kappa}\right]\hat{X}_0 - \kappa\hat{X}_0 - p\hat{X}_0 + \mathcal{O}(\hat{\xi}_j). \end{aligned} \tag{9.4}$$

Transforming back to real space yields a non-Markovian description of the dynamics of X_0, in particular,

$$\dot{X}_0(t) = -pX_0(t) - \kappa X_0(t) + \int_0^t K(t-t')X_0(t')\mathrm{d}t' + \xi_0(t) + \nu(t), \tag{9.5}$$

with a colored noise ν, and the memory kernel

$$\begin{aligned} K(t-t') =& \frac{e^{\left(-\sqrt{p\kappa}+p+\kappa\right)t' - t\left(\sqrt{p\kappa}+p+q\right)}}{2\sqrt{p\kappa}} \times \\ & \left[\left(p^{3/2}+\kappa^{3/2}\right)^2 e^{2\sqrt{p\kappa}t} - \left(p^{3/2}-\kappa^{3/2}\right)^2 e^{2\sqrt{p\kappa}t'}\right]. \end{aligned} \tag{9.6}$$

In the same manner, we have derived the corresponding equation for a network, where all couplings are different, e.g., X_0 is coupled to X_1 via κ_{01}, and back via p_{01} and so on. However, because the results have the same overall characteristics, while yielding very cumbersome expressions; we omit to discuss this case here, and focus on this simpler case. Let us now consider the memory kernel (9.6) generated by the network (9.2) in more detail, for different cases.

Reciprocal coupling For $\kappa = p$, the network has *reciprocal* coupling. Then, it can be considered as a model for three overdamped particles connected by harmonic springs with spring constant κ, which has the corresponding Hamiltonian $H = (\kappa/2)\sum_{i\neq j}(X_i - X_j)^2$. Similar models have been considered from a thermodynamic perspective in, e.g., [117, 124].

In this case, the memory kernel (9.6) simplifies to an exponential decay,

$$K(T) = 2\kappa^2 e^{-\kappa T}. \tag{9.7}$$

We observe the same qualitative behavior for larger networks with $n > 2$, or more general coupling topologies.

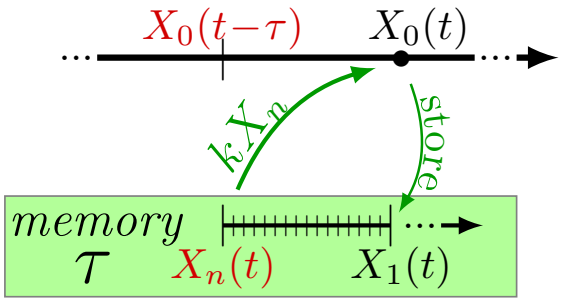

Figure 9.2: *Relation between the temporal evolutions of the "system" X_0 and the n memory cells $X_{j>0}$ of the controller.*

Non-reciprocal coupling In any other case, $\kappa \neq p$, the nodes have NR interactions. In our model, the coupling is then stronger in one direction of the ring compared to the other direction. Considering (9.6), we observe that the memory automatically becomes *non-monotonic*. Specifically, a maximum at a finite $t - t'$ emerges. Now, let us consider an extreme case, where the equation simplifies.

In the limit $p \to 0$, the network is unidirectionally coupled. In this limit, it in fact becomes a specific realization of our embedded model (9.1) introduced in Sec. 5.1 (specifically, see the inset of Fig. 5.2). In this case, the memory kernel (9.6) converges to

$$K(T) = \kappa^3 T e^{-\kappa T}, \tag{9.8}$$

which is a Gamma-distribution with a maximum near $\kappa/3$. We have displayed such a kernel for $\kappa = 3$ in Fig. 5.2 (b).

Playing around with different coupling topologies, we generally find that reciprocal couplings yield monotonic memory kernels. We thus conclude that NR coupling is a crucial ingredient to generate non-monotonic memory[1]. However, we aim to note that we have not tried out nonlinear couplings, or nodes with nonlinear "local" dynamics.

In the following, we will consider our embedded model (9.1), which is essentially a generalization of the unidirectional ring network to arbitrary lengths. Before we proceed with exploring the thermodynamic properties, we discuss a possible interpretation of the d.o.f. $X_{j>0}$ in the case of the feedback controller.

9.1.1 Interpretation of $X_{j>0}$ in the case of a feedback controller

Let us take a closer look at our prime example of non-Markovian systems considered throughout the thesis, i.e., feedback control of colloidal particles via optical traps.

In experimental setups, the delay is typically created within a memory device, e.g., the computer's memory, storing the past trajectory of length τ. To model the memory device, we step away from an actual computer (which would be hopelessly complicated to model and treat analytically). In a simplistic manner, the additional

[1]This observation seems reasonable. A non-monotonic memory indicates that the information is passing through the network as a "package". To realize this, one somehow has to break the symmetry of the dynamics.

d.o.f. in the network (9.1) may be interpreted as the *memory cells* of the controller. Let us explain this a bit better.

Imagine a memory device with a "shift register" type of architecture, as sketched in Fig. 9.2. At every instant in time, the information stored in each cell X_j is shifted one step further from j to $j+1$, while the new information (gathered by a measurement operation) is stored in the first cell X_1. Thus X_n stores the oldest information and is used to perform the feedback control force kX_n. Mathematically, this operation can be represented in discretized form as $\Delta X_j = \frac{\Delta t}{\tau/n}[X_{j-1} - X_j] + \epsilon_j$. The ϵ_j are random numbers accounting for errors, e.g., from measurement, rounding operations, or a finite temperature of the device. The updates of the memory cells are scaled with τ/n to enable storage of a trajectory of length τ in n cells.

Since we aim to model *continuously operating* feedback loops, we assume that the updates are performed (infinitely) fast, that is, $\Delta t \to 0$, yielding the dynamical equations of X_j (9.1b). For the sake of simplicity, we have set all $X_{j>0}$ to the same temperature, $\mathcal{T}'$. In the present context this implies the magnitude of error is identical in each memory cell.

We indeed found a similar consideration in [289]. As suggested there, such a memory device could be realized by n colloidal particles at positions X_j, each trapped in a harmonic trap of stiffness n/τ centered around the position of the proceeding colloid, X_{j-1}. This idea is, of course, rather an idealized thought experiment. Still, it is useful to illustrate how a memory kernel concentrated around a specific instant in time of the past (the delay time), can result from coupled exponential relaxation processes. It further demonstrates that an implementation of (non-monotonic) memory, with physical d.o.f. is inevitably associated with thermodynamic cost and dissipation, which we will discuss in the following[2].

Complex bio-chemical systems as controller Turning away from artificial systems or computer models described above, and towards actual physical, biological, chemical systems, such a "memory device" could for example also be a cascade of chemical reactions.

In fact, in [323] a model for a *cellular sensor* (corresponding to X_1) with one "memory cell" (i.e., X_2) is proposed, which exactly corresponds to our unidirectionally coupled model (9.1) with $n = 2$ and $k = 0$ (i.e., no backcoupling from X_2 to X_0, since [323] doesn't consider feedback). The focus of Ref. [323] is to calculate information flows for $n = 1, 2$. We will as well consider our systems from that perspective below, but mostly focus on other thermodynamic quantities, particularly the heat and total EP, and study how they depend on the number of d.o.f., n.

The apparent connection between our feedback controller and this cellular sensor model, which is an active system, motivates us to elucidate the relation between memory, activity and feedback from a fundamental perspective. First, we will need to clarify the formal connection. This is the objective of the following Sec. 9.2.

[2]Interestingly, this thought experiment also illustrates that the friction terms $-\gamma'\dot{X}_i$ may be crucial to memorize, because they enable "forgetting" the old information.

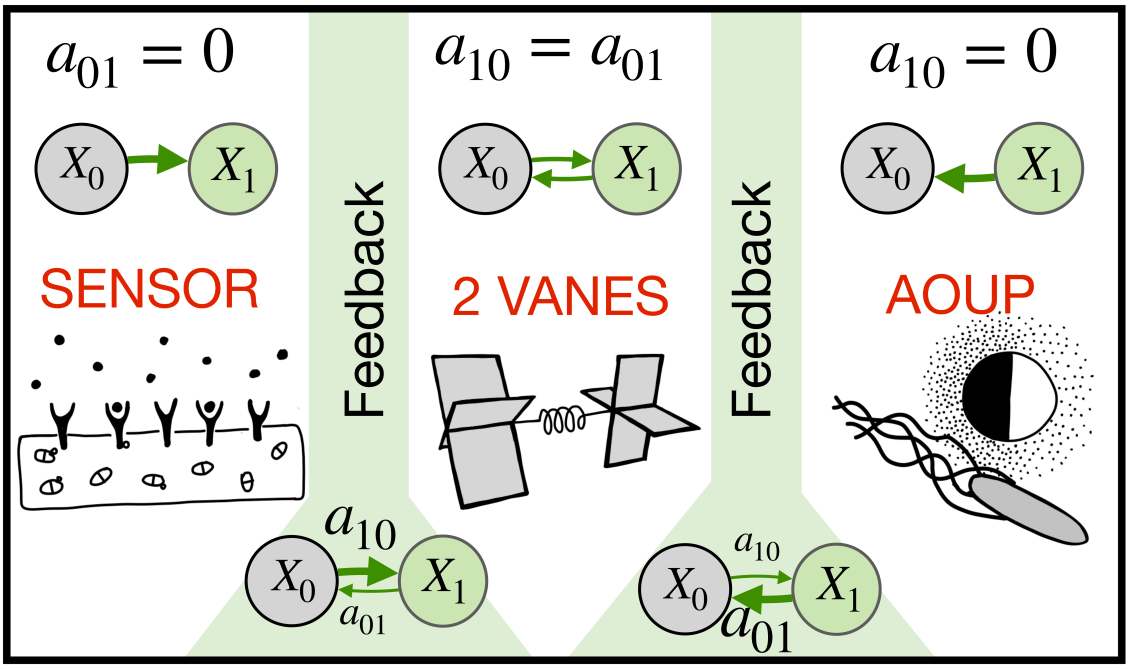

Figure 9.3: *Overview over different examples included in the generic network model (9.9) at $n = 1$ (with increasing ratio a_{01}/a_{10} from left to right). Left: If X_1 is unidirectionally coupled to the system d.o.f. X_0, i.e., $a_{01} = 0$, such that only X_1 "sees" X_0, but not vice versa, X_1 corresponds to a cellular sensor in the model [323]. Center: For reciprocal coupling $a_{01} = a_{10}$, X_1 represents a mechanically coupled d.o.f., and X_0, X_1 correspond to the angles of two vanes coupled by a torsion spring [124]. Right: If X_0 is unidirectionally coupled to X_1, i.e., $a_{10} = 0$, X_0 corresponds to the position of a microswimmer within the AOUP model, where X_1 models the effect of the flagella of a bacterium or the asymmetric flow field around a colloid. In the intermediate cases, where the coupling is bidirectional but non-reciprocal, X_1 corresponds to the controller d.o.f. and X_0 to the position of the feedback-controlled colloid in our model with exponentially distributed delay, and colored noise.*

9.2 The role of non-reciprocal coupling – Connection to active matter

In Sec. 3.7.2 we have mentioned some general links between active matter and colloids subject to time-delayed feedback. Here, we resume this discussion, and bring it to a formal level.

9.2.1 A generic model with non-reciprocal coupling

As we have already mentioned in Section 3.7, Langevin-like equations are currently used by many researchers to describe the motion of active systems. In particular, we have introduced a very simple model for a microswimmer, the active Ornstein-Uhlenbeck particle, see Sec. 3.7.1. Like the cellular sensor, the AOUP model as well has a two-dimensional network representation, which also involves unidirectionally

coupling. In fact, all three examples can be considered as special cases of the model

$$\gamma_j \dot{X}_j = a_{jj} X_j + a_{jj-1} X_{j-1} + \xi_j, \tag{9.9}$$

with periodic index $j \in \{0, 1, ..., n\}$, or with the corresponding projected equation

$$\dot{X}_0(t) = a_{00} X_0(t) + \int_0^t K(t-t') X_0(t') \mathrm{d}t' + \xi_0(t) + \nu(t). \tag{9.10}$$

ν is a colored noise, while ξ_j are Gaussian white noises with $\langle \xi_i(t) \xi_j(t') \rangle = 2k_\mathrm{B} \mathcal{T}_j \gamma_j \, \delta_{ij} \delta(t-t')$. Note that the temperatures $\mathcal{T}_j$ are now scaled into the variance of the noise, for the sake of a shorter notation. Let us now make the connection with the aforementioned examples explicit.

Feedback controller with time delay First of all, our original feedback control model (9.1) is recovered as a special case at $a_{0n} = -k$, $a_{ii-1} = -a_{ii} = \gamma' n/\tau$, $\mathcal{T}_{j>0} = \mathcal{T}'$, $\gamma_{j>0} = \gamma'$. The time-delayed part of the force can be written in two ways as $-kX_n$ or, equivalently, as $-k \int_0^t K(t-t') X_0(t') \mathrm{d}t'$, corresponding to the two representations. n can be any natural number, giving the number of "memory cells", and yielding monotonically decaying colored noise and a Gamma-distributed memory kernel given in (5.3, 5.4), recall

$$C_\nu(T) = \langle \nu(t) \nu(t+T) \rangle = k^2 \frac{k_\mathrm{B} \mathcal{T}'}{\gamma'} \sum_{p=0}^{n-1} \sum_{l=0}^{p} \frac{2^{l-2p} (2p-l)!}{p!(p-l)! \; l!} \frac{e^{-nT/\tau} T^l}{(\tau/n)^{l-1}},$$
$$K(T) = \frac{n^n \, T^{n-1}}{\tau^n (n-1)!} e^{-\frac{nT}{\tau}}. \tag{9.11}$$

Thus, n decides on the functional form of the distributed delay. The discrete delay, white noise LE is recovered in the limit $n \to \infty$, as we have explicitly shown in Sec. 5.1.2.

Cellular sensor, AOUP, feedback controller with exponential delay At $n = 1$, the cellular sensor model [323] is obtained for the unidirectional coupling with $a_{01} = 0$, $a_{10} > 0$, while the AOUP model[3] corresponds to reversed unidirectional coupling $a_{10} = 0$, $a_{01} > 0$. The corresponding projected equation has exponentially decaying colored noise and memory kernel

$$C_\nu(T) := \langle \nu(t) \nu(t+T) \rangle = k_\mathrm{B} \mathcal{T}_1 (a_{01}^2 / a_{11}) e^{a_{11} T/\gamma_1},$$
$$K(T) = (a_{10}/\gamma_1) e^{a_{11} T/\gamma_1}. \tag{9.12}$$

At *bi*directional, non-reciprocal coupling $a_{01} = k$, $a_{10} = -a_{11} = \gamma' n/\tau$, this model with $n = 1$, also reproduces the smallest version of our feedback controller model, i.e., where the entire memory device has one cell, or in other words, where the controller has one d.o.f. X_j (and the memory kernel is exponential). This case is readily included in (9.12).

[3] In comparison with the notation in Eq. (3.37), here $\gamma_1 = 1/\tau_\mathrm{P}$, $2k_\mathrm{B} \mathcal{T}' \gamma_1 = 1$.

Reciprocal coupling – Mechanical models Last, in the specific case of reciprocal coupling, i.e., $a_{ij} = a_{ji} \forall i, j \in \{0, 1, ..., n\}$, the network may model mechanical systems. For example, the $n = 1$ network

$$\begin{aligned} \gamma_0 \dot{X}_0 &= -\kappa X_0 + \kappa X_1 + \xi_0, \\ \gamma_1 \dot{X}_1 &= +\kappa X_0 - \kappa X_1 + \xi_j, \end{aligned} \tag{9.13}$$

with $\mathcal{T}_0 \neq \mathcal{T}_1$ corresponds to a minimal model for *heat conducting* through mechanical motion, discussed in [124] (see p. 154). Here, the two degrees of freedom, $X_{0,1}$ could model the *angles of two vanes* that rotate in two different heat baths, and are coupled by a torsion spring with spring constant κ. More generally, (9.9) corresponds to a system of n colloidal particles trapped in quadratic potentials where the colloids are further coupled among each other by harmonic springs. We have already considered such a model for $n = 2$ in Sec. 9.1.

To summarize, depending on the symmetry properties of the coupling topology, the generic network (9.9) can be used to model various different systems. Specifically, the $n = 1$ case includes a model for a cellular sensor, a minimal version of a time-delayed feedback controller (with exponential delay), a mechanical system of two reciprocally coupled vanes, and the motion of an active swimmer. Figure 9.3 provides an overview over the different examples, and shows the corresponding coupling topologies. Importantly, the coupling is non-reciprocal in all these cases but the mechanical two-vane model. This motivates us to take a closer look at the *thermodynamic implications of NR coupling*. As the case $n = 1$ already includes very different, intriguing example systems, and further represents the smallest version of our model, with a single non-reciprocal link, we will focus on the generic model (9.9) at $n = 1$. At the end, we will also consider higher values of n to come back to the original example of the feedback controller with time delay.

Interpretation of non-reciprocal coupling To avoid confusion, we aim to emphasize the following, concerning the meaning of the NR interactions in (9.9). First, we would like to remind the Reader that the LE is a coarse-grained, i.e., effective equation, and the dynamical variable is a microscopic one. Fundamentally, physical interactions are typically reciprocal and symmetric, including mechanical coupling, as states Newton's third law. However, in complex systems, involving chemical reactions, multiscale processes, or coupling between a chemical component and a mechanical d.o.f., much more complicated relationships between subsystems can emerge [110, 155, 340, 341], which might be representable as coupling between few mesoscopic, stochastic d.o.f.. This is the situation considered here.

9.3 Non-reciprocal coupling & Non-equilibrium

First we demonstrate that NR coupling alone "drives" the system out of equilibrium. To this end, we consider two typical measures, first the fluctuation-dissipation relation, then the total entropy production.

9.3.1 Fluctuation-Dissipation relation

We recall the FDR (or Kubo relation of second kind) [95]

$$\langle \mu(t)\mu(s)\rangle = k_{\rm B}\mathcal{T}_0\,\gamma(|t-s|),$$

which we have introduced earlier in Part I. It describes a balance between friction kernel γ and thermal noise μ [note that contrary to Eq. (1.35), we have here used μ instead of ν to differentiate the different noises]. As well-known for e.g., viscoelastic-fluids, the validity of a FDR implies that the system will equilibrate in the absence of external driving [15, 95].

To check the FDR for the present model, we rewrite (9.10) in the form of a generalized LE by converting K via partial integration into a friction-like kernel. This yields

$$\int_0^t \gamma(|t-s|)\dot{X}_0(s)\,{\rm d}s = a_{00}X_0 + \widetilde{K}(0)X_0(t) + \mu(t), \tag{9.14}$$

involving $\mu(t) = \xi_0(t) + \nu(t)$ and $\gamma(t-s) = 2\gamma_0\,\delta(t-s) + \widetilde{K}(t-s)$, with the integrated memory kernel $\widetilde{K}(t-s)$. In this way, an equation with a memory kernel in an integral over $X(t)$ can be transformed in a dynamical equation with a *time-nonlocal friction force*. The friction kernel then involves the integrated memory kernel $\widetilde{K}$. For the generic network (9.9) at $n=1$, the latter reads

$$\widetilde{K}(T) = (a_{01}a_{10}/a_{11})e^{a_{11}T/\gamma_1}\gamma(t-s) = 2\gamma_0\,\delta(t-s) + \widetilde{K}(t-s). \tag{9.15}$$

By construction, the instantaneous friction $-\gamma_0\dot{X}$ and the white noise ξ_0 readily balance each other. Further, comparing the additional "friction" kernel (9.15) and the correlations of the colored noise ν given in (9.12), it can easily be seen that the FDR holds if

$$a_{01}\mathcal{T}_1 = a_{10}\mathcal{T}_0. \tag{9.16}$$

Remarkably, (9.16) suggests that the system with NR coupling can reach equilibrium despite $\mathcal{T}_0 \neq \mathcal{T}_1$. This is in sharp contrast to reciprocally coupled ("passive") systems, which generally never equilibrate in the presence of temperature gradients. For NR couplings, (9.16) can only be true when both coupling strengths are nonzero and have the same sign, i.e., $a_{01}a_{10} > 0$. We will later in Sec. 9.4 show that this case corresponds to a "mild" form of symmetry-breaking.

For the feedback controller model with $n > 1$ corresponding to (9.11), the integrated kernel reads $\widetilde{K}(T) = k\,\Gamma\left(n, nT/\tau\right)/(n-1)!$ with the upper incomplete Gamma function, and the FDR is violated. An explicit proof is given in the Appendix A.5.

9.3.2 Total entropy production

Now we turn to the total entropy production, which gives a more fundamental notion of non-equilibrium. While the broken FDR indicates non-equilibrium, it does not

quantify the distance from equilibrium. In contrast, the total EP directly quantifies the irreversibility, i.e., time-reversal symmetry breaking.

However, as we have already argued in Chapter 3, the calculation of the EP for a non-Markovian process (9.10) is not straightforward [93, 116, 118, 152]. We recall that there is no closed Fokker-Planck equation to (9.10), and no standard path integral formalism, such that both standard routes [125] are not readily applicable. We will instead exploit from now on the Markovian version of our model (9.9), which allows application of the standard framework. We have introduced the formulae in Chapter 3, particularly, in Eq. (3.15).

We recall that if all X_j of the network are physical d.o.f., this is the usual total EP as traditionally defined in thermodynamics. For cases where the variables $X_{j>0}$ are auxiliary variables without clear physical meaning, the interpretation of the EP calculated in this way, is indeed not entirely clear. It comprises open questions and leaves room for discussion. At the end of this Chapter, we will comment on this issue. But let us first explore, how much we can learn from this quantity about the considered processes.

Specifically, we apply the standard definition (3.15), to calculate the total EP along a fluctuating trajectory $\underline{\mathbf{X}} = \{X_0(t'), .., X_n(t')\}$, $t' \in [t_s, t_f]$ belonging to the network (9.9). Recall

$$\frac{\Delta s_{\rm tot}[\underline{\mathbf{X}}]}{k_{\rm B}} = \ln \frac{\mathcal{P}[\underline{\mathbf{X}}]}{\hat{\mathcal{P}}[\underline{\hat{\mathbf{X}}}]} = \frac{\Delta s_{\rm sh}}{k_{\rm B}} + \ln \frac{\mathcal{P}[\underline{\mathbf{X}}|\underline{x}_s]}{\hat{\mathcal{P}}[\underline{\hat{\mathbf{X}}}|\underline{x}_f]}.$$

Due to the $(n+1)-$dimensionality of the network, the total EP involves the *multivariate (joint) Shannon entropy* $s_{\rm sh} = -k_{\rm B} \ln[\rho_{n+1}(\underline{x})]$ of the $(n+1)$-point joint PDF ρ_{n+1}.

Just as in the one-dimensional case, the total EP, by construction, fulfills an integral fluctuation theorem, $\langle e^{-\Delta s_{\rm tot}/k_{\rm B}} \rangle = 1$. In Sec. 3.4 we have explicitly discussed this theorem, and shown that it implies the second law of thermodynamics $\Delta S_{\rm tot} \geq 0$, i.e., non-negativity of the mean total EP.

Furthermore, the path probabilities conditioned on the starting point $\underline{X}(t_s) = \underline{x}_s$, readily follow from the corresponding LEs, as we have given in Eqs. (1.27, 3.18). Specified to the present case, they read

$$\begin{aligned} \mathcal{P}[\underline{\mathbf{X}}|\underline{x}_s] &\propto \quad e^{-\sum_{j=0}^{n} \int_{t_s}^{t_f} \frac{\left(\gamma_j \dot{X}_j(t') - F_j[\underline{X}(t')]\right)^2}{4k_{\rm B}\mathcal{T}_j\gamma_j} {\rm d}t'}, \\ \hat{\mathcal{P}}[\underline{\hat{\mathbf{X}}}|\underline{x}_f] &\propto \quad e^{-\sum_{j=0}^{n} \int_{t_s}^{t_f} \frac{\left(-\gamma_j \dot{X}_j(t') - F_j[\underline{X}(t')]\right)^2}{4k_{\rm B}\mathcal{T}_j\gamma_j} {\rm d}t'}, \end{aligned} \tag{9.17}$$

with $F_{j>0} = a_{jj}X_j + a_{j-1j}X_{j-1}$, and $F_0 = a_{00}X_0 + a_{0n}X_n$. Importantly, we here assume that $X_{j>0}$ are *even* under time-reversal, like positions, angles, or orientations. This is more suitable for the examples considered throughout the thesis (e.g., when $X_{j>0}$ model memory cells of a feedback-controller as we will argue below, odd parity would imply information storage in particle velocities, which seems less appropriate). However it should be noted that in some cases the parity of $X_{j>0}$ is indeed a nontrivial issue and subject of an ongoing debate, see e. g., [105] for active

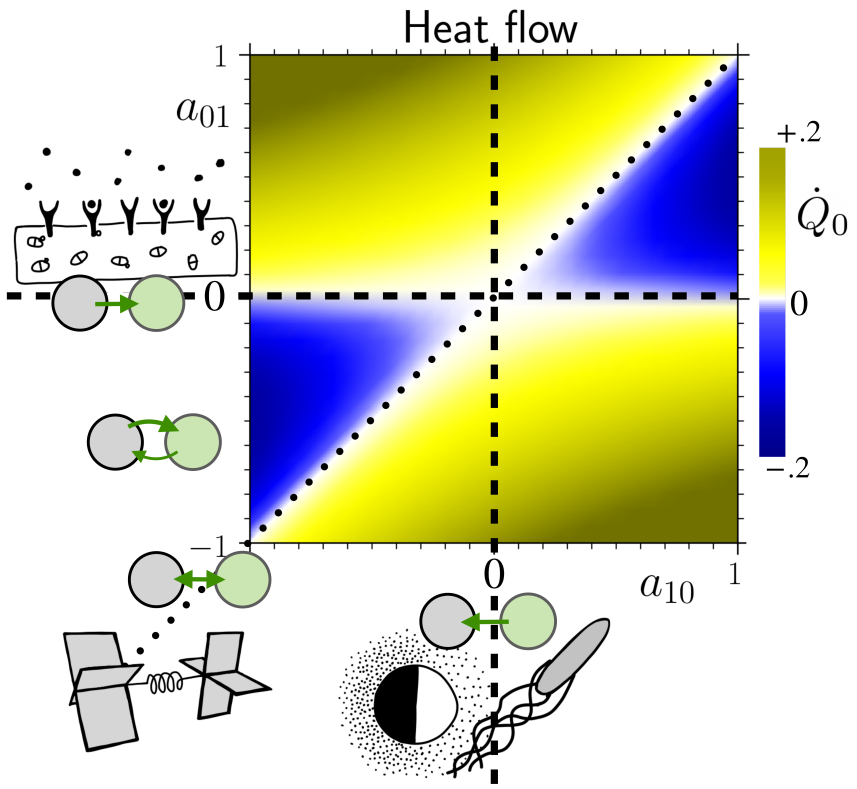

Figure 9.4: *Steady-state mean heat flow $\dot{Q}_0$ in units of $k_\mathrm{B}\mathcal{T}/\tau_\mathrm{B}$ from Eq. (9.22) in the minimal model (9.12) with $a_{11} = a_{00} = -\gamma_0/\tau_\mathrm{B}$, at isothermal conditions $\mathcal{T}_0 = \mathcal{T}_1$, and $\gamma_0 = \gamma_1$. $\dot{Q}_0$ vanishes along the diagonal $a_{01} = a_{10}$, where the system is reciprocally coupled and, thus, "passive" (in this case, $X_{0,1}$ could, e.g., model the angles of vanes coupled by a torsion spring [124]). $\dot{Q}_0$ further vanishes on the unidirectional coupling line $a_{01} = 0$, where X_0 is "sensed" by X_1 (corresponding to a cellular sensor [323]). $\dot{Q}_0 > 0$ along the other unidirectional coupling line $a_{10} = 0$ (where X_0 corresponds to the position of an active swimmer in the AOUP model [105, 107, 109, 307, 308], X_1 modeling the effect of a flagellum of a bacterium, or the asymmetric flow field around a Janus particle). All other parameters are set to unity.*

swimmers. This assumption implies

$$\ln \frac{\mathcal{P}[\underline{\mathbf{X}}|\underline{x}_s]}{\hat{\mathcal{P}}[\hat{\underline{\mathbf{X}}}|\underline{x}_f]} = \int_{t_s}^{t_f} \frac{\delta q_0}{\mathcal{T}_0} + \sum_{j=1}^{n} \int_{t_s}^{t_f} \frac{F_j \circ \mathrm{d}X_j}{\mathcal{T}_j}, \tag{9.18}$$

where δq_0 can be identified as the heat flow from X_0 to its bath [124], which reads $\delta q_0 = [a_{00}X_0 + a_{0n}X_n] \circ \mathrm{d}X_0$. In the steady state, $\Delta S_\mathrm{sh} = 0$, and the ensemble average of the EP is thus

$$\begin{aligned}\dot{S}_\mathrm{tot} =& \frac{a_{0n}^2\langle X_n^2\rangle + a_{0n}\,a_{00}\langle X_0 X_n\rangle}{\mathcal{T}_0} \\ &+ \sum_{j=1}^{n} \frac{a_{jj-1}^2\langle X_{j-1}^2\rangle + a_{jj}a_{jj-1}\langle X_{j-1}X_j\rangle}{\mathcal{T}_j},\end{aligned} \tag{9.19}$$

where we have used that $X_j\dot{X}_j$-correlations vanish in the steady state, since $2\langle X_i\dot{X}_i\rangle = \mathrm{d}\langle X_i^2\rangle/\mathrm{d}t = 0$. The mean EP (9.19) has two contributions, that is, the dissipation of X_0

$$\dot{Q}_0 = a_{0n}^2\langle X_n^2\rangle + a_{0n}\, a_{00}\langle X_0 X_n\rangle, \tag{9.20}$$

and a contribution from the additional d.o.f. $X_{j>0}$.

From the perspective of a non-Markovian process (9.10), the second contribution can be considered as the *entropic cost* of the memory. At this point it should be emphasized, however, that memory is *not* unequivocally connected to EP. For example, the reciprocal case $a_{01} = a_{10}$ of our network (9.12) also generates memory in X_0, but no net dissipation, thus no EP (we will explicitly show this below). We further remind the Reader that there are, of course, real-world examples of physical systems which involve complex memory but nevertheless reach thermal equilibrium (where the mean EP naturally vanishes), e.g., viscoelastic fluids. As we will later see, such *entropic cost*-free memory is, however, *not useful* in terms of extracting net information or for achieving goals of feedback control.

From (9.19, 9.20), analytical expressions for $\dot{Q}_0$ and the total EP can be derived for any n, by means of Green's functions. We will show this below in Sec. 9.3.3.

Let us here already give the results for the case $n = 1$ (which was also discussed in [118]),

$$\dot{S}_{\mathrm{tot}} = k_{\mathrm{B}}\frac{(a_{10}\mathcal{T}_0 - a_{01}\mathcal{T}_1)^2}{\mathcal{T}_0\mathcal{T}_1(-a_{00}/\gamma_0 - a_{11}/\gamma_1)}, \tag{9.21}$$

$$\dot{Q}_0 = \frac{a_{01}(a_{10}\mathcal{T}_0 - a_{01}\mathcal{T}_1)}{(a_{00} + a_{11}\gamma_0/\gamma_1)}. \tag{9.22}$$

Later, we will discuss the EP and the heat flow for cases $n \geq 1$.

From (9.21) one immediately sees that the EP vanishes if, and only if, the FDR (9.16) is fulfilled. In this case, also the mean heat flow (9.22) and associated medium EP vanish. Results for $\dot{Q}_0$ are illustrated in Fig. 9.4, which highlights the role of NR coupling. Since we are primarily interested in the role of NR coupling, we focus in Fig. 9.4 and in the following on *isothermal conditions*, i.e., $\mathcal{T}_j = \mathcal{T}$ for all $j > 0$. Furthermore, we see from Eq. (9.22) that the heat flow may nullify, although the total EP (9.21) is positive, but only if $a_{01} = 0$, $a_{10} \neq 0$.

We conclude that the notion of equilibrium based on the non-Markovian description (from FDR) and based on the total EP of the network representation with additional d.o.f. are consistent. This finding is in accordance with [310], but contrasts the viewpoint in [118], where a different definition of the total EP was used. Besides being a consistency check, this also indicates that the total EP considered here gives some reasonable information in regard to the non-Markovian system.

At this point, we further aim to note an interesting observation concerning the so called "non-Markovian FDR" [94]

$$C_\nu(T) = \langle X_0^2\rangle\, K(T). \tag{9.23}$$

In the derivation [94], it seems that this relation holds under fairly general conditions. However, being broken despite zero EP, it is found to be not suitable to correctly

describe the equilibrium here. We suspect that this is due to the overdamped limit, which requires velocity relaxation as an additional assumption, absent in (9.23).

While the non-equilibrium nature of the non-Markovian Eq. (9.10) is indicated by the violation of FDR (9.16), the mechanism that prevents the system from equilibration, i.e., the driving, might not be very obvious. As we will show in Sec. 9.4, this point is clarified by studying the connection between network topology and energy flows in the corresponding Markovian description.

9.3.3 Analytical solutions

As already indicated by (9.19, 9.20), various thermodynamic and dynamic quantities can be calculated on the basis of (cross-)correlations $\langle X_i X_j \rangle$, including the PDFs, and information flows [see (9.48) below]. For example, the steady-state PDF ρ_{n+1} is, due to the linearity of the model, a Gaussian-distribution with zero mean and the covariance matrix $(\underline{\underline{\Sigma}})_{ij} = \langle X_i X_j \rangle$. Thus, it is fully determined by all the (cross-)correlations.

Let us briefly sketch how analytical expressions for the correlations can be obtained for arbitrary network sizes n. To this end, we transform Eqs. (9.9) via the Fourier-transformation $\tilde{X}_j(s) = \int_{-\infty}^{\infty} X_j(t) e^{-i\omega t} \mathrm{d}t$, which readily yields

$$i\omega \underline{\underline{\gamma}} \underline{\tilde{X}}(\omega) = \underline{\underline{a}} \underline{\tilde{X}}(\omega) + \underline{\tilde{\xi}}(\omega) \tag{9.24}$$

$$\Rightarrow \underline{\tilde{X}}(\omega) = \underbrace{\left(i\omega \underline{\underline{\gamma}} - \underline{\underline{a}} \right)^{-1}}_{= \underline{\underline{\tilde{\lambda}}}(\omega)} \underline{\tilde{\xi}}(\omega). \tag{9.25}$$

$\underline{\underline{\tilde{\lambda}}}(\omega)$ is the Green's function in frequency space, determined by the inverse of the topology matrix and the diagonal friction matrix $\underline{\underline{\gamma}}$ with elements γ_j. Using the well-known relationship (A.14) between spatial correlations and the Green's function from linear response theory [238] (see Appendix A.3 for a derivation) one readily finds

$$\langle X_j^2 \rangle = \frac{k_\mathrm{B}}{\pi} \sum_{p=0}^{n} \mathcal{T}_p \gamma_p \int_{-\infty}^{\infty} \tilde{\lambda}_{jp}(\omega) \tilde{\lambda}_{jp}(-\omega) \, \mathrm{d}\omega, \tag{9.26}$$

$$\langle X_j X_l \rangle = \frac{k_\mathrm{B}}{\pi} \sum_{p=0}^{n} \mathcal{T}_p \gamma_p \int_{-\infty}^{\infty} \tilde{\lambda}_{jp}(\omega) \tilde{\lambda}_{lp}(-\omega) \, \mathrm{d}\omega. \tag{9.27}$$

These are analytical expressions for all correlations for arbitrary network sizes. The matrix inversion required to find $\underline{\underline{\tilde{\lambda}}}(\omega)$ from (9.25) is indeed possible up to very large network sizes due to the sparse coupling. To find closed-form solutions for the integrals (9.26), the residue theorem can be used. However, this requires finding the roots of a polynomial of order $n+1$. Instead, we have evaluated them using computer algebra systems, specifically **Mathematica10.1**. Therewith, this can be done reasonably fast up to about $n = 10$. We also note that for the case $\mathcal{T}_{j>0} = 0$ solutions up to $n \sim 10^4$ can be found in this way.

9.4 Non-reciprocal coupling & Activity

To further unravel the non-equilibrium nature of the systems with NR coupling, we consider the system from an energetic perspective. Since we regard all d.o.f. to be of physical nature, each force applied by, or to, a subsystem j is inevitably connected to an energy exchange.

Let us for a moment take the perspective that each one of the d.o.f. X_j was an entire system. Then, Sekimoto's framework [124] tells us that the fluctuating heat flows along a trajectory of length dt are defined as

$$\delta q_j = (\gamma_j \dot{X}_j - \xi_j) \circ \mathrm{d}X_j = F_j \circ \mathrm{d}X_j, \tag{9.28}$$

the internal energy changes read

$$\mathrm{d}u_j = a_{jj} X_j \circ \mathrm{d}X_j. \tag{9.29}$$

while the work contributions are

$$\delta w_j := a_{ij} X_{j-1} \circ \mathrm{d}X_j. \tag{9.30}$$

Recall the usage of Stratonovich calculus throughout the paper (see Sec. 1.2.5). Now we return to the original perspective where the entire set of $X_{\{0,1,\dots,n\}}$ is the system, and the X_j are subsystems. How does the energetic consideration change? The important difference is the notion of work.

From a mechanical point of view, the fluctuating work δw_j as given by Eq. (9.30) is the energy associated with the coupling term $a_{ij}X_{j-1}$. It corresponds to the work applied to X_j via the single interaction between X_j and X_i ("coupling force times displacement"). If the coupling is *reciprocal*, i.e., stems from an interaction Hamiltonian $H_{\text{interaction}}$, the sum of work applied to X_i and X_j corresponds to its total differential

$$\delta w_i + \delta w_j = -\frac{\partial H_{\text{interaction}}}{\partial X_i} \circ \mathrm{d}X_i - \frac{\partial H_{\text{interaction}}}{\partial X_j} \circ \mathrm{d}X_j = -\mathrm{d}H_{\text{interaction}}. \tag{9.31}$$

Thus, in the reciprocally coupled case the energy exchange (9.31) is, in this picture, actually part of the *internal energy change* of the network $X_{\{0,1,\dots,n\}}$. For stability reasons, the latter must on average be conserved in steady states, hence, $\sum_{j=0}^{n} \dot{W}_j \equiv 0$.

In sharp contrast, a NR coupling turns out to be associated with a steady *energy input* into the system at rate

$$\dot{E}_{\text{input}} := \dot{W}_0 + \sum_{j=1}^{n} \dot{W}_j \geq 0. \tag{9.32}$$

Equality is reached for reciprocal coupling, and the positivity holds for isothermal conditions. The individual terms in (9.32) are given by Eq. (9.30). We recall that capital letters denote ensemble averages.

To prove the non-negativity of (9.32), we consider the energy balance. One shall expect that all picked up energy must ultimately dissipate. The validity of the first law of thermodynamics for each d.o.f: $\delta q_j = \delta w_j + \mathrm{d}u_j$ (with the energetic contributions given above), is indeed already implied in the model (9.9). As the average $\langle \mathrm{d}u_j \rangle = \langle a_{jj} X_j \circ \mathrm{d}X_j \rangle = a_{jj} \mathrm{d}\langle X_j^2 \rangle = 0$ is conserved in steady states, the first law implies $\dot{Q}_j = \dot{W}_j$. Thus, the energy input $\dot{E}_{\mathrm{input}}$, which we define in (9.32), is indeed identical to the total dissipation rate $\sum_{j=0}^{n} \dot{Q}_j \geq 0$. The latter is non-negative, as directly follows from the connection to the EP [see (9.19)]. A (positive) energy input on the level of the individual constituents is considered a defining property of active systems [102, 103, 305, 310, 342]. Hence, it is reasonable to use NR networks to model active matter, which is indeed commonly done.

We note that the energy flow through the system could be initiated by a (bio)chemical energy source, or some sort of external gradient, or force field. In any case, the source will be exploited, such that, looking at the entire system plus surrounding, this steady energy flow must result in free energy loss (or another respective thermodynamic potential). This can be seen by the fact that the energy flow is accompanied by total EP. For this reason, the energy input $\dot{E}_{\mathrm{input}}$ (9.32) can be associated with "free energy consumption".

The NR networks discussed here also include, as a specific case, a temperature gradient as underlying driving mechanism. In particular, if $a_{01}a_{10} > 0$ (upper right and lower left quadrants in Figs. 9.4,9.5), the NR coupling is (potentially) ascribable to a (hidden[4]) temperature gradient, as can be shown by mapping the network onto a symmetric system. We will show this in the next subsection.

9.4.1 "Mild cases"

We here show that for the specific case $a_{01}a_{10} > 0$, the NR coupling can potentially be traced back to a hidden temperature gradient. Hence, this case represents a "mild" form of intrinsic non-equilibrium. Consider the NR network of X_0 and X_1 with $a_{01}a_{10} \neq 0$

$$\begin{cases} \gamma_0 \dot{X}_0 &= a_{00} X_0 + a_{01} X_1 + \xi_0 \\ \gamma_1 \dot{X}_1 &= a_{10} X_0 + a_{11} X_1 + \xi_1. \end{cases} \tag{9.33}$$

Via a scaling, (9.33) can be mapped onto another network, in particular, by introducing $\widetilde{X}_0 = \sqrt{|a_{10}|}\, X_0$, $\widetilde{X}_1 = \sqrt{|a_{01}|}\, X_1$, and $\widetilde{\mathcal{T}}_0 = |a_{10}|\mathcal{T}_0$, $\widetilde{\mathcal{T}}_1 = |a_{01}|\mathcal{T}_1$. (We note that scaling of positional d.o.f. shall indeed be accompanied with scaling of the temperatures due to the connection between temperatures and the time-derivative of the positions, e.g., think of the equipartition theorem). In this way, we find

$$\begin{cases} \dot{\widetilde{X}}_0 &= a_{00} \widetilde{X}_0 + \mathrm{sgn}(a_{01}) \sqrt{a_{01}a_{10}} \widetilde{X}_1 + \widetilde{\xi}_0 \\ \dot{\widetilde{X}}_1 &= \mathrm{sgn}(a_{10}) \sqrt{a_{10}a_{01}} \widetilde{X}_0 + a_{11} \widetilde{X}_1 + \widetilde{\xi}_1, \end{cases} \tag{9.34}$$

with $\langle \widetilde{\xi}_i(t) \widetilde{\xi}_j(t) \rangle = 2k_{\mathrm{B}} \widetilde{\mathcal{T}}_j \gamma_j \delta_{ij} \delta(t - t')$. If $a_{01}a_{10} > 0$, this system has *reciprocal* coupling. Further, even if $\mathcal{T}_0 = \mathcal{T}_1$, it involves a temperature gradient ($\widetilde{\mathcal{T}}_0 \neq \widetilde{\mathcal{T}}_1$).

[4] If X_1 is hidden, i.e., when the non-Markovian description is employed.

This symmetric network (9.34) corresponds to the aforementioned model of two vanes at the angles $\widetilde{X}_0, \widetilde{X}_1$ in different heat baths, coupled by a torsion spring (see Sec. 9.2.1). As is well known, such a reciprocally coupled system equilibrates if, and only if, $\widetilde{\mathcal{T}}_1 = \widetilde{\mathcal{T}}_0$. This implies $|a_{01}|\mathcal{T}_1 = |a_{10}|\mathcal{T}_0$, which is indeed identical to the equilibrium-condition (9.16) found from FDR.

Now we turn to the impact of this scaling on the thermodynamic quantities. For the energy flows, we find the relations

$$\begin{aligned}\delta\widetilde{w}_0 &= \mathrm{sgn}(a_{01})\sqrt{a_{01}a_{10}}\,\widetilde{X}_1 \circ \mathrm{d}\tilde{X}_0 \\ &= a_{01}|a_{10}|\,X_1 \circ \mathrm{d}X_0 = |a_{10}|\,\delta w_0, \\ \delta\widetilde{q}_0 &= (\gamma_0\dot{\widetilde{X}}_0 - \widetilde{\xi}_0) \circ \mathrm{d}\tilde{X}_0 = |a_{10}|\,\delta q_0. \end{aligned} \tag{9.35}$$

Thus, the energy flows to and out of X_0 are both scaled with $|a_{10}|$. Likewise, $\delta\tilde{w}_1 = |a_{01}|\,\delta w_1$ and $\delta\tilde{q}_1 = |a_{01}|\,\delta q_1$. This further means

$$\Delta\tilde{s}_{\mathrm{tot}} = \frac{\delta\tilde{q}_0}{|a_{10}|\mathcal{T}_0} + \frac{\delta\tilde{q}_1}{|a_{01}|\mathcal{T}_1} = \frac{\delta q_0}{\mathcal{T}_0} + \frac{\delta q_1}{\mathcal{T}_1} = \Delta s_{\mathrm{tot}}, \tag{9.36}$$

i.e., the EP in the scaled model is *identical* to the EP in the original model, implying that equilibration of *both* models is expected, if (9.16) holds.

We conclude that the two "driving mechanisms", that is, NR coupling with $a_{01}a_{10} > 0$, or a temperature gradient, can formally not be distinguished on the level of EP. Hence, this represents a "mild" form of intrinsic non-equilibrium. Indeed, (hidden) temperature gradients have been discussed in the literature as possible mechanisms that fuel active motion, see, e.g., [343, 344].

It should be emphasized that a scaling as employed here cannot be found, if $a_{01}a_{10} \leq 0$. Thus, the NR coupling studied here is the more general case.

9.4.2 Reversed heat flow

Let us now consider the direction of the heat flow (9.22) induced by a NR coupled d.o.f.. We recall that in a passive system a non-conservative force acting on X_0 can induce a steady-state heat flow. The latter is strictly *nonnegative*, as dictated by the second law of thermodynamics $\dot{Q}_0/\mathcal{T}_0 = \dot{S}_{\mathrm{tot}} \geq 0$. Thus, heat is never flowing (on average) from the bath to the system.

In contrast, we find that the coupled subsystem X_n, which also applies a non-conservative force onto X_0 (namely $a_{0n}X_n$), *can* induce a reversed heat flow $\dot{Q}_0 < 0$. This is possible due to the usage of extracted information, as we discuss below. This implies a steady extraction of energy from the bath, which is converted into work $\dot{W}_0$, i.e., a (potentially useful) form of energy. As we have already mentioned in Part I (Chapter 3), it is well-known that such a reversed heat flow can be induced by "Maxwell-demon" type devices [136, 149]. In the following, we show that the $(n = 1)$-network considered here represents a minimal, time-continuous version of such a device.

Figure 9.4 reveals that there are parameter regimes, where the heat flow of the d.o.f. X_0 in the $(n = 1)$-network is reversed (blue regions in Fig. 9.4). Specifically,

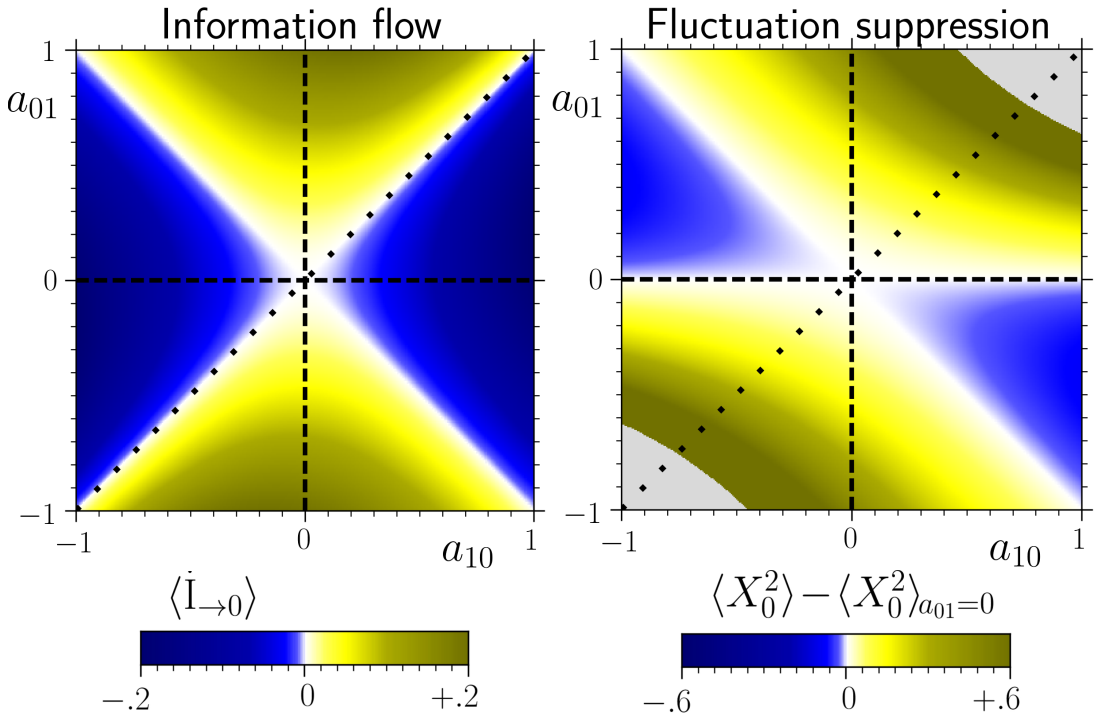

Figure 9.5: *Two d.o.f. X_0 and X_1, with NR coupling as in Fig. 9.4. Left: Information flow to X_0 in units of $1/\tau_{\mathrm{B}}$. Blue areas indicate that X_1 "knows" more about X_0 than vice versa corresponding to feedback control regimes ($|a_{10}| > |a_{01}|$). Right: Thermal fluctuations of X_0 measured by the second moment compared to the uncoupled case ($a_{01} = 0$), $\langle X_0^2 \rangle - \langle X_0^2 \rangle_{a_{01}=0}$ in units of σ^2. Blue areas indicate thermal fluctuation suppression. The gray areas indicate unstable regions (where $\langle X_0^2 \rangle \to \infty$). All other parameters are set to unity and $a_{11} = a_{00} = -1$.*

the energy extraction from the bath occurs if $a_{01}a_{10} > 0$. At this point, we recall the observation from the previous Chapter that a time-delayed feedback control force, in a linear system can induce as well a negative heat flow. There, it occurred for negative feedback[5] only. To compare our present system with the one from Chapter 8, we consider the projected equation (9.10) for the case $n = 1$. Here, the feedback force is $a_{01}X_1 \sim a_{01}a_{10}/\gamma_1 \int e^{a_{11}(t-t')/\gamma_1} X_0(t')\mathrm{d}t'$, i.e., has exponentially distributed delay instead of discrete, delta-distributed delay. Using the feedback controller analogy, the reversed heat flow here occurs only if $a_{01}a_{10} = k/\tau > 0$, which corresponds to *negative feedback*, i.e., analogous to the discrete delay case. Thus, the functional form of the memory kernel and the presence of colored noise, do not seem to qualitatively change the "energy extraction" regime. Later in this Chapter, we will return to this point and consider the behavior for larger n (see, e.g., Figs. 9.8, 9.11).

Taking a closer look at Fig. 9.4 further reveals that a reversed heat current occurs, if the "sensing" is stronger than the "active force" applied to X_0, i.e., $|a_{10}| > |a_{01}|$. To elucidate this observation, we consider the information flows.

[5] i.e., a force directed *towards* the past trajectory.

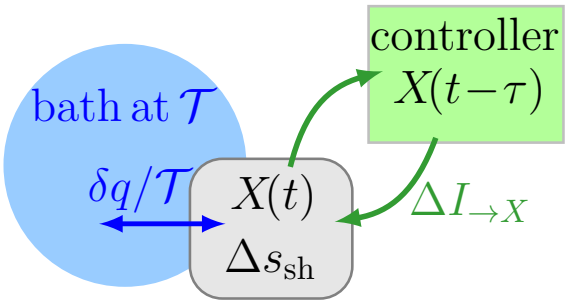

Figure 9.6: *Fluctuating entropy balance of the system $X(t)\hat{=}X_0(t)$: the Shannon entropy may vary upon a change of the PDF or $X(t)$. It corresponds to $s_{\rm sh}\hat{=}s^0_{\rm sh}$. In the steady state, this quantity is on average preserved. Further, the medium entropy might change due to a heat flow $\delta q\hat{=}\delta q_0$. Last, due to measurement and control action, an additional entropic term arises in the entropy balance (9.43) given by the information flow (3.30). The corresponding energetic picture is given in Fig. 3.2. In Fig. 3.3, we have formulated the expectations for the entropy balance.*

9.5 Non-reciprocal coupling & Information

In the previous Section, we have established that NR coupling implies an energy input on the level of individual constituents, which is a defining property of active matter. Now we change the focus onto information and feedback. We will address two questions. We have argued that a unidirectional coupling to another d.o.f. can model an *active force* onto X_i (of strength a_{ij}), i.e, "active propulsion"; or a *sensing operation* (of strength a_{ji}) [see example (9.12)]. *Do these interpretations conform with the information-thermodynamic properties?* Furthermore, a combination of both, i.e., sensing and applying an (active) force according to the measurement outcome, is in line with the *intuitive notion of feedback*. But, *is the existence of a single NR coupling sufficient to model a feedback loop?*

We clarify these questions by analyzing information flows and their connection to energy flows and entropy in detail. To this end, we use a framework similar to [345], where the reciprocally coupled case of (9.9) at $n = 1$ was studied in detail.

9.5.1 Information flow and generalized second law

"Information is physical"

– Rolf Landauer, 1991 [346]

In Section 9.3.2, we have considered the entropy balance of the entire network $X_{\{0,1,\dots,n\}}$. Now we want to consider the *entropy balance* of a single d.o.f. X_j. The motivation is to go back to the original picture, where X_0 represents the colloidal position X and the other d.o.f. represent the feedback controller, which are usually left out in the description. We recall that we expect an information-theoretic contribution in the balance of X_0, as sketched in Fig. 3.3. Here we derive the unknown information-theoretic term from Fig. 3.3.

We start by considering the total temporal derivative of the multivariate (joint) Shannon entropy (3.15) of the $(n+1)$-point PDF, i.e.,

$$\frac{\dot{s}_{\rm sh}}{k_{\rm B}} = \frac{-\partial_t \rho_{n+1}}{\rho_{n+1}} + \sum_{j=0}^{n} \frac{-\left(\partial_{x_j}\rho_{n+1}\right)\dot{X}_j}{\rho_{n+1}}. \tag{9.37}$$

In the steady state, the first term naturally vanishes. Therefore, we concentrate on the sum. To calculate the ensemble average of the latter, use $\langle \dot{X}_j A(x_j,t)\rangle = \int J_j A(x_j,t) \mathrm{d}x_j$ [125, 267] and apply the natural boundary conditions (see Chapter 2). This readily yields

$$\begin{aligned}
\left\langle \frac{-\left(\partial_{x_j}\rho_{n+1}\right)\dot{X}_j}{\rho_{n+1}} \right\rangle &= \int \frac{-\left(\partial_{x_j}\rho_{n+1}\right) J_j}{\rho_{n+1}} \,\mathrm{d}\underline{x} \\
&= -\int \underbrace{[\ln(\rho_{n+1}) J_j]_{-\infty}^{\infty}}_{\to 0} \,\mathrm{d}\underline{x}_{i\neq j} + \int \ln \rho_{n+1}(\underline{x}) \partial_{x_j} J_j \,\mathrm{d}\underline{x} \\
&= -\underbrace{\int \ln \frac{\rho_1(x_j)}{\rho_{n+1}(\underline{x})} \partial_{x_j} J_j \,\mathrm{d}\underline{x}}_{=\dot{I}_{\to j}} + \int \ln \rho_1(x_j)\, \partial_{x_j} J_j \,\mathrm{d}\underline{x}.
\end{aligned} \tag{9.38}$$

In the last step, we have identified the information flow $\dot{I}_{\to j}$, which we have introduced before in Sec. 3.5 Eq. (3.30). Importantly, the involved information flow is from *all* other d.o.f. $\{X_{l\neq j}\}$ to X_j. Even if not directly coupled with each other, two d.o.f. can exchange information through a third d.o.f.

To further proceed, we utilize the closed, multivariate Fokker-Planck equation for the $(n+1)$-point probability distribution $\rho_{n+1}(\underline{x},t)$ of $\underline{x} = (x_0, ..., x_n)^T$, which we have given in Chapter 2 [in Eq. (2.14)]. We aim to stress that the existence of the closed FPE is a key benefit of employing a Markovian description[6]. More specifically, we use the FPE to replace the $\partial_{x_j} J_j$-term in the integral, and find

$$\begin{aligned}
\left\langle \frac{-\left(\partial_{x_j}\rho_{n+1}\right)\dot{X}_j}{\rho_{n+1}} \right\rangle &= -\dot{I}_{\to j} + \int \ln \rho_1(x_j) \left\{ -\partial_t \rho_{n+1} + \sum_{i\neq j} \partial_{x_i} J_i \right\} \mathrm{d}\underline{x} \\
&= -\dot{I}_{\to j} + \frac{\dot{S}^j_{\rm sh}}{k_{\rm B}} - \sum_{i\neq j} \int \ln \rho_1(x_j) \underbrace{[J_i]_{-\infty}^{\infty}}_{\to 0} \,\mathrm{d}\underline{x}_{l\neq i},
\end{aligned} \tag{9.39}$$

where we have introduced the *Shannon entropy of the marginal PDF* $\rho_1(x_j)$: $\dot{S}^j_{\rm sh} = -k_{\rm B}\int \ln \rho_1(x_j) \partial_t \rho_1(x_j) \,\mathrm{d}x_j$. To summarize, by considering the ensemble average of each summand of (9.37), we have found that the total Shannon entropy change is given by $\dot{S}_{\rm Sh} = -k_{\rm B}\sum_{j=0}^{n} \dot{I}_{\to j} + \dot{S}^j_{\rm sh}$.

As we have shown in Sec. 3.5, the involved information flows are connected to the mutual information, which is a conserved quantity in steady states. Specifically, $\dot{\mathcal{I}} = \sum_{j=0}^{n} I_{\to j} = 0$. This immediately tells us that the sum over all information

[6]There is in general no closed FPE for non-Markovian systems of type (9.10), as we have discussed in detail in Part II.

flows nullifies in the steady state, which means that we do not have to take them into account. However, when an *individual* subsystem is considered, they turn out to be an important contribution to the *entropy balance*.

To see this, we reconsider the summands of (9.37), and now rewrite them using the FPE (2.14) as

$$\frac{-\left(\partial_{x_j}\rho_{n+1}\right)\dot{X}_j}{k_{\mathrm{B}}^{-1}\rho_{n+1}} = \underbrace{\frac{\gamma_j\, J_j(\underline{x},t)\dot{X}_j}{\mathcal{T}_j\,\rho_{n+1}}}_{:=k_{\mathrm{B}}\dot{s}^j_{\mathrm{tot}}} - \frac{\dot{q}_j}{\mathcal{T}_j}. \tag{9.40}$$

Combining (9.37, 9.39, 9.40) readily yields

$$\dot{S}^j_{\mathrm{tot}} = \dot{S}^j_{\mathrm{sh}} - k_{\mathrm{B}}\dot{I}_{\to j} + \frac{\dot{Q}_j}{\mathcal{T}_j} = \int \frac{\gamma J_j(\underline{x},t)^2}{k_{\mathrm{B}}\mathcal{T}_j\,\rho_{n+1}}\,\mathrm{d}\underline{x} \geq 0, \tag{9.41}$$

$$\dot{S}_{\mathrm{tot}} = \sum_{j=0}^{n}\frac{\dot{Q}_j}{\mathcal{T}_j} + \dot{S}_{\mathrm{sh}} = \sum_{j=0}^{n}\dot{S}^j_{\mathrm{tot}} \geq 0. \tag{9.42}$$

With Eq. (9.42) we have recovered the mean *total EP*, which we have equivalently deduced from the path probabilities in Eq. (3.15). From Eq. (9.41) we see that the definition (9.40) of a fluctuating "total EP per subsystem", $\dot{s}^j_{\mathrm{tot}}$, is meaningful, as it has the non-negative ensemble average. Equation (9.41) is a *generalized second law* for each d.o.f., which involves the information flow.

Returning to the picture of the colloid controlled by an external feedback controller, this equation gives the appropriate entropy balance (with $j = 0$)

$$\dot{S}^0_{\mathrm{tot}} = \dot{S}^0_{\mathrm{sh}}[\rho_1(X_0)] - k_{\mathrm{B}}\dot{I}_{\to 0} + \frac{\dot{Q}_0}{\mathcal{T}_0} \geq 0, \tag{9.43}$$

as depicted in Fig. 9.6. Besides the medium EP and Shannon entropy change (of the marginalized PDF), it contains the information flow. The latter represents the aforementioned "missing entropic term".

In steady states $\dot{S}^j_{\mathrm{sh}} = 0$. Thus, (9.41) simplifies to

$$\dot{Q}_j \geq k_{\mathrm{B}}\mathcal{T}_j\,\dot{I}_{\to j}, \tag{9.44}$$

in accordance with [289, 345]. Equation (9.44) implies that a reversed steady heat flow, $\dot{Q}_0 < 0$, is only possible, if $\dot{I}_{\to 0} < 0$, i.e., *information is flowing from the X_0 to the rest of the system*. The more information about X_0 is gathered by the other $X_{j>0}$ (the controller d.o.f., see Sec. 9.1.1), the more heat can be extracted from the bath.

Figure 9.7 plots the information and heat flows, as well as the total EP of our network at $n = 1$, which are all connected via (9.41, 9.42). It also illustrates that in the reciprocal case, the memory is free of "entropic cost" (zero EP), but, at the same time, no net information extraction is achieved, nor is a heat flow induced. This further implies that the "missing entropic term" is specific to the NR case.

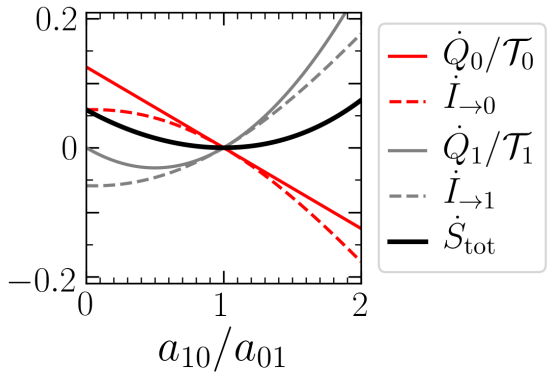

Figure 9.7: *Information $\dot{I}_{\to j}$ (9.48) in units of $1/\tau_{\rm B}$ and medium EP $\dot{Q}_j/\mathcal{T}_j$ (9.22) in units of $k_{\rm B}/\tau_{\rm B}$, for two NR coupled d.o.f., vs. the ratio of coupling strength, a_{01}/a_{10}. At $a_{01} = a_{10}$, the system is reciprocally coupled and displays no net heat and information flow, and has zero total EP (black line). The medium EP (solid red and gray lines) of each d.o.f. is bounded from below by the information flow (dashed lines), as predicted by the generalized second law (9.44). The total EP is given by the sum over all other quantities. $a_{11} = a_{00} = -\gamma_0/\tau_{\rm B}$, all other parameters are unity.*

In this Section, we have derived a generalized second law for individual subsystems X_j. At the same time, we have shown that the ensemble average of the total EP for the *entire* system $\{X_0, ..., X_n\}$

$$\dot{S}_{\rm tot} = \sum_{j=0}^{n} \int \frac{\gamma J_j(\underline{x}, t)^2}{k_{\rm B} \mathcal{T}_j \, \rho_{n+1}} \, {\rm d}\underline{x} \geq 0, \tag{9.45}$$

fulfills the second law (9.42) as is immediately apparent from the nonnegative expression (9.45). It is noteworthy that we have therewith also seen that the mean total EP can be derived by means of the FPE (instead of path integrals) [as we have already noted in Chapter 3, see Eq. (3.20)]. This approach is possible for Markovian systems, where a closed FPE is known [125].

Explicit expressions for information flows Let us now calculate the information flow. As already mentioned in Sec. 9.3.3, due to the linearity of our model the steady-state PDFs are *multivariate Gaussians* with zero mean and with the covariance matrix $(\underline{\underline{\Sigma}})_{ij} = \langle X_i X_j \rangle$. To derive explicit expressions for the steady-state averaged information flows, it turns out to be most convenient to use

$$\dot{I}_{\to j} = \left\langle \frac{\left(\partial_{x_j} \rho_{n+1}\right) \dot{X}_j}{\rho_{n+1}} \right\rangle = - \left\langle (\underline{\underline{\Sigma}}^{-1} \underline{X})_j \dot{X}_j \right\rangle, \tag{9.46}$$

which stems from Eq. (3.30). In the last step, we have utilized a general property of normal distributions $\partial_{x_j} \rho_{n+1}(\underline{x}) = -(\underline{\underline{\Sigma}}^{-1} \underline{X})_j \rho_{n+1}$. Substituting (9.9), we find the general formula

$$\dot{I}_{\to j} = - \sum_{l \neq j} (\underline{\underline{\Sigma}}^{-1})_{lj} \left[\frac{a_{jj}}{\gamma_j} \langle X_l X_j \rangle + \frac{a_{jj-1}}{\gamma_j} \langle X_l X_{j-1} \rangle \right]. \tag{9.47}$$

In combination with (9.26), Eq. (9.47) represents an analytic expression for the information flow to any node in the network of arbitrary size. The explicit closed-form expressions are, however, rather tedious. Let us consider the results for $n = 1, 2$.

For $n = 1$, the information flow to X_0 simplifies to

$$\begin{aligned}\gamma_0 \dot{I}_{\to 0} =& \frac{-\langle X_1 X_0\rangle}{\langle X_1 X_0\rangle^2 - \langle X_1^2\rangle\langle X_0^2\rangle}\frac{\dot{Q}_0}{\mathcal{T}_0\, a_{01}} \\ =& \frac{[a_{01}\mathcal{T}_1 - a_{10}\mathcal{T}_0][a_{00}a_{01}/\mathcal{T}_0 + a_{11}a_{10}\gamma_0/(\mathcal{T}_1\gamma_1)]}{(a_{00}\gamma_1 + a_{11}\gamma_0)^2 + a_{01}^2\mathcal{T}_1/(\mathcal{T}_0\gamma_0\gamma_1) - 2a_{01}a_{10} + a_{10}^2\mathcal{T}_0/\mathcal{T}_1}. \end{aligned} \tag{9.48}$$

According to (9.48), the information flow trivially vanishes if X_0 and X_1 are uncorrelated, as one would expect. It also vanishes for reciprocally coupled systems under isothermal conditions. More generally, at non-isothermal conditions, it nullifies at $a_{10}\mathcal{T}_0 = a_{01}\mathcal{T}_1$. These two cases correspond to thermal equilibrium situations, as shown above.

Furthermore, we see that $\dot{I}_{\to 0}$ is automatically zero, if $\dot{Q}_0 = 0$. This implication $\dot{Q}_0 = 0 \Rightarrow \dot{I}_{\to 0} = 0$ was already noticed in Ref. [345], where the reciprocal case of our model was considered. However, we see here that this implication is only true if $a_{01} \neq 0$.

For $n = 2$, Eq. (9.47) yields

$$\begin{aligned}\psi\gamma_0\dot{I}_{\to 0} &= [\langle X_0X_2\rangle\langle X_1X_2\rangle - \langle X_0X_1\rangle\langle X_2^2\rangle]\dot{Q}_0/(\mathcal{T}_0 a_{02}) \\ &\quad + [\langle X_0X_1\rangle\langle X_1X_2\rangle - \langle X_0X_2\rangle\langle X_1^2\rangle][a_{00}\langle X_0X_2\rangle + a_{02}\langle X_2^2\rangle], \qquad (9.49) \\ \text{with}\quad \psi &:= \langle X_0^2\rangle\langle X_1X_2\rangle^2 + \langle X_1^2\rangle\langle X_0X_2\rangle^2 + \langle X_2^2\rangle\langle X_0X_1\rangle^2 \\ &\quad - \langle X_0^2\rangle\langle X_1^2\rangle\langle X_2^2\rangle - 2\langle X_0X_1\rangle\langle X_0X_2\rangle\langle X_1X_2\rangle. \qquad (9.50)\end{aligned}$$

This term is quite cumbersome and, admittedly, not very enlightening. However, (9.49) suggests that in this case the information flow can be nonzero, despite $\dot{Q}_0 = 0$ and $a_{02} \neq 0$. For even higher n, the results get even more tedious.

Information-theoretic perspective on NR coupling We are now in the position to clarify the meaning of unidirectional coupling from an information-thermodynamic perspective. Inspecting Figs. 9.4 and 9.5 we see that along the unidirectional coupling axis $a_{01} = 0$, there is net information flow from X_0 to X_1, but no net work applied onto X_0 ($\dot{Q}_0 = \dot{W}_0 = 0$). Thus, it is indeed very reasonable to consider X_1 a "sensor" and the unidirectional coupling a *sensing interaction*. On the other hand, if the unidirectional coupling is reversed ($a_{10} = 0$), the heat flow is always positive, $\dot{Q}_0 > 0$. This suits the idea that X_0 models an active swimmer within the AOUP model. In particular, the swimmer heats up the surrounding fluid but never has a net cooling effect, as shall be expected. In this case, there is as well a nonzero information flow, which is directed from the source of propulsion (i.e., the flagella or flow field) to the particle. This is indeed reasonable as the propulsion force "carries" information (one could, on average, reconstruct the position of the flagella by only monitoring X_0).

9.5.2 Information-theoretic perspective on feedback control

According to the results from the last Section, a (unidirectional) interaction from one d.o.f., say X_0, to another, X_1, can model a measurement operation, while the

backwards coupling can model an active (propulsion) force. Thus, a NR interaction can be considered as a combination of a measurement and a force depending on the measurement outcome, or, in other words, *feedback*. In this sense, we can state that our networks with NR bidirectional coupling model feedback loops.

However, it seems intuitive to consider X_0 a feedback-*controlled* system, if and only if the net information flow out of X_0 is positive, i.e., the controller "knows" more about X_0 than vice versa.

According to this definition, the *control regime* for the model (9.12), is given if $|a_{10}| > |a_{01}|$, see the blue regions in Fig. 9.5. The generalized second law (9.44) states that only in this regime, a reversed heat flow is possible, as indeed confirmed by Fig. 9.4. Interestingly, we find that another intricate phenomenon is limited to the control regime, namely, the *suppression of thermal fluctuations* of X_0, measured by a reduced second moment $\langle X_0^2\rangle < \langle X_0^2\rangle_{a_{01}=0}$ (see Fig. 9.5). Such a suppression, i.e., noise-reduction (which can be interpreted as feedback cooling) is desired in various experimental setups, and indeed one of the main applications of feedback control [347–349].

We thus conclude that characteristic features of feedback-controlled systems are described by a single NR coupling, e.g., by (9.9) with $n = 1$ and $|a_{10}| > |a_{01}|$. X_1 represents a controller that utilizes information about X_0 extracted via a sensing operation. The resulting system is characterized by positive information flow from X_0 to X_1, reversed heat flow, and suppression of thermal fluctuations (depending on the strength of the "feedback force" $a_{01}X_1$).

For further discussions about information flow, energetics and control, we refer the interested Reader to [345].

9.6 Total entropy production and heat flow in the presence of non-monotonic memory

Let us return to the picture where X_0 is the position of a colloidal particle, while $X_{\{1,2,..,n\}}$ represent feedback controller degrees of freedom, and now focus on the impact of the number of n. To this end, we revisit the formulae derived in the previous Sections. First, we consider the mean heat flow from X_0 to its bath from Eq. (9.20), which is given by $\dot{Q}_0 = k^2\langle X_n^2\rangle + ka_{00}\langle X_0X_n\rangle$, with k being the strength of the feedback kX_n, and $a_{00}X_0$ being the restoring force. A comparison of the cases $n = 1, 2$ is instructive, since the networks are of comparable size while the memory kernels $K(T)$ have distinct characteristics, i.e., exponential decay with maximum at $T = 0$ ($n = 1$) vs. a non-monotonic kernel with a minimum at $T = 0$ ($n = 2$), see Fig. 5.2 (a,b). Figure 9.8 reveals that the heat flow $\dot{Q}_0$ is qualitatively and quantitatively similar for $n = 1$ and 2. The (blue) area of reversed heat flow is slightly smaller for $n = 2$, but lies in the same region of the (τ, k)-plane. For $k = 0$, there is trivially no heat flow (in this case, X_0 is a passive d.o.f. which does not "see" the other d.o.f.). For both $n = 1, 2$, there is a second line along which the heat flow vanishes. For $n = 1$, this line corresponds to parameters where the FDR (9.16) is fulfilled (dashed line). As shown in Fig. 9.9, the total EP is then zero

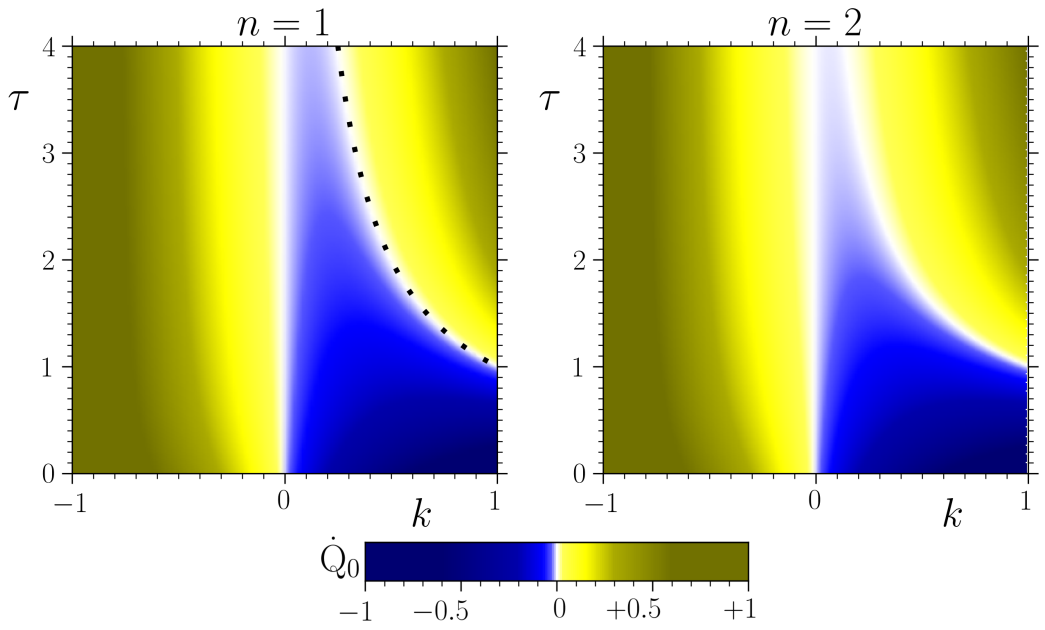

Figure 9.8: *Heat flow from X_0 to its bath [from Eqs. (9.20, 9.22)], in networks of type (9.11) of different sizes; left: $n = 1$ (one "memory cell", exponential memory) and right: $n = 2$ (two "memory cells", non-monotonic memory). For $n = 1$, the FDR (9.16) is fulfilled at $k = 1/\tau$ (dotted line). Since this system equilibrates, $\dot{Q}_0 = 0$. In the uncoupled case ($k = 0$), the subsystem X_0 equilibrates as well (for arbitrary n). All other parameters and k_B are set to unity.*

as well. For $n = 2$, the FDR is broken for all (τ, k) and the system never reaches thermal equilibrium. Correspondingly, the EP is strictly positive, quantifying the "thermodynamic cost" of the non-monotonic memory, or, the dissipation of $X_{j>0}$. For $n > 1$, there are (τ, k) values where the system is out of equilibrium, but nevertheless $\dot{Q}_0 = 0$. Thus, the vanishing medium EP is not sufficient to signal equilibrium. We recall that in the last Chapter, we also observed points of vanishing heat flux in the presence of feedback with discrete delay (see Chapter 8). Note that the case $n = 1$ is significantly different in this regard, and zero dissipation always coincides with equilibrium.

While we generally focus on the mean values of the total EP, which we can access analytically, let us briefly take a look onto the fluctuations. Figure 9.10 displays probability distributions of the total EP fluctuations (obtained by BD simulations). The distributions have very similar characteristics for $n = 1$ and 2, despite the distinct memory kernels (exponential decay vs. non-monotonic memory). In particular, they both have exponentially decaying tails, and fulfill the IFT $\langle e^{-\Delta s_\mathrm{tot}/k_\mathrm{B}} \rangle = 1$, as expected (see Sec. 3.4). The characteristics of the entropy fluctuations are, in fact, identical for all cases, whether the network models an active microswimmer, a cellular sensor, or a feedback-controlled colloid with time delay. In the following, we will again return to the discussion of the mean values.

How do the mean values behave upon further increase of n, corresponding to more sharply peaked memory kernels appearing in the dynamics of X_0? In our

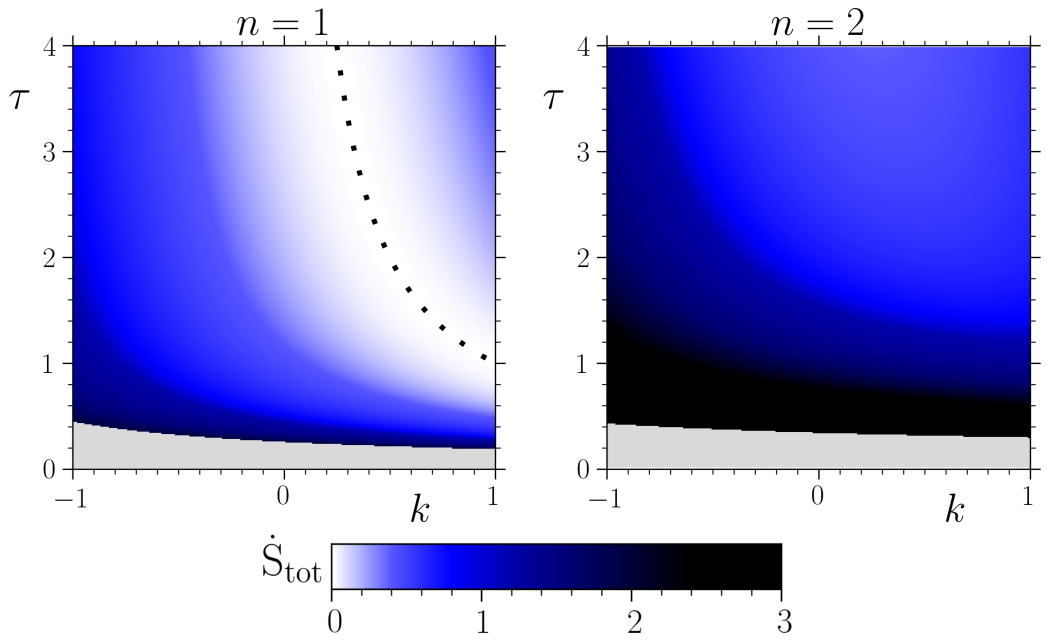

Figure 9.9: *Total EP $\dot{S}_{\rm tot}$ [from Eqs. (9.19, 9.21)] in units of $k_{\rm B}/\tau_{\rm B}$, in networks of sizes $n = 1, 2$, as in Fig. 9.8. Along the dashed line, FDR (9.16) is fulfilled, and the total EP is zero, showing consistency between the non-Markovian description (9.10) and the Markovian network description (9.9). The gray areas indicate unstable regions (it is slightly larger in the case $n = 2$).*

interpretation, this means an increase of the number of memory cells. Figure 9.11 shows analytical results for the mean total EP and heat flow as functions of n, for three different values of the feedback strength k, i.e., the coupling between X_n and X_0. We observe that the total EP generally increases quadratically with n. This observation is robust against the variation of system parameters, e.g., the temperatures, coupling strengths, or delay time. In fact, it appears to hold even in nonlinear systems. To check this, we have added a nonlinear force F_0 in the dynamical equation of X_0, Eq. (9.9), and performed BD simulations to calculate the EP (not shown here). For all cases, we find quadratic increase with n.

In sharp contrast, the heat flow saturates with n, as shown in the right panel of Fig. 9.11. Moreover, in agreement with our previous observation for the case $n = 1$ (e.g., see Fig. 9.4), we again find a reversed heat flow ($\dot{Q}_0 < 0$) for negative feedback, i.e., when the feedback force is directed towards the past trajectory.

9.6.1 Limit of discrete delay

In realistic setups, a controller has a finite memory capacity, i.e., the amount of stored information is limited. This implies, on the one hand, that $\mathcal{T}'$ is finite (finite precision), and on the other hand, that the number of memory cells n should be finite. The resulting memory kernel has a peak around τ with non-zero width (distributed delay).

However, in theoretical studies, the delay is often assumed to be discrete, i.e.,

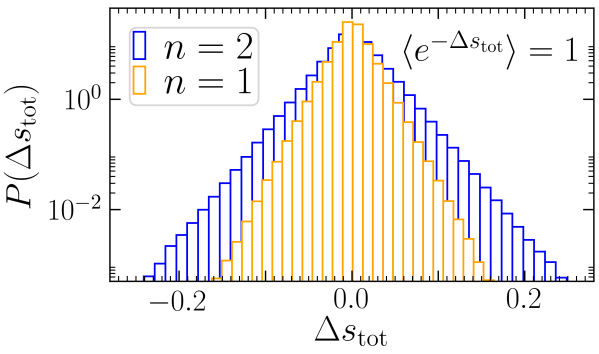

Figure 9.10: *Distributions of the fluctuating total EP $\Delta s_{\rm tot}$ (3.15) for $n = 1, 2$ from Brownian dynamics simulations. The linear decays in this logarithmic plots indicate exponential tails. The mean values are in good agreement with the analytical results ($\Delta S_{\rm tot} = 2$ for $n = 1$, and $\Delta S_{\rm tot} = 5$ for $n = 2$), and both distributions fulfill the integral fluctuation theorem (3.22). To calculate the distributions, 5×10^6 steady-state trajectories of length 10^{-4} were generated. $k = -1$, and all other parameters and $k_{\rm B}$ are set to unity.*

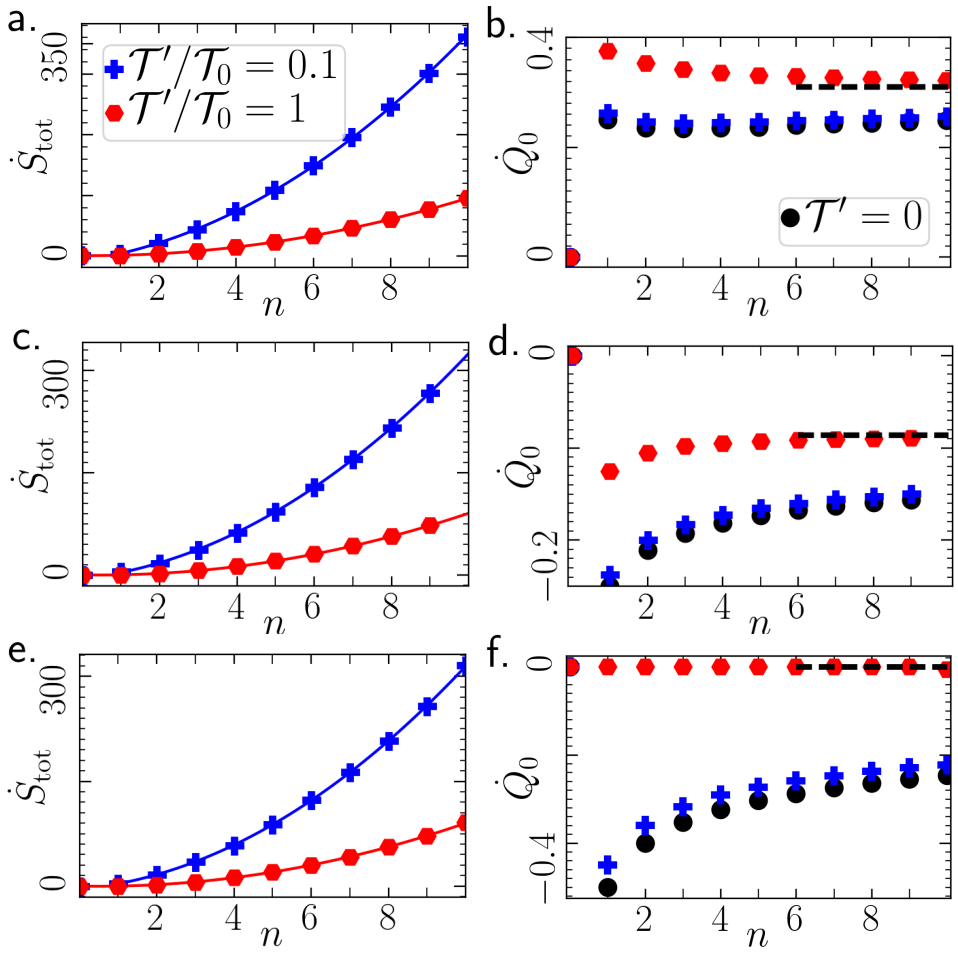

Figure 9.11: *(a.+c.+e.) Total EP [from Eq. (9.19)], and (b.+d.+f.) heat flow $\dot{Q}_0$ [from Eq. (9.20)], vs. the number n of $X_{j>0}$. The EP plots are complemented by quadratic fits $\dot{S}_{\rm tot} \sim n^2$ (solid lines). For all temperature ratios $\mathcal{T}'/\mathcal{T}_0$ (different colors), the heat flow converges in the limit $n \to \infty$ to the discrete delay solution (dashed lines). At $\mathcal{T}' = 0$ (black disks) $\dot{S}_{\rm tot}$ diverges. (a.+b.) $k = -1/2$ (positive feedback), (c.+d.) $k = 1/2$ (negative feedback). (e.+f.) $k = 1$, where the system with $n = 1$ is reciprocally coupled. All other parameters and $k_{\rm B}$ are set to unity.*

infinitely sharp. In our approach, this corresponds to the limit $n \to \infty$, where $K(T) \to \delta(T - \tau)$. In this limit, we find a surprising result for the colored noise. As we have shown in Chapter 5 (see Sec. 5.1.2), the noise correlations completely *vanish*

$$\lim_{n\to\infty} C_\xi = 0 \tag{9.51}$$

(irrespective of the value of $\mathcal{T}'$). Hence, a delay LE with white noise is recovered from (9.10), that is,

$$\dot{X}_0(t) = a_{00}\, X_0(t) + k\, X_0(t-\tau) + \xi_0(t). \tag{9.52}$$

This implies that the heat flow (9.20) approaches the value of the heat flow for discrete delay and white noise, which we have intensely studied in the previous Chapter, for both, linear and nonlinear systems. The corresponding values [the formula is given in Eq. (8.17)] are plotted as dashed lines in Fig. 9.11. Indeed, the analytical results for finite n seem to converge to this value. However, the convergence is relatively slow.

Remarkably, we generally find that $\dot{Q}_0$ approaches the limit fastest for isothermal conditions $\mathcal{T}_0 = \mathcal{T}'$. We will come back to this observation below.

As we have discussed above, the total entropy production increases quadratically with n, which implies a *divergent* EP in the limit $n \to \infty$. Taking the perspective that the network models a memory device, where each memory cell can contribute to dissipation, a divergent EP is indeed somewhat expected due to the infinite system size. Also from an information-theoretical viewpoint, this result is in fact not too surprising. To realize (9.52), a random trajectory of length τ needs to be memorized. However, the information content of a Brownian trajectory is infinite (even if it is arbitrarily short), since it is governed by *white noise*, which yields infinitely many jumps within arbitrarily short time intervals. This infinite amount of information results in a divergent EP.

To avoid confusion, we aim to differentiate our findings from [93, 116, 152], where a finite measure for irreversibility was reported. We have briefly reviewed this approach earlier in Sec 3.6.4. Specifically, in this line of work, an irreversibility measures based on X_0 alone is employed, for the case of underdamped motion. In contrast, we here embed the delay problem into a higher dimensional structure and then use the standard framework of thermodynamics (and do not use a specifically adapted formalism). Therefore, our result is very different. We find a diverging EP owing to the infinite-dimensional character of this case, which is apparent from the infinite network size, as well as the infinite-dimensional nature of the delta-distribution kernel.

9.6.2 Impact of measurement errors

Finally, we briefly consider the impact of $\mathcal{T}'$, that is, the strength of the noise terms ξ_j in Eq. (9.9) (accounting for measurement or rounding errors). As visible in Fig. 9.11, the heat flow converges to the previously obtained result as $n \to \infty$ (dashed lines), fastest for isothermal conditions $\mathcal{T}_0 = \mathcal{T}'$. This implies that the

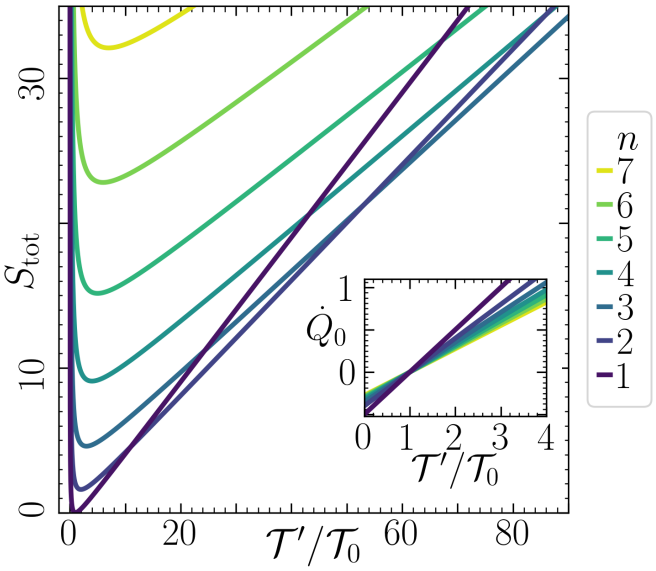

Figure 9.12: *Total EP $\dot{S}_{\rm tot}$ [from Eq. (9.19)] as a function of the ratio of the temperatures of $X_{j>0}$ and X_0, for networks with different numbers n of d.o.f. $X_{j>0}$. The parameters are identical to Fig. 9.11 (e.+f.). The network with $n = 1$ is reciprocally coupled, and fulfills FDR (9.16) if $\mathcal{T}'/\mathcal{T}_0 = 1$. In accordance, the total EP vanishes. For each n, the EP has a minimum at a finite $\mathcal{T}'/\mathcal{T}_0$ of order 1, and diverges in the limits $\mathcal{T}'/\mathcal{T}_0 \to 0$ or $\mathcal{T}'/\mathcal{T}_0 \to \infty$. The inset shows the heat flow $\dot{Q}_0$, which for all n increases linearly and is zero at $\mathcal{T}'/\mathcal{T}_0 = 1$.*

details of the memory kernel (which depends on n) have a larger impact, if the temperatures are inhomogeneously distributed over the network. This is somehow puzzling, especially when one considers the colored noise. We recall that the colored noise is present at finite n, but vanishes in the limit $n \to \infty$ for all cases $\mathcal{T}'$. At $\mathcal{T}' = 0$, it is not present even for finite n, and its impact increases with $\mathcal{T}'$. Given these qualitative arguments, it is very unexpected that the case $\mathcal{T}' = \mathcal{T}_0$ converges fastest to the limit. So far, we have not resolved this seeming contradiction, but it would be very interesting to study this in more detail in the future (especially, because the finite n case may be used to approximate the discrete delay case).

Figure 9.12 shows that the heat flow $\dot{Q}_0$ increases linearly with $\mathcal{T}'/\mathcal{T}_0$. Due to the specific parameter choice in Fig. 9.12, where $n = 1$ corresponds to a reciprocally coupled network, the heat flow vanishes at $\mathcal{T}_0 = \mathcal{T}'$ for all n. Thus, at $\mathcal{T}_0 = \mathcal{T}'$, all lines cross, and the number n matters least. For different parameter settings, we also observe that the lines cross near this point (not shown here).

On the contrary, the total EP has for each n a pronounced minimum at a finite $\mathcal{T}'/\mathcal{T}_0$. Interestingly, at some values of $\mathcal{T}'/\mathcal{T}_0$ above the minima, the total EP for different n are *identical* (i.e., the lines in Fig. 9.12 cross). This implies that two networks at the same temperature but with a different number of nodes can produce the same amount of entropy. This is somewhat surprising, as one would expect that a larger system automatically has a higher EP.

The EP diverges if one of the temperatures goes to zero, corresponding to the cases of vanishing or infinite measurement errors. This is in accordance with the result from [135], whose limit of "precise and infinitely fast" control corresponds in our time-continuous model to the limit $\mathcal{T}' \to 0$. We also note that when we interpret X_0 as the position of an active swimmer, $\mathcal{T}_0 \to 0$ corresponds to the limit of zero translational ("passive") noise (and pure "active" fluctuations). The divergence of EP in this limit was noted in [104] for a similar active swimmer model.

9.7 Irreversibility & Coarse-graining

In this last Section, we elucidate the relationship between our network model and the representation via X_0 alone from a more fundamental perspective in the context of *irreversibility* and *effective thermodynamics*. We have already mentioned several times that the investigation of the network model in this Chapter can apply to two situations. First, if all X_j are d.o.f. with a direct physical meaning, we have considered the total EP $\Delta s_{\rm tot}$ of that actual model, and the interpretation of all entropic terms is straightforward. Here, we now consider the second, more subtle situation, i.e., if the $X_{j>0}$ are only considerable as auxiliary variables, but have no direct physical interpretation. As we have seen, we can indeed get an interesting and reasonable view on the process by considering the total EP $\Delta s_{\rm tot}$, although the additional entropic terms have a somewhat debatable meaning.

As mentioned in Sec. 3.6.4, finding thermodynamic notions for systems where X_0 is the only observable (non-Markovian) process (9.10), is the objective of several recent publications aiming for an effective thermodynamics, see e.g., [93, 105, 107, 109, 116, 152]. Indeed, in many situations of particular interest, X_0 is the only experimentally accessible quantity. Then, a natural irreversibility measure, which can directly be calculated from the observed trajectories, would be [107]

$$\frac{\Delta s^0}{k_{\rm B}} := \ln \frac{\mathcal{P}[\mathbf{X}_0]}{\hat{\mathcal{P}}[\hat{\mathbf{X}}_0]}, \tag{9.53}$$

[as we have already given in Eq. (3.36)]. With this picture in mind, one may question the meaning of the formula for the total EP $\Delta s_{\rm tot}$, which explicitly involves $X_{j>0}$. However, our framework also yields a statement about Δs^0.

First, we notice that Δs^0 corresponds to one part of the total EP (3.15). To see this, we use the definition of conditional path probabilities, i.e., $\mathcal{P}[\mathbf{X}_0;\underline{\mathbf{X}}_{\rm c}] = \mathcal{P}[\underline{\mathbf{X}}_{\rm c}|\mathbf{X}_0]\mathcal{P}[\mathbf{X}_0]$, with $\underline{\mathbf{X}}_{\rm c} = \{\mathbf{X}_1,...,\mathbf{X}_n\}$, to split up the total EP

$$\frac{\Delta s_{\rm tot}[\underline{\mathbf{X}}]}{k_{\rm B}} = \ln \frac{\mathcal{P}[\underline{\mathbf{X}}]}{\hat{\mathcal{P}}[\underline{\hat{\mathbf{X}}}]} = \frac{s^0[\mathbf{X}_0]}{k_{\rm B}} + \underbrace{\ln \frac{\mathcal{P}[\mathbf{X}_{\rm c}|\mathbf{X}_0]}{\hat{\mathcal{P}}[\underline{\hat{\mathbf{X}}}_{\rm c}|\hat{\mathbf{X}}_0]}}_{=s^{{\rm c}|0}/k_{\rm B}}. \tag{9.54}$$

Further, by exploiting our Markovian representation of the process, we can show that the fluctuations of Δs^0 fulfill an integral fluctuation theorem, as do those of $\Delta s_{\rm tot}$ (see Sec. 3.4, Fig. 9.10),

$$\begin{aligned}\left\langle e^{-\frac{\Delta s^0}{k_{\rm B}}} \right\rangle &= \int .. \int \frac{\hat{\mathcal{P}}[\hat{\mathbf{X}}_0]}{\mathcal{P}[\mathbf{X}_0]} \mathcal{P}[\underline{\mathbf{X}}] \, \mathfrak{D}\underline{\mathbf{X}} \\ &= \int \hat{\mathcal{P}}[\hat{\mathbf{X}}_0] \underbrace{\int \mathcal{P}[\mathbf{X}_1,..,\mathbf{X}_n|\mathbf{X}_0] \, \mathfrak{D}\mathbf{X}_1..\mathfrak{D}\mathbf{X}_n}_{=1} \, \mathfrak{D}\mathbf{X}_0 = \int \hat{\mathcal{P}}[\hat{\mathbf{X}}_0] \, \mathfrak{D}\mathbf{X}_0 = 1.\end{aligned} \tag{9.55}$$

This implies $\Delta S^0 \geq 0$. The same can be shown for the other contribution in (9.54), i.e., the entropy-like quantity $\Delta s^{\mathrm{c}|0}$, because

$$\left\langle e^{-\Delta s^{\mathrm{c}|0}/k_{\mathrm{B}}} \right\rangle = \iint \frac{\hat{\mathcal{P}}[\underline{\hat{\mathbf{X}}}_{\mathrm{c}}|\hat{\mathbf{X}}_0]}{\mathcal{P}[\underline{\mathbf{X}}_{\mathrm{c}}|\mathbf{X}_0]} \mathcal{P}[\underline{\mathbf{X}}]\, \mathfrak{D}\underline{\mathbf{X}}$$
$$= \int \mathcal{P}[\mathbf{X}_0] \left[\int \hat{\mathcal{P}}[\underline{\hat{\mathbf{X}}}_{\mathrm{c}}|\hat{\mathbf{X}}_0]\, \mathfrak{D}\underline{\mathbf{X}}_{\mathrm{c}} \right] \mathfrak{D}\mathbf{X}_0 = \int \mathcal{P}[\mathbf{X}_0]\, \mathfrak{D}\mathbf{X}_0 = 1. \quad (9.56)$$

These considerations yield the inequality

$$\Delta S_{\mathrm{tot}} \geq \Delta S^0 \geq 0. \quad (9.57)$$

Thus, the irreversibility of X_0 alone underestimates the irreversibility of the entire system, consistent with what one would expect [126]. Furthermore, (9.57) states that the mean total EP considered in this thesis yields an upper bound to the average of the irreversibility measure Δs^0.

Importantly, the fluctuation theorems can only be shown, when the ensemble average $\langle ... \rangle = \iint ... \mathcal{P}[\underline{\mathbf{X}}] \mathfrak{D}\underline{\mathbf{X}}$ is evaluated using the *full* path probability $\mathcal{P}[\underline{\mathbf{X}}]$, while averages with respect to marginalized measures do not yield the correct relations.

By considering Δs^0, we have addressed entropy-like quantities based on marginalized path probabilities. An obvious alternative to get rid of the additional d.o.f. $X_{j>0}$ would be to study instead the marginalized EP: $\iint \Delta s_{\mathrm{tot}}[\underline{\mathbf{X}}]\, \mathfrak{D}\mathbf{X}_1 ... \mathfrak{D}\mathbf{X}_n$. We stress that this quantity is different from Δs^0 [as can be seen from Eq. (9.54)] and there is no direct relation between both. The relevance of such a marginalized EP is, in fact, questionable. It does *not* solely rely on measurable quantities (as opposed to Δs^0), neither does it have a direct link to the well-understood framework of statistical mechanics.

In contrast, the total EP Δs_{tot} is the central quantity in statistical mechanics with a direct connection to the energy flows, the first and second law of thermodynamics, and the underlying microscopic laws of motion [350].

In this Chapter, we have investigated thermodynamic and information-theoretic notions of systems with time-delay. We have focused on a specific question relevant in this context, that is, the consequences of non-reciprocal coupling. By application of standard definitions, we were able to obtain a consistent picture, and gained some interesting insights. However, concerning the underlying bigger problem, various important open questions remain.

In the following, last Part of this thesis, we will give an outlook on some further objectives of future research.

Part IV

Concluding remarks

"Thermodynamik ist ein komisches Fach!
Das erste Mal, wenn man sich damit befasst, versteht man nichts davon. Beim zweiten Durcharbeiten denkt man, man hätte nun alles verstanden, mit Ausnahme von ein oder zwei kleinen Details. Das dritte Mal, wenn man den Stoff durcharbeitet, bemerkt man, dass man fast garnichts davon versteht, aber man hat sich inzwischen so daran gewöhnt, dass es einen nicht mehr stört."

– Arnold Sommerfeld, as quoted in [351]

In this last Part IV, we summarize the main results presented in this thesis. We further put our ideas into a wider context, and critically challenge our findings. At the end, we draw the attention of the Reader to open questions building upon our results, and give some perspectives for future research on stochastic systems with time delay.

10 | Summary

This thesis is concerned with the non-Markovian dynamics of stochastic, fluctuation-dominated systems with time delay. The prime example studied here is a colloidal (Brownian) particle subject to a position-dependent, time-delayed feedback force and nonlinear static forces stemming from a potential. We describe its noisy motion by a Langevin equation, which is, due to the delay, non-local in time and infinite-dimensional. Being experimentally realizable and analytically treatable (by means of comparably simple equations), this system can be considered as a paradigm for a non-Markovian, nonequilibrium system.

While nonequilibrium systems are in general a major focus of modern statistical physics, the incorporation of time delays, or generally memory, in the theoretical description is still in its infancy. We have focused on two concepts that play a central role in statistical physics: the probabilistic treatment and thermodynamic notions. Both concepts are, so far, poorly understood for systems with time delay, especially in the presence of nonlinearities.

Probabilistic treatment

First, we have discussed the probabilistic treatment of delayed stochastic systems. It was already known that, in the presence of delay, the Fokker-Planck equation for the one-time probability density is not closed, but explicitly involves the two-time (joint) probability density. In fact, an infinite hierarchy of coupled Fokker-Planck equations emerges [1, 87, 116], reflecting the infinite-dimensionality of the underlying Langevin equation.

While this was, in principle, already known, we could not find explicit expressions for the higher members of this hierarchy in the literature. One goal of this thesis was to investigate these higher members. As a starting point, we have reviewed an earlier derivation based on Novikov's theorem, and extended this approach to derive the second member of the Fokker-Planck hierarchy. Secondly, we have suggested a novel strategy, by means of a *Markovian embedding*, from which we derive an alternative representation of the Fokker-Planck hierarchy.

Let us summarize the idea and technical details of this embedding procedure here. Mainly, one changes the system state from $X(t)$, which has a non-Markovian dynamics, to the entire interval $\{X(s)\}$, $s \in [t-\tau, t]$, which has a Markovian dynamics (τ being the delay time). More precisely, by introducing n additional (auxiliary) variables, which represent n timesteps between the times t and $t-\tau$, we embed the one-variable delayed process $X(t)$ in an $(n+1)$-variable space

$\{X_0(t), X_1(t), ..., X_n(t)\} \hat{=} \{X(t), X(t - \mathrm{d}t), ..., X(t - \tau)\}$ ($\mathrm{d}t$ corresponds to τ/n). The original delay process is represented as one variable $X_0 \hat{=} X$ in this $(n + 1)$-dimensional space. Importantly, this new representation is only exact in the limit $n \to \infty$. This is because the original dynamics is time-continuous, thus, there are infinitely many timesteps between t and $t-\tau$. Considering the equations of motion, this "blown up" Markovian representation is given by a ring network with $n+1$ nodes which are unidirectionally coupled among each other, and follow linear equations, even if the original process itself is nonlinear.

Due to the Markovianity, this $(n+1)$-variable process has a corresponding closed, $(n + 1)$-variate Fokker-Planck equation. By projecting this Fokker-Planck equation onto a lower dimensional subspace (e.g., onto X_0 to derive the first member), and taking the limit $n \to \infty$ in the projected equations, we derived a hierarchy of Fokker-Planck equations. While the first member is identical, the representation of the higher members derived in this way differs from the one obtained by Novikov's theorem. This is because the underlying mathematical representation is different. Specifically, the Novikov-approach is based on an interpretation of the delay process where $\{X(t), X(t - \tau), X(t - 2\tau), ...\}$ are the relevant variables, rather than $\{X(t), X(t - \mathrm{d}t), ..., X(t - \tau)\}$ like here.

The main theoretical advancements due to this thorough investigation are the following two aspects: On the one hand, we have derived explicit expressions for the higher Fokker-Planck hierarchy members, on the other hand, we have gained a more profound understanding of the mathematical structure of the delay process.

Next, we have turned to a more practical issue, that is, the development of suitable *approximation* schemes. At this point we should recall that, by not being closed, the Fokker-Planck equation for the one-time probability density alone is insufficient to find solutions for the probability density (an exception being the case of linear delay Langevin equations). Specifically, we have asked ourselves in the beginning whether it is possible to find a scheme, which yields an approximate, simplified, yet non-Markovian, description; contrasting previous approaches from the literature, which effectively render Markovian descriptions. The two most established earlier strategies are the "small delay expansion", which is a Taylor expansion of the Langevin equation in τ [82, 83], and a perturbation theory on the Fokker-Planck level [86, 87], where the entire delay force is treated as a small deviation from the Markovian dynamics. To this end, we have here proposed a new scheme, called the "force-linearization closure". This is a closure of the infinite Fokker-Planck hierarchy, achieved by linearizing the total deterministic force in all members of the hierarchy starting from the second. In this way one obtains an approximate steady-state one-time probability density of nonlinear delay systems.

Applying the "force-linearization closure" to two exemplary model systems, a bistable and a periodic system with linear delay force, we present numerical evidence that our approach generates indeed a very good approximation of the one-time probability density. In fact, the force-linearization closure performs better than the small delay expansion and the perturbation theory. We also note that the perturbation theory has a further technical disadvantage as it actually relies on an analytical expression for the conditional probability for the corresponding system without delay

force, which is usually not available and requires additional approximations.

A further aspect touched upon is the calculation of transport-related quantities, such as escape times or waiting time distributions. For non-Markovian processes, an exact calculation would require analytical expressions of the two-time probability or even higher n-time densities. The force-linearization closure does not provide access to these probabilistic quantities. Still, we have demonstrated that a reasonable estimate of the escape times is possible by combining the force-linearization closure and the Kramers theory. This estimate, however, breaks down when the delay times are in the range of (or larger than) the jump duration times. Under such conditions, the interplay between the dynamics of the system state variable and the time-dependent energy landscape causes oscillatory motion, i.e., stochastic particle oscillations between two adjacent minima, whose description is clearly beyond this approximation. This exemplifies that an extension of the theory towards higher members of the Fokker-Planck hierarchy would be worthwhile.

Thus, while we have achieved our goal of finding a "true non-Markovian approximation" of the one-time probability density, we have actually learned along the way that the insights provided by this probabilistic quantity alone are somewhat limited. This is due to the fact that the dynamics is crucially determined by the memory, and, subsequently, the temporal correlations, particularly between the times t and $t - \tau$. The latter are given directly by the two-time probability density, which, in turn, is the solution of the second Fokker-Planck hierarchy member. In this spirit, we have reconsidered the second member of the Fokker-Planck hierarchy, which we have derived by means of the Markovian embedding technique. Based on this equation, we have introduced an approach that yields an approximate two-time probability density. This approximation finally enables us to describe truly non-Markovian, dynamical effects. We have demonstrated this by again considering the delay- and noise-induced interwell oscillations in our bistable example system. In particular, the proposed scheme renders very good results, for both cases, feedback forces that are directed towards, or away from the delayed position, capturing the main features of the two-time probability density related to the oscillations, in the regime of large delay times. Moreover, the corresponding approximate one-time probability density (which can be obtained by marginalization) outperforms the approaches discussed before (small delay, perturbation theory, force-linearization closure). To the best of our knowledge, this is the first approach to calculate the two-time probability density, which is indeed a central quantity in the probabilistic description of stochastic systems with time delay.

Thermodynamics

The second main topic of this thesis is the search for physically consistent thermodynamic descriptions of systems with time delay. The generalization of the concepts of (stochastic) thermodynamics towards non-Markovian systems turns out to be a major problem that is not addressed in its full extent within the scope of this work. However, by pointing out fundamental theoretical issues, deriving some concrete solutions for the heat flow, discussing information flows and the total entropy production, and further adding some new ideas and arguments, we hope that our

investigations have advanced the theoretical understanding of this thriving new field in statistical physics. Let us summarize our research questions and main findings.

To begin with, we have considered the net steady-state heat flow induced by a delay force focusing on nonlinear systems (with polynomial potentials). To gain analytical access to this quantity, we have first derived a formula for the heat rate that only involves positional moments, and can be combined with the approximations for the one-time probability density discussed before. Thus, despite the inherent memory, no expressions for the temporal correlations are needed to evaluate the mean heat flow.

Initially, we have asked ourselves whether the feedback may induce a negative heat flow, corresponding to a steady energy extraction from the heat bath. We have seen that such a reversed heat flow is indeed encountered in linear systems as well as nonlinear systems. As a concrete example, we have again considered the bistable model with linear delay force. Furthermore, we have presented analytical results for the Markovian limits and thereby observed a *growing* net heat flow when the delay time is reduced towards zero. We have argued that this pronounced heat flow at small τ above and around the (ballistic) velocity-relaxation timescale, is a consequence of the interplay of noise and delay linked to the phenomenon of "entropy pumping". To be more specific, in this regime of small delay times, the feedback becomes effectively dependent on the *velocity*. Further on, velocity-dependent forces are known to automatically induce heat flows, as they "pump entropy into the system".

Another point that we wanted to clarify concerns the consequences of the overdamped limit in the thermodynamic description. Indeed, we found that, when the inertia term is neglected in the Langevin equation, the regime of τ values below the (ballistic) velocity-relaxation timescale becomes problematic. In particular, we have uncovered discontinuous behavior in the limit $\tau \to 0$, which is absent in the corresponding underdamped description. The discontinuity in the limit of vanishing delay times is a consequence of the white noise assumption in combination with the overdamped limit. From a technical perspective, this is a notable result of our investigation, as the limit of small delay times, is, in fact, a very common regime to look at in the context of dynamical and probabilistic quantities of overdamped systems. As an example, recall the small delay expansion, which is a quite conventional approach to approximate (overdamped) delay systems. On the level of dynamical and probabilistic quantities, this limit is usually well-behaved for overdamped and underdamped descriptions, and rather unproblematic. Here we see that the thermodynamic quantities are more sensitive and one indeed needs to be careful with combining different approximations (white noise, the overdamped limit, the limit of small delay). Physically, we suspect that the reason for the apparent offset at $\tau \to 0$ is connected to the fact that the overdamped limit implies that the kinetic energy is not described. In turn, the kinetic energy, would yield additional contribution in the energy balance and could thus affect the calculated heat flow.

For experimental realizations of feedback traps these theoretical results imply that the (unavoidable) delay causes a finite steady-state heat flow unless the delay is significantly smaller than the (ballistic) relaxation timescale. This might, for example, be relevant in the following context. Feedback loops offer a way to precisely

control the effective temperature of a particle by suppressing its thermal fluctuations [352]. When this technique is used to change the temperature, e.g., during the cyclic process of a heat engine [353], the heat flow discussed here, which is caused even by tiny delays, might play a non-negligible role and should therefore be taken into account in the theoretical calculations of efficiency and entropy production[1].

We have further explicitly shown that the direction of induced heat flow is tunable by adjusting the delay time and the (positive) strength of the feedback force (which would correspond to the laser intensity in an experimental setup with an optical trap). This might be a useful feature in view of possible applications.

Lastly, another interesting result of this investigation concerns the heat flow and the closely connected medium entropy production in the parameter regimes where the delay-induced interwell oscillations are prominent. As was already known, these oscillations are especially pronounced and regular around a certain finite temperature, a phenomenon which is called *coherence resonance* [10, 40, 256, 257, 266, 354–356]. Here we have shown that the medium entropy production has a maximum right at the point where the oscillations are coherent-resonant; linking a nonlinear-dynamical effect to a thermodynamic phenomenon. By combining a discretized (Master) equation approach, called 2-state reduction [256], with the framework of stochastic energetics, we have found a way to approximate the thermodynamic quantities in this parameter regime and further provided an analytical explanation for this maximum, which was previously only detected numerically [266].

At the end of this thesis, we have turned to the question of how to quantify the "irreversibility" of delay processes, i.e., their distance from thermal equilibrium. For Markovian systems, the total entropy production provides the entire entropy balance of the process. Thereby, the total entropy production of a single fluctuating trajectory can be calculated based on the path probabilities of the observed trajectory and the path probability of its time-reversed picture in the backward process. This is a very interesting concept directly linking the notion of irreversibility and the thermodynamic arrow of time to the laws of thermodynamics. However, a generalization of this concept to *non-Markovian* systems is not straightforward, and in fact subject of many recent debates in stochastic thermodynamics.

Leaving aside the technical and conceptual challenges implied by the non-Markovianity, we have here pursued the following question: What do we learn about the thermodynamics of the time-delayed process, if we study the thermodynamic properties of the corresponding Markovian network representation with degrees of freedom $\{X_0(t), X_1(t), ..., X_n(t)\}$ obtained by the Markovian embedding?

There is also another way to tell the story. One could also say that we have now included the feedback-controller explicitly into our model. In particular, we have added n additional variables $\{X_1(t), ..., X_n(t)\}$ representing the "memory cells" of the controller, and in this way, obtained a Markovian description for a specific

[1] Although the investigations presented here are restricted to static energy landscapes, we indeed expect the results to also hold for time-dependent external potentials, at least, if their changes are slow as compared to the other dynamics.

representation of the feedback controller. An important feature of the memory device is the *non-reciprocal* coupling between the cells, which corresponds to a "shift register architecture" and reflects that the updates of the device are a one-way operation in time.

We have noticed that such non-reciprocal interactions between mesoscale degrees of freedom are, in fact, also common in the context of *active* systems, examples being the active Ornstein-Uhlenbeck particle model [105, 107, 109, 307, 308], or the cellular sensor model [322, 323]. Indeed, this immediately matches our intuition. The state of one component (the "sensor"), say X_1, depends on the state of the other component ("the measured system"), say X_0, but not *vice versa*. In reverse, in the case of active self-propelled motion, the temporal evolution of the particle's position is assumed to be affected by the noisy orientation, but there is no backcoupling.

Thus, non-reciprocal coupling is an important and quite universal concept, motivating us to discuss the connections between coupling topology and thermodynamic properties of the model in a general sense. We have shown that a single non-reciprocal interaction already implies that the system is intrinsically out of equilibrium as indicated by a positive total entropy production and a broken fluctuation-dissipation relation. A single unidirectional interaction can be viewed as an active "propulsion force", which will ultimately always have a heating effect on the environment, and never cool it down. A reversed unidirectional interaction can be considered as a "sensing" operation, inducing a net information flow from the measured system to the sensor. Furthermore, feedback can be viewed as a combination of measurement and active force. When the measurement is the dominating operation (i.e., the controller "knows" more about the sensed system than vice versa, indicated by a positive information flow out of the system) the following goals of feedback control can be achieved: thermal fluctuation suppression (i.e., feedback cooling), and energy extraction from the heat bath (a reversed heat flow). Taken together, from our analysis emerges a unifying perspective on activity, feedback and memory.

For the special case of discrete delay with white noise, which emerges as the limit of an infinitely large ring network, we confirm that the heat exchange of the system of interest, X_0, approaches the expected result for linear white noise delay Langevin equations. On the contrary, the total entropy production diverges in this limit, suggesting that the cost of storing a Brownian trajectory (of arbitrary length) is unbounded. We can understand this from an information-theoretic perspective. Recall that the white noise is a process with zero temporal correlation (yielding infinitely many jumps on each trajectory, no matter how short it is), thus, resulting in an unbounded information content.

An interesting way to think about the approach discussed in this thesis is that it points out a connection between irreversibility and symmetry-broken coupling topology. The entropy production, which gives the *thermodynamic arrow of time*, directly measures the breaking of time-reversal symmetry, i.e., how sure can we be that a system's evolution is running from past to future, and not vice versa. In a way, memory breaks that symmetry in a fundamentally different way than, e.g., an external field, because memorizing is *per se* only one-way in time.

11 | Outlook – Open questions and further perspectives

We close with some comments putting the findings presented in this thesis into a broader perspective and giving an outlook on future research. We begin with formulating a few obvious follow-up questions and take a wider view towards the end.

Probability densities In this thesis, we have derived explicit equations for the second member of the Fokker-Planck hierarchy, and based thereon formulated a new approximation scheme, by relatively crude simplifications. This approximation is much simpler as compared to the multi-step procedure of the force-linearization closure, but still yields better results in the parameter regime where the particle motion is dominated by the interplay of nonlinearities, noise and memory. Thus, we have seen that an approximation scheme that operates on this second member of the Fokker-Planck hierarchy is, in some sense, superior as opposed to approaches based on the first one. This is due to the fact that the starting equation is already more sophisticated and, in particular, explicitly addresses the temporal correlations essential to capture non-Markovian behavior.

For this reason, we think it would be worthwhile to further explore the approximation proposed here or alternative approaches aiming for the two-time probability density, or other quantities that directly describe the temporal correlations. To this end, it might also be beneficial to inspect even higher members of the Fokker-Planck hierarchy.

Fluctuations An obvious follow-up question concerns the *fluctuations* of the thermodynamic quantities. In this thesis, we have mainly discussed the mean rates. However, another central topic in stochastic thermodynamics is the description of the fluctuations around those mean values. For Markovian systems, the probability distributions in non-equilibrium steady states are already quite well-understood, an example being the knowledge of several universal fluctuation theorems [124, 125]. Corresponding statements for general systems with memory have not been developed so far. Only for some specific non-Markovian processes, fluctuation theorems have been found in the past, e.g., [138, 139, 143–146], including linear Langevin equations with delay [116, 141]. However, we could not find general results applicable to the delay processes considered here, especially not for nonlinear systems.

Our preliminary, numerical investigations have shown that the heat distributions are non-Gaussian, and violate fluctuation relations. Deriving analytical expressions for the probability distributions, and searching for universal symmetry properties of the heat fluctuations are open problems for future research.

Quantum effects Stepping away from the (classical) colloidal example, and reconsidering the stochastic delay processes and the meaning of non-reciprocal coupling, one notices some connections to quantum systems. Indeed, also in quantum systems, researchers have recently started to analyze, on the one hand, feedback with memory [45, 357], and, on the other hand, non-reciprocal interactions [315, 316].

In the present work, we have developed a unifying viewpoint on these two concepts, delayed feedback and non-reciprocity, for classical (stochastic) models. It would be very interesting to generalize this perspective to the quantum world, where entropy, information and energy play an analogous role, while measurement and stochasticity have different notions.

Uniqueness of the Markovian embedding In this thesis, we have employed a Markovian embedding to represent the time delayed process by means of a Markovian network, specifically, a unidirectionally coupled ring network, where all variables but X_0 follow linear equations of motion. Let us rethink the Markovian embedding and the approach to thermodynamics of delayed systems taken in this thesis.

Focusing on such linear networks, we have shown that non-reciprocal interactions are an essential ingredient to generate non-monotonic memory kernels, and are thus essential to represent the memory kernels characteristic for delayed feedback. However, one important question that we have not yet addressed in this context, is the *uniqueness* of the Markovian (network) representation: Is it possible to find the exact same memory in another network of different size, coupling topology, or with another type of dynamics? This question is particularly relevant in situations, where only some of the degrees of freedom are *observable*, e.g., only X_0, while all others, $\{X_1, ..., X_n\}$, are "auxiliary" variables without physical interpretation.

We indeed think that the representation is *not* unique. In fact, exploring more complex network architectures (with nonlinear couplings, other coupling topologies, or nonlinear local dynamics)is an interesting question and shall be subject of future work. However, by considering linearly coupled, overdamped nodes with linear dynamics, we have, in fact, studied the most simple possibility.

In view of the approach to thermodynamics of delayed systems taken in this thesis, one should clarify whether and how the calculated entropic and information-theoretic quantities depend on the (probably ambiguous) representation. We aim to stress that we already know that the heat flow is independent (which follows from the fact that it can equivalently be derived from the non-Markovian Langevin equation alone). However, the entropic picture and information flows will probably not be.

We have also noted the possibility to build something like an effective thermodynamics upon our findings. For example, one may consider the total entropy production $\dot{S}_{\rm tot}(X_0, X_1, ..., X_n)$ of the "super-system" consisting of all degrees of free-

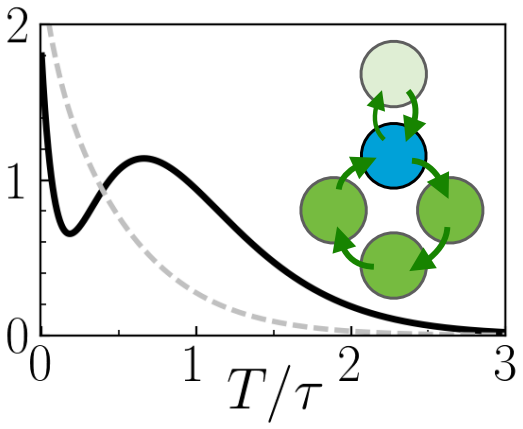

Figure 11.1: *Sketch of a Markovian network of a colloidal particle (blue disk) coupled to four (auxiliary) degrees of freedom (green disks). Due to the coupling, the corresponding non-Markovian, one-variable Langevin equation for the colloid features a (bimodal) memory kernel and a colored noise with monotonically decaying correlations (gray line). Also compare with Fig. 5.2, to see the corresponding kernels and noise correlations for simpler coupling topologies (specifically, for ring networks with 2,3,4 nodes).*

dom, $\{X_0, X_1, ..., X_n\}$, and then in a second step trace out all "auxiliary" degrees of freedom $\{X_1, ..., X_n\}$. It will be particularly interesting to see whether the resulting effective "irreversibiltiy measure" will be independent on the chosen Markovian representation, or not.

A conceptionally interesting aspect regarding the way of modeling memory employed in this work, is that it can be considered as a "discretization of the past" [3]. This is a delicate point, as the dynamics itself is time-*continuous*. As opposed to that, a discretization of the entire dynamics was proposed in [358]. Surprisingly, the approach taken here yields a simpler thermodynamic framework, despite the apparent contradiction of its discrete–continuous nature.

Further physical systems with memory and activity Finally, we aim to discuss the Markovian embedding in view of possible further applications. So far, we have seen that a unidirectionally coupled ring network, where all auxiliary variables follow linear equations of motion, automatically generates a Gamma-distributed kernel and monotonically decaying colored noise (recall Fig. 5.2 from the Part II). Indeed, besides yielding discrete (delta-distributed) delay in the limit of an infinitely large network, such Gamma-distributed delays themselves have many applications. For example, in biological contexts they appear in gene regulatory networks [55–58], or in the dynamics of virus spreading between cells [359]. Furthermore, by combining multiple Gamma-distributed kernels, one can obtain models for a very general type of memory. From the perspective of the Markovian network, this amounts to generalizing the coupling topology. As a concrete example, we have depicted in Fig. 11 the (bimodal) kernel and noise correlations arising from a linear, unidirectional ring network and an additional, separately coupled degree of freedom. As a matter of fact, any kernel that eventually decays exponentially can be approximated in this way. We find that a memory kernel with N extrema can be represented by N non-reciprocally coupled subsystems (while reciprocally coupled degrees of freedom, always yield monotonically decaying memory kernels).

This further implies that more general types of stochastic processes with memory can, in principle, be treated with this approach. For example, one may think of other sources of memory like viscoelasticity of the surrounding fluid, and hydrodynamic

backflow [16]. While there exist already sophisticated theories to treat these two kinds of memory (e.g., the Basset force for hydrodynamic backflow [360, 361]), the network representation employed here yields a simplistic, effective description with relatively few variables. Furthermore, the established theories usually describe systems that can (in the absence of external time-dependent drivings) reach thermal equilibrium. However, in various real-world scenarios, the heat bath does not only have memory, but at the same time is an active environment. Think for example of biocellular processes, where motor proteins drive actin filaments [362], creating a highly active and complex environment with memory effects. Having all this in mind, it appears to be an interesting and fruitful direction of future research to investigate whether a low-dimensional Markovian network with non-reciprocal couplings can be used as a simplistic model for the motion of a colloid in such a complex, active surrounding that is far from thermal equilibrium.

Computer science The Markovian network representation and the information-thermodynamic analysis from Part III might further be interesting from the viewpoint of computer science. On an abstract level, our thermodynamic investigation might be useful to incorporate the thermodynamic cost of memory in the theoretical description of computation in living [363] and artificial systems [276]. A more concrete link is the similarity between the ring network architecture studied here and the reservoir computers investigated in [364, 365]. A reservoir computer of this type may be experimentally realized by a laser network [366, 367], or by coupled RC circuits [171, 368]. The ring network is further very similar to a recurrent neural network [369], used for example for reinforcement learning. In these contexts, the connection between non-reciprocal coupling and information flow discussed here might be of particular interest. Notably, the topology of the unidirectionally coupled ring considered here also resembles the architecture of a Brownian clock [370], which, in contrast, has discrete dynamics.

Collective behavior Lastly, we want to come back to an initial thought formulated at the beginning of this thesis. We have considered (colloidal) systems subject to time-delayed feedback. We have seen that these particles display complex dynamical behavior already on the level of individual particles. Therefore, one shall expect intriguing collective effects in a collection of multiple systems with delay (or delayed interactions). In this thesis, we have pointed out fundamental connections between the thermodynamic properties of individual active and feedback-controlled systems with memory. Given these connections, one should also expect common aspects of their collective behaviors. For active systems, particularly microswimmer suspensions [20, 102, 103, 371], clustering and pattern formation, are well-known phenomena. Indeed, recent studies report similar phenomena in collections of feedback-controlled systems [38, 66, 89, 113].

Reconsidering this from a broader perspective, we have argued that within the framework of modern statistical physics, single colloids under delayed feedback can be considered as a prototype of a "novel type of particle". Representing important phenomena of the complex world, which are, so far, not incorporated in the frame-

work of statistical physics, like feedback, intrinsic energy and information flows, they are suitable building blocks for the construction of new, generalized theories. Understanding the (thermo-)dynamic properties of individual constituents is a first step in this direction.

A | Appendix

A.1 Numerical methods

In this appendix, we lay out all relevant information for the specific simulations used to generate the results presented in this thesis. We refer the Reader who is interested in detailed descriptions and background information about simulation technique for stochastic systems to [181], where extensive discussion can be found.

In order to test our theoretical results against quasi-exact data, we perform Brownian dynamics (BD) simulations of the delayed LE within C++. We use the Euler-Maruyama integration scheme [181, 249], with a varying temporal discretization (see below). We perform each simulation multiple times (with different random number seed), and take the average over at least $N = 10^4$ realizations (usually more, as occasionally stated in the main text). The (pseudo) random numbers are generated with the algorithm "Mersenne Twister" [372] and the Box-Müller method [181].

Density profiles from the BD simulations are obtained from histograms with a spatial resolution of $\Delta x = 0.005\,\sigma$. In general, steady-state ensemble averages are obtained after appropriate transient times have been cut off. That means, before we measure, we let at least $100\,\tau_\mathrm{B}$ times and at least 20τ (whichever is higher) pass, in the presence of the delay force, such that the system has reached a steady state. The simulation time thereafter is more than $200\,\tau_\mathrm{B}$. Furthermore, steady-state ensemble averages are obtained by using a sufficiently high number of realizations N. Both is checked by insensitivity against further increase of simulated time or sample size.

We have tested different initial conditions, including equidistantly distributed, as well as delta-distributed ones, both with fixed position within the interval $[-\tau, 0]$. We found that an equilibrium configuration in the respective static potential without delay force often yields the fastest relaxation towards the steady state.

The temporal discretization is varied between $\Delta t = 10^{-6}$ and 10^{-4}, such that the typical time scales are all properly resolved. In particular, the delay time $\tau > 1000\Delta t$ and the intrawell relaxation time $\tau_\mathrm{ir} > 1000\Delta t$. Indeed, calculating properties of a system with very short delay times is therefore very expensive, as the step size must be extremely small. For example, in Fig. 8.3, where $\tau < 10^{-2}\tau_\mathrm{B}$, we have used $\Delta t = 10^{-6}$.

For the confining potentials (quadratic trap, doublewell potential), we did not have to pay special attention to the boundary conditions, as the probability to find the particle quickly vanishes for $|x| \gg 0$. We have just made sure that no particle may escape by stopping the simulations in case a position far away from zero occurs. On the contrary, in the case of the washboard potential, the particle is not confined. We have implemented a comoving simulation box. Note that whenever moving the box, the delayed particle position must be transformed accordingly.

A.2 Derivation of Novikov's theorem

In this appendix, we sketch a derivation of the Novikov's theorem (4.3), which reads

$$\left\langle\, \Lambda[\xi]\xi(t)\,\right\rangle = \left\langle \frac{\delta\Lambda[\xi]}{\delta\xi(t)} \right\rangle, \tag{A.1}$$

for a functional Λ of a *Gaussian white noise* ξ [327]. It links the functional derivative w.r.t. the noise to the cross-correlation between functional and noise. First, we consider the case where the ensemble is w.r.t. a *fixed* initial condition $\phi = X(-\tau \leq t \leq t)$, yielding $\Lambda(0) = \Lambda_0$. In a second step we will generalize towards initial conditions drawn from an arbitrary distribution.

We express the ensemble average $\langle..\rangle_{\Lambda_0}$ of the l.h.s. of (4.3) via the path integral over all possible paths ξ between the times 0 and t, accounting for all possible realizations of the random process ξ at each instant in time

$$\langle\Lambda[\xi]\xi(t)\rangle_{\Lambda_0} = \int_{\xi_0}^{\xi_t} \xi\, \Lambda\left[\xi|\Lambda_0\right]\ \mathcal{P}[\xi]\, \mathcal{D}[\xi], \tag{A.2}$$

with arbitrary but fixed $\xi_0 := \xi(t_0)$ and $\xi_t := \xi(t)$. Please note that specifying the noise process at the boundaries does *not* impose a restriction on the generality, as we deal with *white* noise. We recall that the weight $\mathcal{P}[\xi]$ of each white noise realization is given by the Gaussian path probability (for arbitrary $t > 0$)

$$\mathcal{P}[\xi] = \mathcal{J} e^{-\frac{1}{2}\int_0^t \xi(t')^2 \mathrm{d}t'}, \tag{A.3}$$

with Jacobian $\mathcal{J}$ [as already given in Eq. (1.4) in the main text, note that we there use bold symbols for paths, $\boldsymbol{\xi}$, different than here]. Now, we rewrite the integrand in (A.2) using that the functional derivative of $\mathcal{P}[\xi]$ w.r.t. ξ is simply $(-\xi)\mathcal{P}[\xi]$, yielding

$$\begin{aligned}\langle\Lambda[\xi]\xi(t)\rangle_{\Lambda_0} &= -\int_{\xi_0}^{\xi_t} \Lambda\left[\xi|\Lambda_0\right] \frac{\delta}{\delta\xi}\left\{ e^{-\frac{1}{2}\int_0^t \xi(t')^2 \mathrm{d}t'} \right\} \mathcal{D}[\xi] \\ &= -\int_{\xi_0}^{\xi_t} \frac{\delta\left\{\Lambda\left[\xi|\Lambda_0\right]\mathcal{P}[\xi]\right\}}{\delta\xi}\mathcal{D}[\xi] + \int_{\xi_0}^{\xi_t} \frac{\delta\Lambda[\xi]}{\delta\xi}\mathcal{P}[\xi]\,\mathcal{D}[\xi]. \end{aligned} \tag{A.4}$$

In the last step we have simply used the product rule. The functional derivative in the first path integral yields

$$\int_{\xi_0}^{\xi_t} \frac{\delta\left\{\Lambda\left[\xi|\Lambda_0\right]\mathcal{P}[\xi]\right\}}{\delta\xi}\mathcal{D}[\xi] = \int_{\xi_0}^{\xi_t} \Lambda\left[\xi+\delta\xi|\Lambda_0\right]\mathcal{P}[\xi+\delta\xi]\mathcal{D}[\xi] - \int_{\xi_0}^{\xi_t} \Lambda\left[\xi|\Lambda_0\right]\mathcal{P}[\xi]\mathcal{D}[\xi]. \tag{A.5}$$

Now we use that the variations of the paths $\xi + \delta\xi$ are already contained in the integral over *all* paths, and find

$$\int_{\xi_0}^{\xi_t} \frac{\delta\left\{\Lambda\left[\xi|\Lambda_0\right]\mathcal{P}[\xi]\right\}}{\delta\xi}\mathcal{D}[\xi] = \int_{\xi_0}^{\xi_t} \Lambda\left[\xi|\Lambda_0\right]\mathcal{P}[\xi]\mathcal{D}[\xi] - \int_{\xi_0}^{\xi_t} \Lambda\left[\xi|\Lambda_0\right]\mathcal{P}[\xi]\mathcal{D}[\xi] = 0, \tag{A.6}$$

(as long as $\int \Lambda[\xi]\mathcal{P}[\xi]\mathcal{D}[\xi] < \infty$), see also [373] (on p. 273f). Hence, we obtain from (A.4),

$$\langle\Lambda[\xi]\xi(t)\rangle_{\Lambda_0} = \int_{\xi_0}^{\xi_t} \frac{\delta\Lambda\left[\xi|\Lambda_0\right]}{\delta\xi}\mathcal{P}[\xi]\,\mathcal{D}[\xi] = \left\langle \frac{\delta\Lambda[\xi]}{\delta\xi} \right\rangle_{\Lambda_0}, \tag{A.7}$$

for a specific initial condition Λ_0.

This result can readily be generalized to ensembles where the initial conditions are instead drawn from an arbitrary, normalized distribution $P(\Lambda_0)$. In particular, Eq. (A.7) implies

$$\langle \Lambda[\xi]\xi(t)\rangle = \int \langle \Lambda[\xi]\xi(t)\rangle_{\Lambda_0} P(\Lambda_0)\mathrm{d}\Lambda_0 = \int \left\langle \frac{\delta\Lambda[\xi]}{\delta\xi}\right\rangle_{\Lambda_0} P(\Lambda_0)\mathrm{d}\Lambda_0 = \left\langle \frac{\delta\Lambda[\xi]}{\delta\xi}\right\rangle. \quad \text{(A.8)}$$

This is the relation (4.3), which is often referred to as *Novikov's theorem*.

A.3 Green's function method

In Sec. 4.2.2, we have derived a closed-form expression of the autocorrelation function using the method of steps. An alternative strategy to obtain analytical expressions for correlation functions is via the the Green's function method. To this end, one transforms the LE by Laplace-Transformation $\tilde{X}_j(\omega) = \int_0^\infty X_j(t)e^{-i\omega t}\mathrm{d}t$, which readily yields

$$\omega\gamma\tilde{X}(\omega) = -\alpha\tilde{X}(\omega) - \beta e^{-i\tau\omega}\tilde{X}(\omega) + \sqrt{2D_0}\tilde{\xi}(\omega) \quad \text{(A.9)}$$

$$\Rightarrow \tilde{X}(\omega) = \underbrace{\left(i\omega\gamma + \alpha + \beta e^{-i\omega\tau}\right)^{-1}}_{=\tilde{\lambda}(\omega)} \sqrt{2D_0}\tilde{\xi}(\omega). \quad \text{(A.10)}$$

with the Green's function in Fourier-space $\tilde{\lambda}(\omega)$. As we are mainly interested in steady states, we have set the entire history function to zero, to facilitate the calculation. In Ref. [89] it is shown how this calculation in principle can be generalized to arbitrary initial conditions. If initial conditions matter, it is more handy to use the Laplace-Transformation instead. Transforming back to real space this yields

$$X(t) = \int_0^t \lambda(t-t')\sqrt{2D_0}\xi(t')\mathrm{d}t', \quad \text{(A.11)}$$

describing $X(t)$ as a functional of the noise process $\boldsymbol{\xi} = \{\xi(0 \le t' \le t)\}$. In this form, the Green's function λ tells, how the system state X responds to the external perturbation, and is hence also called a (linear) response function.

From Eq. (A.11) one can deduce a well-known relationship between position-autocorrelation and Green's function. Due to its importance, and because we will later apply it (in Chapter 9), let us briefly show how one can derive it [238]. Consider $\Delta t \ge 0$,

$$\begin{aligned} C(\Delta t) = \langle X(t)X(t+\Delta t)\rangle &= (2D_0)\int_0^t \int_0^{t+\Delta t} \lambda(t-t')\lambda(t+\delta t-t'')\langle\xi(t')\xi(t'')\rangle\,\mathrm{d}t'\mathrm{d}t'' \\ &= \int_0^t \lambda(t-t')\lambda(t+\Delta t-t')\,\mathrm{d}t' = 2D_0\int_0^t \lambda(T)\lambda(T+\Delta t)\,\mathrm{d}T \\ &\overset{\text{ss}}{=} 2D_0\int_0^\infty \lambda(T)\lambda(T+\Delta t)\,\mathrm{d}T. \end{aligned} \quad \text{(A.12)}$$

In the last step we have specified to steady states. Now we use $\lambda(t<0) = 0$, and again

transform to Fourier-space

$$\begin{aligned}\tilde{C}(\omega) &= 2D_0 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \lambda(T)\lambda(T+\Delta t)e^{-i\omega\Delta t}\,\mathrm{d}T\,\mathrm{d}\Delta t \\ &= 2D_0 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \lambda(T)e^{i\omega T}\lambda(T+\Delta t)e^{-i\omega(\Delta t+T)}\,\mathrm{d}T\,\mathrm{d}\Delta t \\ &= 2D_0 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \lambda(T)e^{i\omega T}\lambda(T')e^{-i\omega T'}\,\mathrm{d}T\,\mathrm{d}T' = 2D_0\tilde{\lambda}(\omega)\tilde{\lambda}(-\omega) \end{aligned} \tag{A.13}$$

Transforming back to real space, this implies

$$C(\Delta t) = \frac{D_0}{\pi}\int_{-\infty}^{\infty} \tilde{\lambda}(-\omega)\tilde{\lambda}(-\omega)e^{-i\omega\Delta t}\,\mathrm{d}\omega. \tag{A.14}$$

While this strategy in principle yields analytical expressions for various (linear) systems, explicit closed-form solutions are only available for specific cases (which can, e.g., be numerically integrated), where the inverse Fourier-Transformation of the Green's function (A.10) is known, and the integral in Eq. (A.14) [see [89, 99] for some explicit results]. We could not find general closed-from solutions for the problem considered here.

A.4 Connection to Fokker-Planck hierarchy from Novikov's theorem

In this appendix, we consider the two representations of the second members of the Fokker-Planck hierarchy derived and discussed in Chapter 5. We recall that, while the FPE for ρ_1 obtained by our approach is equivalent to the first member of the FP hierarchy from Novikov's theorem (4.3), the second members differ. However, the apparent disagreement can be resolved. Indeed, as we show here, our result (5.44) can be transformed to the representation obtained from Novikov's theorem [given in (4.7)]. We focus on the term which differs, i.e., the drift term $\partial_{x_\tau}[\rho_2(x,t;x_\tau,t-\tau)\,\langle\dot{X}(t-\tau)|x=X(t),x_\tau=X(t-\tau)\rangle]$. To this end, we first use the definition of ρ_2 via delta-distributions, plug in the LE (1.42), which yields

$$\begin{aligned}&\partial_{x_\tau}\Big[\langle\dot{X}(t-\tau)|x=X(t),x_\tau=X(t-\tau)\rangle\rho_2(x,t;x_\tau,t-\tau)\Big] \\ &= \partial_{x_\tau}\Big\langle\dot{X}(t-\tau)\,\delta[X(t)-x]\delta[X(t-\tau)-x_\tau]\Big\rangle \\ &\overset{\text{LE}}{=} \partial_{x_\tau}\langle F[X(t-\tau),X(t-2\tau)]\,\delta[X(t)-x]\delta[X(t-\tau)-x_\tau]\rangle \\ &\quad + \partial_{x_\tau}\Big\langle\sqrt{2D_0}\xi(t-\tau)\,\delta[X(t)-x]\delta[X(t-\tau)-x_\tau]\Big\rangle. \end{aligned} \tag{A.15}$$

At this point, we have two correlations to deal with. The first one (involving F) can easily be simplified by expressing the ensemble average with the help of the PDF ρ_3, i. e.,

$$\begin{aligned}&\partial_{x_\tau}\Big\langle F\left[X(t-\tau),X(t-2\tau)\right]\delta[X(t)-x]\delta[X(t-\tau)-x_\tau]\Big\rangle \\ &\qquad = \partial_{x_\tau}\iiint_\Omega F(x_\tau,x_{2\tau})\,\delta[X(t)-x]\delta[X(t-\tau)-x_\tau]\times \\ &\qquad\qquad \rho_3(x,t;x_\tau,t-\tau;x_{2\tau},t-2\tau)\mathrm{d}x\mathrm{d}x_\tau\mathrm{d}x_{2\tau} \\ &\qquad = \partial_{x_\tau}\int_\Omega F(x_\tau,x_{2\tau})\rho_3(x,t;x_\tau,t-\tau;x_{2\tau},t-2\tau)\mathrm{d}x_{2\tau}. \end{aligned} \tag{A.16}$$

The remaining term on the right side of (A.15) (involving ξ) is simplified with the help of Novikov's theorem (4.3) (see Appendix A.2). Specifically, we define $\Lambda[\xi] := \delta[x - X(t)]\,\delta[x_\tau - X(t-\tau)]$, which yields

$$\begin{aligned}
\partial_{x_\tau}\Big\langle \sqrt{2D_0}\xi(t-\tau)\,\delta[X(t)-x]\delta[X(t-\tau)-x_\tau]\Big\rangle \\
&= \sqrt{2D_0}\partial_{x_\tau}\left\langle \xi(t-\tau)\Lambda[\xi]\right\rangle, \\
&= \sqrt{2D_0}\partial_{x_\tau}\left\langle \frac{\delta\Lambda[\xi]}{\delta\xi(t-\tau)}\right\rangle \\
&= \sqrt{2D_0}\partial_{x_\tau}\left[\left\langle \frac{\delta\Lambda[\xi]}{\delta X(t)}\frac{\delta X(t)}{\delta\xi(t-\tau)}\right\rangle + \left\langle \frac{\delta\Lambda[\xi]}{\delta X(t-\tau)}\frac{\delta X(t-\tau)}{\delta\xi(t-\tau)}\right\rangle\right] \\
&= -\sqrt{2D_0}\partial_{x_\tau}\left\langle \partial_x\delta[x-X(t)]\,\delta[x_\tau - X(t-\tau)]\frac{\delta X(t)}{\delta\xi(t-\tau)}\right\rangle \\
&\quad -\sqrt{2D_0}\partial_{x_\tau}\left\langle \partial_{x_\tau}\delta[x-X(t)]\,\delta[x_\tau - X(t-\tau)]\,\underbrace{\frac{\delta X(t-\tau)}{\delta\xi(t-\tau)}}_{=\sqrt{\frac{D_0}{2}}}\right\rangle. \qquad \text{(A.17)}
\end{aligned}$$

In the last step we have plugged in the definition of Λ, employed the chain rule, and used Eq. (4.4) (derived in the main text, Chapter 4). We note that the functional derivative of the trajectory X w.r.t. the *earlier* noise ξ which also appears in (A.17) has no such general solution, as discussed in Chapter 4. Based on these considerations, we can rewrite (A.17) as

$$\begin{aligned}
\sqrt{2D_0}\,\partial_{x_\tau}\left\langle \xi(t-\tau)\Lambda[\xi]\right\rangle = &-\sqrt{2D_0}\,\partial_{x_\tau}\partial_x\left\langle \frac{\delta X(t)}{\delta\xi(t-\tau)}\Big|_{\substack{X(t)=x\\X(t-\tau)=x_\tau}}\right\rangle \rho_2(x,t;x_\tau,t-\tau) \\
&- D_0\,\partial^2_{x_\tau}\rho_2(x,t;x_\tau,t-\tau). \qquad \text{(A.18)}
\end{aligned}$$

In combination with (A.15, A.16), we obtain the identity

$$\begin{aligned}
\partial_{x_\tau}\Big\langle \dot X(t-\tau)\,\delta[X(t)-x]\delta[X(t-\tau)-x_\tau]\Big\rangle \\
&= -\sqrt{2D_0}\partial_{x_\tau}\partial_x\left\langle \frac{\delta X(t)}{\delta\xi(t-\tau)}\Big|_{\substack{X(t)=x\\X(t-\tau)=x_\tau}}\right\rangle \rho_2(x,t;x_\tau,t-\tau) \\
&\quad - D_0\partial^2_{x_\tau}\rho_2(x,t;x_\tau,t-\tau) + \partial_{x_\tau}\int_\Omega F(x_\tau,x_{2\tau})\rho_3(x,t;x_\tau,t-\tau;x_{2\tau},t-2\tau)\mathrm{d}x_{2\tau} \\
&= \partial_{x_\tau}\left[\langle\dot X(t-\tau)|x=X(t),x_\tau=X(t-\tau)\rangle\rho_2(x,t;x_\tau,t-\tau)\right], \qquad \text{(A.19)}
\end{aligned}$$

proving that the second member of the here presented FP hierarchy, Eq. (5.44), is identical to the corresponding one (4.7) obtained from Novikov's theorem.

A.5 Fluctuation-dissipation relation for unidirectional ring of arbitrary length n

In this Appendix, we generally consider the existence of the FDR for a network of type (9.11) with arbitrary length n. The FDR is also discussed in the main text in Secs. 1.3.1

and 9.3.1. We need to show the equivalence of the following two terms

$$k_{\mathrm{B}}\mathcal{T}_0\gamma(\Delta t) = k_{\mathrm{B}}\mathcal{T}_0\left[2\gamma_0\,\delta(\Delta t) + \frac{k\Gamma\left(n,\frac{n\Delta t}{\tau}\right)}{(n-1)!}\right] \tag{A.20}$$

and

$$C_\mu(\Delta t) = k_{\mathrm{B}}\mathcal{T}_0\,2\gamma_0\delta(\Delta t) + \frac{k_{\mathrm{B}}k^2}{\gamma'}\sum_{p=0}^{n-1}\sum_{l=0}^{p}\mathcal{T}'\frac{2^{l-2p}(2p-l)!}{p!l!(p-l)!}\frac{e^{-n\Delta t/\tau}\Delta t^l}{(\tau/n)^{l-1}}. \tag{A.21}$$

The first summands on the right hand side of (A.20) and (A.21) are already equivalent. Further, the other term in (A.20) an be rewritten by partially integrating the incomplete Gamma function as

$$\Gamma\left(n, n\Delta t/\tau\right) = (n-1)!\sum_{p=0}^{n-1}\left(\frac{n\Delta t}{\tau}\right)^p\frac{e^{-n\Delta t/\tau}}{p!}. \tag{A.22}$$

Thus, the FDR holds, if at all $\Delta t \geq 0$

$$\sum_{p=0}^{n-1}\sum_{l=0}^{p}\frac{\mathcal{T}'}{\mathcal{T}_0}\left[\frac{k\,2^l(2p-l)!}{n\,2^{2p}l!(p-l)!}\right]\frac{\tau(n\Delta t)^l e^{-n\Delta t/\tau}}{\gamma'\,\tau^l p!} \stackrel{?}{=} \sum_{p=0}^{n-1}\frac{(n\Delta t)^p e^{-n\Delta t/\tau}}{\tau^p p!}. \tag{A.23}$$

At $\Delta t = 0$ Eq. (A.23) is fulfilled, if and only if

$$\mathcal{O}(\Delta t^0): \quad \frac{\gamma'}{\tau} \stackrel{?}{=} \sum_{p=0}^{n-1}\frac{\mathcal{T}'}{\mathcal{T}_0}\frac{k(2p)!}{n\,2^{2p}(p!)^2}. \tag{A.24}$$

This is the zero$^{\text{th}}$-order condition (in Δt). For $n = 1$, Eq. (A.24) reads $k\tau\mathcal{T}' = \gamma'\mathcal{T}_0$ [in agreement with (9.16)], and is a *sufficient* condition. It can only be fulfilled for $k \geq 0$. Further, for $n > 1$, it yields a *necessary*, but not sufficient condition due to higher order contributions in Eq. (A.23). For example, the first order in Δt, which is relevant for all $n > 1$, yields the additional condition

$$\mathcal{O}(\Delta t^1): \quad \frac{\gamma'}{\tau}(n\Delta t) \stackrel{?}{=} \frac{\mathcal{T}'}{\mathcal{T}_0}\frac{k}{2\,n}(n\Delta t). \tag{A.25}$$

Combining this with (A.24) gives

$$\sum_{p=0}^{n-1}\frac{\mathcal{T}'(n\Delta t)}{\mathcal{T}_0}\frac{2(2p)!}{2^{2p}(p!)^2} \stackrel{?}{=} \frac{\mathcal{T}'(n\Delta t)}{\mathcal{T}_0}, \tag{A.26}$$

which can only be fulfilled if $\mathcal{T}' = 0$. However, at $\mathcal{T}' = 0$ condition (A.24) is violated. Hence, the FDR is broken for all choices $\mathcal{T}'$ when $n > 1$.

We have, in fact, also investigated a generalization of (9.11), where all temperatures are allowed to be different, i.e., X_j has temperature $\mathcal{T}_j$ (generalizing the derivations presented in Sec. 5.1.1 to this case is straightforward). Even in this case, the FDR cannot be fulfilled if $n > 1$.

Bibliography

[1] Loos, S. A. M. and Klapp, S. H. L. *Force-linearization closure for non-Markovian Langevin systems with time delay.* Phys. Rev. E **96**, 012106 (2017).

[2] Loos, S. A. M. and Klapp, S. H. L. *Heat flow due to time-delayed feedback.* Sci. Rep. **9**, 2491 (2019).

[3] Loos, S. A. M. and Klapp, S. H. L. *Fokker-Planck equations for time-delayed systems via Markovian Embedding.* J. Stat. Phys. **177**, 95–118 (2019).

[4] Loos, S. A. M., Hermann, S. M. and Klapp, S. H. L. *Non-reciprocal hidden degrees of freedom: A unifying perspective on memory, feedback, and activity.* ArXiv:1910.08372 (2019).

[5] Milton, J. G. *et al. Delayed pursuit-escape as a model for virtual stick balancing.* Nonlinear Theory and Its Applications, IEICE **4**, 129–137 (2013).

[6] Milton, J. G. *et al. Balancing with Vibration: A Prelude for "Drift and Act" Balance Control.* PLoS ONE **4**, e7427 (2009).

[7] Ohira, T. *Delay, noise, and resonance: human balancing and temporal non-locality.* Artif, Life Robot. **15**, 436–438 (2010).

[8] Bechhoefer, J. *Feedback for physicists: A tutorial essay on control.* Rev. Mod. Phys. **77**, 783 (2005).

[9] Longtin, A. *Complex time-delay systems: theory and applications, edited by Fatihcan M. Atay* (Springer-Verlag Berlin Heidelberg, 2010).

[10] Schöll, E. and Schuster, H. G. (eds.) *Handbook of chaos control* (John Wiley & Sons, 2008).

[11] Schöll, E., Klapp, S. H. L. and Hövel, P. (eds.) *Control of self-organizing nonlinear systems* (Springer, 2016).

[12] Zwanzig, R. *Nonequilibrium statistical mechanics* (Oxford University Press, 2001).

[13] Goychuk, I. *Viscoelastic subdiffusion: From anomalous to normal.* Phys. Rev. E **80**, 046125 (2009).

[14] Rouse Jr, P. E. *A theory of the linear viscoelastic properties of dilute solutions of coiling polymers.* The Journal of Chemical Physics **21**, 1272–1280 (1953).

[15] Maes, C., Safaverdi, S., Visco, P. and Van Wijland, F. *Fluctuation-response relations for nonequilibrium diffusions with memory.* Phys. Rev. E **87**, 022125 (2013).

[16] Franosch, T. *et al. Resonances arising from hydrodynamic memory in Brownian motion.* Nature **478**, 85 (2011).

[17] Kheifets, S., Simha, A., Melin, K., Li, T. and Raizen, M. G. *Observation of Brownian motion in liquids at short times: instantaneous velocity and memory loss.* Science **343**, 1493–1496 (2014).

[18] Marchetti, M. C. *et al. Hydrodynamics of soft active matter.* Rev. Mod. Phys. **85**, 1143 (2013).

[19] Narinder, N., Bechinger, C. and Gomez-Solano, J. R. *Memory-induced transition from a persistent random walk to circular motion for achiral microswimmers.* Phys. Rev. Lett. **121**, 078003 (2018).

[20] Nagai, K. H., Sumino, Y., Montagne, R., Aranson, I. S. and Chaté, H. *Collective motion of self-propelled particles with memory.* Phys. Rev. Lett. **114**, 168001 (2015).

[21] Scholz, C., Jahanshahi, S., Ldov, A. and Löwen, H. *Inertial delay of self-propelled particles.* Nat. Commun. **9**, 5156 (2018).

[22] Driver, R. D. *A" backwards" two-body problem of classical relativistic electrodynamics.* Phys. Rev. **178**, 2051 (1969).

[23] Will, C. M. *Propagation speed of gravity and the relativistic time delay.* Astrophys. J. **590**, 683 (2003).

[24] Goel, N. S., Maitra, S. C. and Montroll, E. W. *On the Volterra and other nonlinear models of interacting populations.* Rev. Mod. Phys. **43**, 231 (1971).

[25] Das, K., Srinivas, M., Srinivas, M. and Gazi, N. *Chaotic dynamics of a three species prey–predator competition model with bionomic harvesting due to delayed environmental noise as external driving force.* C. R. Biol. **335**, 503–513 (2012).

[26] Schiering, C. *et al. Feedback control of AHR signalling regulates intestinal immunity.* Nature **542**, 242 (2017).

[27] Pardee, A. B. *Regulation of Cell Metabolism* (J. and A. Churchill LTD, London, 1959).

[28] Polettini, M. and Esposito, M. *Irreversible thermodynamics of open chemical networks. I. Emergent cycles and broken conservation laws.* The Journal of chemical physics **141**, 07B610_1 (2014).

[29] Waibel, A., Hanazawa, T., Hinton, G., Shikano, K. and Lang, K. J. *Backpropagation: Theory, Architectures and Applications* (Lawrence Erlbaum Assoc. Mahwah, NJ, 1995).

[30] Longtin, A., Milton, J. G., Bos, J. E. and Mackey, M. C. *Noise and critical behavior of the pupil light reflex at oscillation onset.* Phys. Rev. A **41**, 6992 (1990).

[31] Cabral, J. R. *et al. Exploring mechanisms of spontaneous functional connectivity in MEG: How delayed network interactions lead to structured amplitude envelopes of band-pass filtered oscillations.* Neuroimage **90**, 423–435 (2014).

[32] Karmeshu, Gupta, V. and Kadambari, K. *Neuronal model with distributed delay: analysis and simulation study for gamma distribution memory kernel.* Biol. Cybern. .

[33] Voss, H. U. and Kurths, J. *Modelling and forecasting financial data: techniques of nonlinear dynamics*, vol. 2 (Springer Science & Business Media, 2002).

[34] Stoica, G. *A stochastic delay financial model.* Proc. Am. Math. Soc. **133**, 1837–1841 (2005).

[35] Tambue, A., Brown, E. K. and Mohammed, S. *A stochastic delay model for pricing debt and equity: Numerical techniques and applications.* Commun. Nonlinear Sci. Numer. Simul. **20**, 281–297 (2015).

[36] Callen, J. L., Khan, M. and Lu, H. *Accounting quality, stock price delay, and future stock returns.* Contemp. Account. Res. **30**, 269–295 (2013).

[37] Német, N. and Parkins, S. *Enhanced optical squeezing from a degenerate parametric amplifier via time-delayed coherent feedback.* Phys. Rev. A **94**, 023809 (2016).

[38] Mijalkov, M., McDaniel, A., Wehr, J. and Volpe, G. *Engineering sensorial delay to control phototaxis and emergent collective behaviors.* Phys. Rev. X **6**, 011008 (2016).

[39] Bruot, N., Damet, L., Kotar, J., Cicuta, P. and Lagomarsino, M. C. *Noise and synchronization of a single active colloid.* Phys. Rev. Lett. **107**, 094101 (2011).

[40] Masoller, C. *Noise-Induced Resonance in Delayed Feedback Systems.* Phys. Rev. Lett. **88**, 034102 (2002).

[41] Kane, D. M. and Shore, K. A. (eds.) *Unlocking dynamical diversity: optical feedback effects on semiconductor lasers* (John Wiley & Sons, 2005).

[42] Liu, Y., Chen, H., Liu, J., Davis, P. and Aida, T. *Synchronization of optical-feedback-induced chaos in semiconductor lasers by optical injection.* Phys. Rev. A **63**, 031802 (2001).

[43] Hein, S. M., Schulze, F., Carmele, A. and Knorr, A. *Entanglement control in quantum networks by quantum-coherent time-delayed feedback.* Phys. Rev. A **91**, 052321 (2015).

[44] Lu, Y. *et al. Intensified antibunching via feedback-induced quantum interference.* Phys. Rev. A **95**, 063840 (2017).

[45] Carmele, A., Kabuss, J., Schulze, F., Reitzenstein, S. and Knorr, A. *Single photon delayed feedback: a way to stabilize intrinsic quantum cavity electrodynamics.* Phys. Rev. Lett. **110**, 013601 (2013).

[46] Schneider, I. *Delayed feedback control of three diffusively coupled Stuart–Landau oscillators: a case study in equivariant Hopf bifurcation.* Philos. Trans. Royal Soc. A **371**, 20120472 (2013).

[47] Zakharova, A. *et al. Controlling chimera patterns in networks: interplay of structure, noise, and delay* (Springer, 2016).

[48] Schneider, I. and Bosewitz, M. *Eliminating restrictions of time-delayed feedback control using equivariance.* Disc. Cont. Dyn. Syst. A **36**, 451–467 (2016).

[49] Yi, T.-M., Huang, Y., Simon, M. I. and Doyle, J. *Robust perfect adaptation in bacterial chemotaxis through integral feedback control.* PNAS **97**, 4649–4653 (2000).

[50] Aguilar-Hidalgo, D. *et al. Critical point in self-organized tissue growth.* Phys. Rev. Lett. **120**, 198102 (2018).

[51] Schöll, E., Hiller, G., Hövel, P. and Dahlem, M. A. *Time-delayed feedback in neurosystems.* Philos. Trans. Royal Soc. A **367**, 1079–1096 (2009).

[52] Parmar, K., Blyuss, K. B., Kyrychko, Y. N. and Hogan, S. J. *Time-delayed models of gene regulatory networks.* Comput. Math. Meth. M. **2015**, 347273 (2015).

[53] Bratsun, D., Volfson, D., Tsimring, L. S. and Hasty, J. *Delay-induced stochastic oscillations in gene regulation.* PNAS **102**, 14593–14598 (2005).

[54] Gupta, C., López, J. M., Ott, W., Josić, K. and Bennett, M. R. *Transcriptional Delay Stabilizes Bistable Gene Networks.* Phys. Rev. Lett. **111**, 058104 (2013).

[55] Lai, X., Wolkenhauer, O. and Vera, J. *Understanding microRNA-mediated gene regulatory networks through mathematical modelling.* Nucleic acids research **44**, 6019–6035 (2016).

[56] Josić, K., López, J. M., Ott, W., Shiau, L. and Bennett, M. R. *Stochastic delay accelerates signaling in gene networks.* PLoS computational biology **7**, e1002264 (2011).

[57] Rateitschak, K. and Wolkenhauer, O. *Intracellular delay limits cyclic changes in gene expression.* Math. Biosci. **205**, 163–179 (2007).

[58] Friedman, N., Cai, L. and Xie, X. S. *Linking stochastic dynamics to population distribution: an analytical framework of gene expression.* Phys. Rev. Lett. **97**, 168302 (2006).

[59] Ito, S. and Sagawa, T. *Maxwells demon in biochemical signal transduction with feedback loop.* Nat. Commun. **6**, 7498 (2015).

[60] Micali, G. and Endres, R. G. *Bacterial chemotaxis: information processing, thermodynamics, and behavior.* Curr. Opin. Microbiol. **30**, 8–15 (2016).

[61] Wördemann, M. *Structured Light Fields: Applications in Optical Trapping, Manipulation, and Organisation* (Springer Science & Business Media, 2012).

[62] Kotar, J., Leoni, M., Bassetti, B., Lagomarsino, M. C. and Cicuta, P. *Hydrodynamic synchronization of colloidal oscillators.* PNAS **107**, 7669–7673 (2010).

[63] Kim, K. H. and Qian, H. *Fluctuation theorems for a molecular refrigerator.* Phys. Rev. E **75**, 022102 (2007).

[64] Qian, B., Montiel, D., Bregulla, A., Cichos, F. and Yang, H. *Harnessing thermal fluctuations for purposeful activities: the manipulation of single micro-swimmers by adaptive photon nudging.* Chem. Sci. **4**, 1420–1429 (2013).

[65] Balijepalli, A., Gorman, J. J., Gupta, S. K. and LeBrun, T. W. *Significantly improved trapping lifetime of nanoparticles in an optical trap using feedback control.* Nano Lett. **12**, 2347–2351 (2012).

[66] Khadka, U., Holubec, V., Yang, H. and Cichos, F. *Active Particles Bound by Information Flows.* Nat. Commun. **9**, 3864 (2018).

[67] Braun, M. and Cichos, F. *Optically Controlled Thermophoretic Trapping of Single Nano-Objects.* ACS Nano **7**, 11200–11208 (2013).

[68] Braun, M., Bregulla, A. P., Günther, K., Mertig, M. and Cichos, F. *Single Molecules Trapped by Dynamic Inhomogeneous Temperature Fields.* Nano Lett. **15**, 5499–5505 (2015).

[69] Haeufle, D. F. B. *et al. External control strategies for self-propelled particles: Optimizing navigational efficiency in the presence of limited resources.* Phys. Rev. E **94**, 012617 (2016).

[70] Loos, S. A. M., Gernert, R. and Klapp, S. H. L. *Delay-induced transport in a rocking ratchet under feedback control.* Phys. Rev. E **89**, 052136 (2014).

[71] Gernert, R., Emary, C. and Klapp, S. H. L. *Waiting time distribution for continuous stochastic systems.* Phys. Rev. E **90**, 062115 (2014).

[72] Gernert, R. and Klapp, S. H. L. *Enhancement of mobility in an interacting colloidal system under feedback control.* Phys. Rev. E **92**, 022132 (2015).

[73] Gernert, R., Loos, S. A. M., Lichtner, K. and Klapp, S. H. L. *Feedback control of colloidal transport.* In Schöll, E., Klapp, S. H. L. and Hövel, P. (eds.) *Control of Self-Organizing Nonlinear Systems*, 375–392 (Springer, 2016).

[74] Lichtner, K., Pototsky, A. and Klapp, S. H. L. *Feedback-induced oscillations in one-dimensional colloidal transport.* Phys. Rev. E **86**, 051405 (2012).

[75] Lichtner, K. and Klapp, S. H. L. *Feedback-controlled transport in an interacting colloidal system.* EPL **92**, 40007 (2010).

[76] Hennig, D. *Current control in a tilted washboard potential via time-delayed feedback.* Phys. Rev. E **79**, 041114 (2009).

[77] Küchler, U. and Mensch, B. *Langevins stochastic differential equation extended by a time-delayed term.* Stoch. Stoch. Rep. **40**, 23–42 (1992).

[78] Ohira, T. and Hosaka, T. *Control in delayed stochastic systems.* Tech. Rep. (2005).

[79] Ohira, T. and Yamane, T. *Delayed stochastic systems.* Phys. Rev. E **61**, 1247–1257 (2000).

[80] Wu, D. and Zhu, S. *Brownian motor with time-delayed feedback.* Phys. Rev. E **73**, 051107 (2006).

[81] Milton, J., Cabrera, J. and Ohira, T. *Unstable dynamical systems: delays, noise and control.* EPL **83**, 48001 (2008).

[82] Guillouzic, S., L'Heureux, I. and Longtin, A. *Small delay approximation of stochastic delay differential equations.* Phys. Rev. E **59**, 3970 (1999).

[83] Guillouzic, S., L Heureux, I. and Longtin, A. *Rate processes in a delayed, stochastically driven, and overdamped system.* Phys. Rev. E **61**, 4906 (2000).

[84] Frank, T. D. and Beek, P. J. *Stationary solutions of linear stochastic delay differential equations: Applications to biological systems.* Phys. Rev. E **64**, 021917 (2001).

[85] Frank, T. D., Beek, P. J. and Friedrich, R. *Fokker-Planck perspective on stochastic delay systems: Exact solutions and data analysis of biological systems.* Phys. Rev. E **68**, 021912 (2003).

[86] Frank, T. D. *Delay Fokker-Planck equations, perturbation theory, and data analysis for nonlinear stochastic systems with time delays.* Phys. Rev. E **71**, 031106 (2005).

[87] Frank, T. D. *Delay Fokker-Planck equations, Novikov's theorem, and Boltzmann distributions as small delay approximations.* Phys. Rev. E **72**, 011112 (2005).

[88] Van Vu, T. and Hasegawa, Y. *Uncertainty relations for time-delayed Langevin systems.* Phys. Rev. E **100**, 012134 (2019).

[89] Geiss, D., Kroy, K. and Holubec, V. *Brownian molecules formed by delayed harmonic interactions.* New J. Phys. **21**, 093014 (2019).

[90] Mackey, M. C. and Nechaeva, I. G. *Noise and stability in differential delay equations.* J. Dyn. Diff. Equat. **6**, 395–426 (1994).

[91] McKetterick, T. J. and Giuggioli, L. *Exact dynamics of stochastic linear delayed systems: Application to spatiotemporal coordination of comoving agents.* Phys. Rev. E **90**, 042135 (2014).

[92] Giuggioli, L., McKetterick, T. J., Kenkre, V. M. and Chase, M. *Fokker–Planck description for a linear delayed Langevin equation with additive Gaussian noise.* J. Phys. A **49**, 384002 (2016).

[93] Munakata, T. and Rosinberg, M. L. *Entropy production and fluctuation theorems for Langevin processes under continuous non-Markovian feedback control.* Phys. Rev. Lett. **112**, 180601 (2014).

[94] Zwanzig, R. *Nonlinear generalized Langevin equations.* J. Stat. Phys. **9**, 215–220 (1973).

[95] Kubo, R. *The fluctuation-dissipation theorem.* Rep. Prog. Phys **29**, 255 (1966).

[96] Krüeger, M. and Maes, C. *The modified Langevin description for probes in a nonlinear medium.* J. Phys.: Condens. Matter **29**, 064004 (2016).

[97] Bao, J.-D., Hänggi, P. and Zhuo, Y.-Z. *Non-Markovian Brownian dynamics and nonergodicity.* Phys. Rev. E **72**, 061107 (2005).

[98] Maes, C. *On the second fluctuation–dissipation theorem for nonequilibrium baths.* J. Stat. Phys. **154**, 705–722 (2014).

[99] Munakata, T., Iwama, S. and Kimizuka, M. *Linear stochastic system with delay: Energy balance and entropy production.* Phys. Rev. E **79**, 031104 (2009).

[100] Bloom, G. S. and Endow, S. A. *Motor proteins 1: kinesins.* Protein profile **2**, 1105–1171 (1995).

[101] Kuo, S. C. and Sheetz, M. P. *Force of single kinesin molecules measured with optical tweezers.* Science **260**, 232–234 (1993).

[102] Ramaswamy, S. *The mechanics and statistics of active matter.* Annu. Rev. Condens. Matter Phys. **1**, 323–345 (2010).

[103] Ramaswamy, S. *Active matter.* J. Stat. Mech. Theor. Exp. **2017**, 054002 (2017).

[104] Pietzonka, P. and Seifert, U. *Entropy production of active particles and for particles in active baths.* J. Phys. A: Math. and Theor. **51**, 01LT01 (2017).

[105] Shankar, S. and Marchetti, M. C. *Hidden entropy production and work fluctuations in an ideal active gas.* Phys. Rev. E **98**, 020604(R) (2018).

[106] Speck, T. *Stochastic thermodynamics for active matter.* EPL **114**, 30006 (2016).

[107] Dabelow, L., Bo, S. and Eichhorn, R. *Irreversibility in active matter systems: Fluctuation theorem and mutual information.* Phys. Rev. X **9**, 021009 (2019).

[108] Ehrich, J. and Kahlen, M. *Approximating microswimmer dynamics by active Brownian motion: Energetics and efficiency.* Phys. Rev. E **99**, 012118 (2019).

[109] Caprini, L., Marconi, U. M. B., Puglisi, A. and Vulpiani, A. *The entropy production of Ornstein–Uhlenbeck active particles: a path integral method for correlations.* J. Stat. Mech. Theor. Exp. **2019**, 053203 (2019).

[110] Durve, M., Saha, A. and Sayeed, A. *Active particle condensation by non-reciprocal and time-delayed interactions.* Eur. Phys. J. E **41**, 49 (2018).

[111] Kürsten, R., Sushkov, V. and Ihle, T. *Giant Kovacs-like memory effect for active particles.* Phys. Rev. Lett. **119**, 188001 (2017).

[112] Khadem, S. M. J. and Klapp, S. H. *Delayed feedback control of active particles: a controlled journey towards the destination.* Phys. Chem. Chem. Phys. **21**, 13776 (2019).

[113] Tarama, S., Egelhaaf, S. U. and Löwen, H. *Traveling band formation in feedback-driven colloids.* Phys. Rev. E **100**, 022609 (2019).

[114] Piwowarczyk, R., Selin, M., Ihle, T. and Volpe, G. *Influence of sensorial delay on clustering and swarming.* Phys. Rev. E **100**, 012607 (2019).

[115] Schwabl, F. *Statistical mechanics* (Springer, Berlin, 2002).

[116] Rosinberg, M. L., Munakata, T. and Tarjus, G. *Stochastic thermodynamics of Langevin systems under time-delayed feedback control: Second-law-like inequalities.* Phys. Rev. E **91**, 042114 (2015).

[117] Puglisi, A. and Villamaina, D. *Irreversible effects of memory.* EPL **88**, 30004 (2009).

[118] Crisanti, A., Puglisi, A. and Villamaina, D. *Nonequilibrium and information: The role of cross correlations.* Phys. Rev. E **85**, 061127 (2012).

[119] Siegle, P., Goychuk, I., Talkner, P. and Hänggi, P. *Markovian embedding of non-Markovian superdiffusion.* Phys. Rev. E **81**, 011136 (2010).

[120] Siegle, P., Goychuk, I. and Hänggi, P. *Markovian embedding of fractional superdiffusion.* EPL **93**, 20002 (2011).

[121] Villamaina, D., Baldassarri, A., Puglisi, A. and Vulpiani, A. *The fluctuation-dissipation relation: how does one compare correlation functions and responses?* J. Stat. Mech. Theory Exp **2009**, P07024 (2009).

[122] Smith, H. L. *An introduction to delay differential equations with applications to the life sciences*, vol. 57 (Springer New York, 2011).

[123] Niculescu, S.-I. and Gu, K. *Advances in time-delay systems*, vol. 38 (Springer Science & Business Media, 2012).

[124] Sekimoto, K. *Stochastic energetics*, vol. 799 (Springer, 2010).

[125] Seifert, U. *Stochastic thermodynamics, fluctuation theorems and molecular machines.* Rep. Prog. Phys. **75**, 126001 (2012).

[126] Esposito, M. *Stochastic thermodynamics under coarse graining.* Phys. Rev. E **85**, 041125 (2012).

[127] Sagawa, T. and Ueda, M. *Nonequilibrium thermodynamics of feedback control.* Phys. Rev. E **85**, 021104 (2012).

[128] Abreu, D. and Seifert, U. *Thermodynamics of genuine nonequilibrium states under feedback control.* Phys. Rev. Lett. **108**, 030601 (2012).

[129] Barato, A. C. and Seifert, U. *Unifying Three Perspectives on Information Processing in Stochastic Thermodynamics.* Phys. Rev. Lett. **112**, 090601 (2014).

[130] Mandal, D. and Jarzynski, C. *Work and information processing in a solvable model of Maxwell's demon.* PNAS **109**, 11641–11645 (2012).

[131] Parrondo, J. M. R., Horowitz, J. M. and Sagawa, T. *Thermodynamics of information.* Nat. Phys. **11**, 131–139 (2015).

[132] Kutvonen, A., Sagawa, T. and Ala-Nissila, T. *Thermodynamics of information exchange between two coupled quantum dots.* Phys. Rev. E **93**, 032147 (2016).

[133] Barato, A. C. and Seifert, U. *Thermodynamic uncertainty relation for biomolecular processes.* Phys. Rev. Lett. **114**, 158101 (2015).

[134] Esposito, M., Lindenberg, K. and Van den Broeck, C. *Thermoelectric efficiency at maximum power in a quantum dot.* EPL **85**, 60010 (2009).

[135] Strasberg, P., Schaller, G., Brandes, T. and Esposito, M. *Thermodynamics of a physical model implementing a Maxwell demon.* Phys. Rev. Lett. **110**, 040601 (2013).

[136] Koski, J. V., Maisi, V. F., Sagawa, T. and Pekola, J. P. *Experimental Observation of the Role of Mutual Information in the Nonequilibrium Dynamics of a Maxwell Demon.* Phys. Rev. Lett. **113**, 030601 (2014).

[137] Imparato, A., Peliti, L., Pesce, G., Rusciano, G. and Sasso, A. *Work and heat probability distribution of an optically driven Brownian particle: Theory and experiments.* Phys. Rev. E **76**, 050101 (2007).

[138] Speck, T. and Seifert, U. *The Jarzynski relation, fluctuation theorems, and stochastic thermodynamics for non-Markovian processes.* J. Stat. Mech. Theory Exp **2007**, L09002 (2007).

[139] Vaikuntanathan, S. and Jarzynski, C. *Dissipation and lag in irreversible processes.* EPL **87**, 60005 (2009).

[140] Kutvonen, A., Ala-Nissila, T. and Pekola, J. *Entropy production in a non-Markovian environment.* Phys. Rev. E **92**, 012107 (2015).

[141] Hasegawa, H. *Classical open systems with nonlinear nonlocal dissipation and state-dependent diffusion: Dynamical responses and the Jarzynski equality.* Phys. Rev. E **84**, 051124 (2011).

[142] García-García, R. *Nonadiabatic entropy production for non-Markov dynamics.* Phys. Rev. E **86**, 031117 (2012).

[143] Roche, B. *et al. Harvesting dissipated energy with a mesoscopic ratchet.* Nat. Commun. **6**, 6738 (2015).

[144] Whitney, R. S. *Non-Markovian quantum thermodynamics: Laws and fluctuation theorems.* Phys. Rev. B **98**, 085415 (2018).

[145] Mai, T. and Dhar, A. *Nonequilibrium work fluctuations for oscillators in non-Markovian baths.* Phys. Rev. E **75**, 061101 (2007).

[146] Schmidt, R., Carusela, M. F., Pekola, J. P., Suomela, S. and Ankerhold, J. *Work and heat for two-level systems in dissipative environments: Strong driving and non-Markovian dynamics.* Phys. Rev. B **91**, 224303 (2015).

[147] Cui, B. and Zaccone, A. *Generalized Langevin equation and fluctuation-dissipation theorem for particle-bath systems in external oscillating fields.* Phys. Rev. E **97**, 060102 (2018).

[148] Sagawa, T. *Thermodynamics of information processing in small systems.* Progress of theoretical physics **127**, 1–56 (2012).

[149] Maxwell, J. C. *Theory of heat* (Dover Publications, inc., 1871).

[150] Roldán, É., Martínez, I. A., Parrondo, J. M. R. and Petrov, D. *Universal features in the energetics of symmetry breaking.* Nat. Phys. **10**, 457 (2014).

[151] Cao, F. J. and Feito, M. *Thermodynamics of feedback controlled systems.* Phys. Rev. E **79**, 041118 (2009).

[152] Rosinberg, M. L., Tarjus, G. and Munakata, T. *Stochastic thermodynamics of Langevin systems under time-delayed feedback control. II. Nonequilibrium steady-state fluctuations.* Phys. Rev. E **95**, 022123 (2017).

[153] Landauer, R. *Irreversibility and heat generation in the computing process.* IBM journal of research and development **5**, 183–191 (1961).

[154] Polettini, M. and Esposito, M. *Effective thermodynamics for a marginal observer.* Phys. Rev. Lett. **119**, 240601 (2017).

[155] Agudo-Canalejo, J. and Golestanian, R. *Active Phase Separation in Mixtures of Chemically Interacting Particles.* Phys. Rev. Lett. **123**, 018101 (2019).

[156] van Kampen, N. G. *Remarks on non-Markov processes.* Braz. J. Phys. **28**, 90–96 (1998).

[157] Maxwell, J. C. *VIII. A dynamical theory of the electromagnetic field.* Philos. Trans. Royal Soc. 459–512 (1865).

[158] Gardiner, C. W. *Handbook of Stochastic Methods* (Springer, 2002), 2 edn.

[159] Hänggi, P., Marchesoni, F. and Nori, F. *Brownian motors.* Annalen der Physik **14**, 51–70 (2005).

[160] Kapral, R. *Multiparticle collision dynamics: Simulation of complex systems on mesoscales.* Advances in Chemical Physics **140**, 89 (2008).

[161] Brown, R. *A brief account of microscopical observations.* The Philosophical Magazine and Annals of Philosophy **4**, 161 (1827).

[162] Dhont, J. K. *An introduction to dynamics of colloids*, vol. 2 (Elsevier, 1996).

[163] Löwen, H. *Colloidal dispersions in external fields: recent developments.* J. Phys. Condens. Matter **20**, 404201 (2008).

[164] Yethiraj, A. and van Blaaderen, A. *A colloidal model system with an interaction tunable from hard sphere to soft and dipolar.* Nature **421**, 513–517 (2003).

[165] Berg, J. *Out-of-equilibrium dynamics of gene expression and the Jarzynski equality.* Phys. Rev. Lett. **100**, 188101 (2008).

[166] Ovaskainen, O. and Meerson, B. *Stochastic models of population extinction.* Trends in ecology & evolution **25**, 643–652 (2010).

[167] Majda, A. J., Timofeyev, I. and Eijnden, E. V. *Models for stochastic climate prediction.* PNAS **96**, 14687–14691 (1999).

[168] Bouchaud, J.-P. and Cont, R. *A Langevin approach to stock market fluctuations and crashes.* The European Physical Journal B-Condensed Matter and Complex Systems **6**, 543–550 (1998).

[169] Verechtchaguina, T., Schimansky-Geier, L. and Sokolov, I. *Spectra and waiting-time densities in firing resonant and nonresonant neurons.* Phys. Rev. E **70**, 031916 (2004).

[170] Van Zon, R., Ciliberto, S. and Cohen, E. *Power and heat fluctuation theorems for electric circuits.* Phys. Rev. Lett. **92**, 130601 (2004).

[171] Kish, L. B. and Granqvist, C.-G. *Electrical Maxwell demon and Szilard engine utilizing Johnson noise, measurement, logic and control.* PloS one **7**, e46800 (2012).

[172] Gillespie, D. T. *The chemical Langevin equation.* The Journal of Chemical Physics **113**, 297–306 (2000).

[173] Risken, H. *The Fokker-Planck equation* (Springer, 1984).

[174] Karatzsas, I. and Shreve, S. E. *Brownian motion and stochastic calculus.* Graduate texts in Mathematics **113** (1991).

[175] Hida, T., Kuo, H.-H., Potthoff, J. and Streit, L. *White noise: an infinite dimensional calculus*, vol. 253 (Springer Science & Business Media, 2013).

[176] Bewersdorff, J. *Statistik - wie und warum sie funktioniert: Ein mathematisches Lesebuch* (Springer-Verlag, 2011).

[177] Kudryashov, N. A. *Nonlinear differential equations with exact solutions.* ArXiv:0311058 (2003).

[178] Einstein, A. *Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen.* Annalen der Physik **322**, 549–560 (1905).

[179] Jeon, J.-H., Leijnse, N., Oddershede, L. B. and Metzler, R. *Anomalous diffusion and power-law relaxation of the time averaged mean squared displacement in worm-like micellar solutions.* New J. Phys. **15**, 045011 (2013).

[180] Metzler, R., Jeon, J.-H., Cherstvy, A. G. and Barkai, E. *Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking.* Phys. Chem. Chem. Phys. **16**, 24128–24164 (2014).

[181] Kloeden, P. and Platen, E. *Numerical solution of stochastic differential equations* (Springer–Verlag, 1992).

[182] Chow, C. C. and Buice, M. A. *Path integral methods for stochastic differential equations.* JMN **5**, 8 (2015).

[183] Mannella, R. and McClintock, P. V. *Itô versus Stratonovich: 30 years later.* Fluctuation and Noise Letters **11**, 1240010 (2012).

[184] Vale, R. D. *The molecular motor toolbox for intracellular transport.* Cell **112**, 467–480 (2003).

[185] Sokolov, I. *Itô, Stratonovich, Hänggi and all the rest: The thermodynamics of interpretation.* Chemical Physics **375**, 359–363 (2010).

[186] Moon, W. and Wettlaufer, J. *On the interpretation of Stratonovich calculus.* New J. Phys. **16**, 055017 (2014).

[187] Sancho, J. M. *Brownian colloidal particles: Itô, Stratonovich, or a different stochastic interpretation.* Phys. Rev. E **84**, 062102 (2011).

[188] Farago, O. and Grønbech-Jensen, N. *Langevin dynamics in inhomogeneous media: Re-examining the Itô-Stratonovich dilemma.* Phys. Rev. E **89**, 013301 (2014).

[189] Yuan, R. and Ao, P. *Beyond Itô versus Stratonovich.* J. Stat. Mech.: Theory Exp. **2012**, P07010 (2012).

[190] Hänggi, P. *Path integral solutions for non-Markovian processes.* Z. Phys. B **75**, 275–281 (1989).

[191] Abe, S. *Generalized molecular chaos hypothesis and the H theorem: problem of constraints and amendment of nonextensive statistical mechanics.* Phys. Rev. E **79**, 041116 (2009).

[192] Ehrenfest, P. and Ehrenfest, T. *The conceptual foundations of the statistical approach in mechanics* (Courier Corporation, 1990).

[193] Mori, H. *Transport, collective motion, and Brownian motion.* Progress of theoretical physics **33**, 423–455 (1965).

[194] Lee, H. K., Lahiri, S. and Park, H. *Nonequilibrium steady states in Langevin thermal systems.* Phys. Rev. E **96**, 022134 (2017).

[195] Hemmer, P. C., Maximon, L. and Wergeland, H. *Recurrence time of a dynamical system.* Phys. Rev. **111**, 689 (1958).

[196] Strogatz, S. H., Friedman, M., Mallinckrodt, A. J., McKay, S. *et al. Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering.* Computers in Physics **8**, 532–532 (1994).

[197] Banasiak, J. *Chaotic linear systems in mathematical biology.* South African Journal of Science **104**, 173–179 (2008).

[198] Klages, R. and Dorfman, J. *Simple maps with fractal diffusion coefficients.* Phys. Rev. Lett. **74**, 387 (1995).

[199] Moran, B., Hoover, W. G. and Bestiale, S. *Diffusion in a periodic Lorentz gas.* J. Stat. Phys. **48**, 709–726 (1987).

[200] Machta, J. and Zwanzig, R. *Diffusion in a periodic Lorentz gas.* Phys. Rev. Lett. **50**, 1959 (1983).

[201] Cecconi, F., del Castillo-Negrete, D., Falcioni, M. and Vulpiani, A. *The origin of diffusion: the case of non-chaotic systems.* Physica D **180**, 129–139 (2003).

[202] Berner, J., Müller, B., Gomez-Solano, J. R., Krüger, M. and Bechinger, C. *Oscillating modes of driven colloids in overdamped systems.* Nat. comm. **9**, 1–8 (2018).

[203] Reinken, H., Klapp, S. H., Bär, M. and Heidenreich, S. *Derivation of a hydrodynamic theory for mesoscale dynamics in microswimmer suspensions.* Phys. Rev. E **97**, 022613 (2018).

[204] Seifert, U. and Speck, T. *Fluctuation-dissipation theorem in nonequilibrium steady states.* EPL **89**, 10007 (2010).

[205] Bohec, P. *et al. Probing active forces via a fluctuation–dissipation relation.* Arxiv:1203.3571 (2012).

[206] Price, H. *Time symmetry in microphysics.* Philosophy of science **64**, S235–S244 (1997).

[207] Wu, T.-Y. *Boltzmann's H theorem and the Loschmidt and the Zermelo paradoxes.* International Journal of Theoretical Physics **14**, 289–294 (1975).

[208] Posch, H. A., Dellago, C., Hoover, W. G. and Kum, O. *Microscopic Time-Reversibility and Macroscopic IrreversibilityStill a Paradox?* In *Pioneering Ideas for the Physical and Chemical Sciences*, 233–248 (Springer, 1997).

[209] Felderhof, B. *On the derivation of the fluctuation-dissipation theorem.* Journal of Physics A: Mathematical and General **11**, 921 (1978).

[210] Noether, E. *Invariante Variationsprobleme.* Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen. 235 – 257 (1918).

[211] Evans, D. J. and Searles, D. J. *The fluctuation theorem.* Advances in Physics **51**, 1529–1585 (2002).

[212] Caticha, A. *Lectures on probability, entropy, and statistical physics.* ArXiv:0808.0012 (2008).

[213] Petrillo, A., Salvi, A., Santini, S. and Valente, A. S. *Adaptive multi-agents synchronization for collaborative driving of autonomous vehicles with multiple communication delays.* Transportation research part C: emerging technologies **86**, 372–392 (2018).

[214] (2019). Weblink: `https://www.nobelprize.org/ceremonies/the-nobel-prize-award-ceremony-2018/`.

[215] Debiossac, M., Grass, D., Alonso, J. J., Lutz, E. and Kiesel, N. *Thermodynamics of continuous non-Markovian feedback control.* ArXiv:1904.04889 (2019).

[216] Wallin, A. E., Ojala, H., Hæggström, E. and Tuma, R. *Stiffer optical tweezers through real-time feedback control.* Appl. Phys. Lett. **92**, 224104 (2008).

[217] Braun, M., Würger, A. and Cichos, F. *Trapping of single nano-objects in dynamic temperature fields.* Phys. Chem. Chem. Phys. **16**, 15207 (2014).

[218] Ashkin, A. *Acceleration and trapping of particles by radiation pressure.* Phys. Rev. Lett. **24**, 156 (1970).

[219] (2019). Weblink: `https://www.nobelprize.org/prizes/physics/2018/ashkin/facts/`.

[220] Blickle, V. and Bechinger, C. *Realization of a micrometre-sized stochastic heat engine.* Nature Physics **8**, 143 (2012).

[221] Killian, J. L., Ye, F. and Wang, M. D. *Optical Tweezers: A Force to Be Reckoned With.* Cell **175**, 1445–1448 (2018).

[222] Schirber, M. *Focus: Nobel PrizeLasers as Tools.* Physics **11**, 100 (2018).

[223] Kaufman, A. M., Lester, B. J. and Regal, C. A. *Cooling a single atom in an optical tweezer to its quantum ground state.* Phys. Rev. X **2**, 041014 (2012).

[224] Zhong, M.-C., Wei, X.-B., Zhou, J.-H., Wang, Z.-Q. and Li, Y.-M. *Trapping red blood cells in living animals using optical tweezers.* Nat. Commun. **4**, 1768 (2013).

[225] Ashkin, A. and Dziedzic, J. M. *Optical trapping and manipulation of viruses and bacteria.* Science **235**, 1517–1520 (1987).

[226] Ashkin, A. *The study of cells by optical trapping and manipulation of living cells using infrared laser beams.* Gravitational and Space Research **4** (2007).

[227] Peterman, E. J., Gittes, F. and Schmidt, C. F. *Laser-induced heating in optical traps.* Biophysical journal **84**, 1308–1316 (2003).

[228] Liphardt, J., Dumont, S., Smith, S. B., Tinoco, I. and Bustamante, C. *Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality.* Science **296**, 1832–1835 (2002).

[229] Martínez, I. A. *et al. Brownian carnot engine.* Nature physics **12**, 67 (2016).

[230] Juniper, M. P. N., Straube, A. V., Aarts, D. G. A. L. and Dullens, R. P. A. *Colloidal particles driven across periodic optical-potential-energy landscapes.* Phys. Rev. E **93**, 012608 (2016).

[231] Kim, K. H. and Qian, H. *Entropy production of Brownian macromolecules with inertia.* Phys. Rev. Lett. **93**, 120602 (2004).

[232] Levy, N., Pressman, A., Mussa-Ivaldi, F. A. and Karniel, A. *Adaptation to delayed force perturbations in reaching movements.* PLoS one **5** (2010).

[233] Liu, M., Wang, K. and Hong, Q. *Stability of a stochastic logistic model with distributed delay.* Mathematical and Computer Modelling **57**, 1112–1121 (2013).

[234] René, A. and Longtin, A. *Mean, covariance, and effective dimension of stochastic distributed delay dynamics.* Chaos **27**, 114322 (2017).

[235] Kyrychko, Y., Blyuss, K. and Schöll, E. *Amplitude and phase dynamics in oscillators with distributed-delay coupling.* Philos. Trans. Royal Soc. A **371**, 20120466 (2013).

[236] Bernard, S., Bélair, J. and Mackey, M. C. *Sufficient conditions for stability of linear differential equations with distributed delay.* Discrete and Continuous Dynamical Systems Series B **1**, 233–256 (2001).

[237] Hänggi, P. and Talkner, P. *On the equivalence of time-convolutionless master equations and generalized Langevin equations.* Phys. Lett. A **68**, 9–11 (1978).

[238] Hänggi, P. and Thomas, H. *Stochastic processes: Time evolution, symmetries and linear response.* Phys. Rep. **88**, 207–319 (1982).

[239] Hernández-Machado, A., Sancho, J., San Miguel, M. and Pesquera, L. *Joint probability distribution of nonMarkovian sde.* EPJ B **52**, 335–343 (1983).

[240] Budini, A. A. and Cáceres, M. O. *Functional characterization of generalized Langevin equations.* J. Phys. A **37**, 5959 (2004).

[241] Diekmann, O., Van Gils, S. A., Lunel, S. M. and Walther, H.-O. *Delay equations: functional-, complex-, and nonlinear analysis*, vol. 110 (Springer Science & Business Media, 2012).

[242] Driver, R. D. *Ordinary and delay differential equations*, vol. 20 (Springer Science & Business Media, 2012).

[243] Hale, J. K. and Lunel, S. M. V. *Introduction to functional differential equations*, vol. 99 (Springer Science & Business Media, 2013).

[244] Hale, J. K. and Lunel, S. M. V. *Effects of small delays on stability and control.* In *Operator theory and analysis*, 275–301 (Springer, 2001).

[245] Erneux, T. *Applied delay differential equations*, vol. 3 (Springer Science & Business Media, 2009).

[246] Pyragas, K. *Control of chaos via extended delay feedback.* Phys. Lett. A **206**, 323–330 (1995).

[247] Pyragas, K. *Continuous control of chaos by self-controlling feedback.* Phys. Lett. A **170**, 421–428 (1992).

[248] Parmananda, P. *et al. Stabilization of unstable steady states and periodic orbits in an electrochemical system using delayed-feedback control.* Phys. Rev. E **59**, 5266 (1999).

[249] Buckwar, E. *Introduction to the numerical analysis of stochastic delay differential equations.* J. Comput. Appl. Math. **125**, 297–307 (2000).

[250] Luo, J. *Fixed points and stability of neutral stochastic delay differential equations.* Journal of Mathematical Analysis and Applications **334**, 431–440 (2007).

[251] Wu, F., Mao, X. and Szpruch, L. *Almost sure exponential stability of numerical solutions for stochastic delay differential equations.* Numer. Math. **115**, 681–697 (2010).

[252] Klosek, M. M. and Kuske, R. *Multiscale analysis of stochastic delay differential equations.* Multiscale Modeling & Simulation **3**, 706–729 (2005).

[253] Baker, C. T., Ford, J. M. and Ford, N. J. *Bifurcations in approximate solutions of stochastic delay differential equations.* International Journal of Bifurcation and Chaos **14**, 2999–3021 (2004).

[254] Kramers, H. A. *Brownian motion in a field of force and the diffusion model of chemical reactions.* Physica **7**, 284–304 (1940).

[255] Hänggi, P., Talkner, P. and Borkovec, M. *Reaction-rate theory: fifty years after Kramers.* Rev. Mod. Phys. **62**, 251 (1990).

[256] Tsimring, L. S. and Pikovsky, A. *Noise-Induced Dynamics in Bistable Systems with Delay.* Phys. Rev. Lett. **87**, 250602 (2001).

[257] Masoller, C. *Distribution of Residence Times of Time-Delayed Bistable Systems Driven by Noise.* Phys. Rev. Lett. **90**, 020601 (2003).

[258] Du, L. and Mei, D. *Global delay induced transition in a bistable system with multiplicative and additive noises.* Indian J. Phys. **89**, 267–272 (2015).

[259] Piwonski, T., Houlihan, J., Busch, T. and Huyet, G. *Delay-Induced Excitability.* Phys. Rev. Lett. **95**, 040601 (2005).

[260] Goulding, D. *et al. Kramers' law for a bistable system with time-delayed noise.* Phys. Rev. E **76**, 031128 (2007).

[261] Reimann, P. *et al. Giant Acceleration of Free Diffusion by Use of Tilted Periodic Potentials.* Phys. Rev. Lett. **87**, 010602 (2001).

[262] Emary, C., Gernert, R. and Klapp, S. H. L. *Minimal model for short-time diffusion in periodic potentials.* Phys. Rev. E **86**, 061135 (2012).

[263] Arteaga, M. A. *et al. Experimental evidence of coherence resonance in a time-delayed bistable system.* Phys. Rev. Lett. **99**, 023903 (2007).

[264] Kappler, J., Daldrop, J. O., Brünig, F. N., Boehle, M. D. and Netz, R. R. *Memory-induced acceleration and slowdown of barrier crossing.* J. Chem. Phys. **148**, 014903 (2018).

[265] Balanov, A. G., Janson, N. B. and Schöll, E. *Control of noise-induced oscillations by delayed feedback.* Physica D: Nonlinear Phenomena **199**, 1–12 (2004).

[266] Xiao, T. *Heat dissipation and information flow for delayed bistable Langevin systems near coherence resonance.* Phys. Rev. E **94**, 052109 (2016).

[267] Reimann, P. *Brownian motors noisy transport far from equilibrium.* Phys. Rep. **361** (2002).

[268] Marconi, U. M. B. and Tarazona, P. *Dynamic density functional theory of fluids.* J. Chem. Phys. **110**, 8032–8044 (1999).

[269] Menzel, A. M., Saha, A., Hoell, C. and Löwen, H. *Dynamical density functional theory for microswimmers.* J. Chem. Phys. **144**, 024115 (2016).

[270] Das, S. P. *Mode-coupling theory and the glass transition in supercooled liquids.* Rev. Mod. Phys. **76**, 785 (2004).

[271] Janssen, L. M. C. and Reichman, D. R. *Microscopic dynamics of supercooled liquids from first principles.* Phys. Rev. Lett. **115**, 205701 (2015).

[272] Hunter, M. C. W. and Davis, E. B. *The Works of Robert Boyle*, vol. 7 (Pickering & Chatto London, 1999).

[273] Carroll, S. M. and Chen, J. *Does inflation provide natural initial conditions for the universe?* International Journal of Modern Physics D **14**, 2335–2339 (2005).

[274] Katz, A. *Principles of statistical mechanics: the information theory approach* (WH Freeman, 1967).

[275] Jaynes, E. T. *Information theory and statistical mechanics.* Phys. Rev. **106**, 620 (1957).

[276] Wolpert, D. H. *Memory systems, computation, and the second law of thermodynamics.* Int. J. Theor. Phys **31**, 743–785 (1992).

[277] Morikuni, Y. and Tasaki, H. *Quantum Jarzynski-Sagawa-Ueda relations.* J. Stat. Phys. **143**, 1–10 (2011).

[278] Neumann, C. *Vorlesungen über die mechanische Theorie der Wärme* (Teubner, 1875).

[279] Aaronson, S., Carroll, S. M. and Ouellette, L. *Quantifying the rise and fall of complexity in closed systems: The coffee automaton.* ArXiv:1405.6903 (2014).

[280] Lebowitz, J. L. and Spohn, H. *A Gallavotti–Cohen-type symmetry in the large deviation functional for stochastic dynamics.* J. Stat. Phys. **95**, 333–365 (1999).

[281] García-García, R., Domínguez, D., Lecomte, V. and Kolton, A. B. *Unifying approach for fluctuation theorems from joint probability distributions.* Phys. Rev. E **82**, 030104 (2010).

[282] Esposito, M. and Van den Broeck, C. *Three detailed fluctuation theorems.* Phys. Rev. Lett. **104**, 090601 (2010).

[283] Tasaki, H. *Jarzynski relations for quantum systems and some applications.* arXiv preprint cond-mat/0009244 (2000).

[284] Wolf, S. and Stock, G. *Targeted molecular dynamics calculations of free energy profiles using a nonequilibrium friction correction.* J. Chem. Theory Comput. **14**, 6175–6182 (2018).

[285] Williams, S. R., Evans, D. J. and Searles, D. J. *Nonequilibrium Umbrella Sampling and the Functional Crooks Fluctuation Theorem.* J. Stat. Phys. **145**, 831–840 (2011).

[286] Sivak, D. A., Chodera, J. D. and Crooks, G. E. *Using nonequilibrium fluctuation theorems to understand and correct errors in equilibrium and nonequilibrium simulations of discrete Langevin dynamics.* Phys. Rev. X **3**, 011007 (2013).

[287] Jarzynski, C. *Nonequilibrium equality for free energy differences.* Phys. Rev. Lett. **78**, 2690 (1997).

[288] Crooks, G. E. *Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences.* Phys. Rev. E **60**, 2721 (1999).

[289] Horowitz, J. M. and Sandberg, H. *Second-law-like inequalities with information and their interpretations.* New J. Phys. **16**, 125007 (2014).

[290] Sagawa, T. and Ueda, M. *Generalized Jarzynski Equality under Nonequilibrium Feedback Control.* Phys. Rev. Lett. **104**, 090602 (2010).

[291] Esposito, M. and Schaller, G. *Stochastic thermodynamics for "Maxwell demon" feedbacks.* EPL **99**, 30003 (2012).

[292] Horowitz, J. M. and Esposito, M. *Thermodynamics with continuous information flow.* Phys. Rev. X **4**, 031015 (2014).

[293] Rosinberg, M., Tarjus, G. and Munakata, T. *Influence of time delay on information exchanges between coupled linear stochastic systems.* Phys. Rev. E **98**, 032130 (2018).

[294] Mehl, J., Lander, B., Bechinger, C., Blickle, V. and Seifert, U. *Role of hidden slow degrees of freedom in the fluctuation theorem.* Phys. Rev. Lett. **108**, 220601 (2012).

[295] Kawaguchi, K. and Nakayama, Y. *Fluctuation theorem for hidden entropy production.* Phys. Rev. E **88**, 022147 (2013).

[296] Altaner, B. and Vollmer, J. *Fluctuation-preserving coarse graining for biochemical systems.* Phys. Rev. Lett. **108**, 228101 (2012).

[297] Puglisi, A., Pigolotti, S., Rondoni, L. and Vulpiani, A. *Entropy production and coarse graining in Markov processes.* J. Stat. Mech.: Theory Exp. **2010**, P05015 (2010).

[298] Strasberg, P. and Esposito, M. *Stochastic thermodynamics in the strong coupling regime: An unambiguous approach based on coarse graining.* Phys. Rev. E **95**, 062101 (2017).

[299] Fodor, É. and Marchetti, M. C. *The statistical physics of active matter: From self-catalytic colloids to living cells.* Physica A **504**, 106–120 (2018).

[300] Ricci, A., Crawford, A. and Fettiplace, R. *Active hair bundle motion linked to fast transducer adaptation in auditory hair cells.* Journal of Neuroscience **20**, 7131–7142 (2000).

[301] Barral, J., Dierkes, K., Lindner, B., Jülicher, F. and Martin, P. *Coupling a sensory hair-cell bundle to cyber clones enhances nonlinear amplification.* PNAS **107**, 8079–8084 (2010).

[302] Speck, T. *Thermodynamic approach to the self-diffusiophoresis of colloidal Janus particles.* Phys. Rev. E **99**, 060602(R) (2019).

[303] Blair, D. L., Neicu, T. and Kudrolli, A. *Vortices in vibrated granular rods.* Phys. Rev. E **67**, 031303 (2003).

[304] Bechinger, C. *et al. Active particles in complex and crowded environments.* Rev. Mod. Phys. **88**, 045006 (2016).

[305] Fodor, É. *et al. How far from equilibrium is active matter?* Phys. Rev. Lett. **117**, 038103 (2016).

[306] Takatori, S. C. and Brady, J. F. *Towards a thermodynamics of active matter.* Phys. Rev. E **91**, 032117 (2015).

[307] Mandal, D., Klymko, K. and DeWeese, M. R. *Entropy production and fluctuation theorems for active matter.* Phys. Rev. Lett. **119**, 258001 (2017).

[308] Bonilla, L. *Active Ornstein-Uhlenbeck particles.* Phys. Rev. E **100**, 022601 (2019).

[309] Marconi, U. M. B., Puglisi, A. and Maggi, C. *Heat, temperature and Clausius inequality in a model for active Brownian particles.* Sci. Rep. **7**, 46496 (2017).

[310] Nardini, C. *et al. Entropy production in field theories without time-reversal symmetry: quantifying the non-equilibrium character of active matter.* Phys. Rev. X **7**, 021007 (2017).

[311] Cates, M. E. *Diffusive transport without detailed balance in motile bacteria: does microbiology need statistical physics?* Rep. Prog. Phys. **75**, 042601 (2012).

[312] Sourjik, V. and Wingreen, N. S. *Responding to chemical gradients: bacterial chemotaxis.* Current opinion in cell biology **24**, 262–268 (2012).

[313] Bourret, R. B. and Stock, A. M. *Molecular information processing: lessons from bacterial chemotaxis.* J. Biol. Chem. **277**, 9625–9628 (2002).

[314] Ivlev, A., Bartnick, J., Heinen, M. and Löwen, H. *Statistical mechanics for nonreciprocal forces.* ArXiv:1403.2417 (2014).

[315] Fang, K. *et al. Generalized non-reciprocity in an optomechanical circuit via synthetic magnetism and reservoir engineering.* Nat. Phys. **13**, 465 (2017).

[316] Metelmann, A. and Clerk, A. A. *Nonreciprocal photon transmission and amplification via reservoir engineering.* Phys. Rev. X **5**, 021025 (2015).

[317] Bernier, N. R. *et al. Nonreciprocal reconfigurable microwave optomechanical circuit.* Nat. commu. **8**, 1–8 (2017).

[318] Wanjura, C. C., Brunelli, M. and Nunnenkamp, A. *Topological framework for directional amplification in driven-dissipative cavity arrays.* ArXiv:1909.11647 (2019).

[319] Chiacchio, E. R. and Nunnenkamp, A. *Dissipation-induced instabilities of a spinor Bose-Einstein condensate inside an optical cavity.* Phys. Rev. Lett. **122**, 193605 (2019).

[320] Loos, S. A. M., Claussen, J. C., Schöll, E. and Zakharova, A. *Chimera patterns under the impact of noise.* Phys. Rev. E **93**, 012209 (2016).

[321] Premalatha, K., Chandrasekar, V. K., Senthilvelan, M. and Lakshmanan, M. *Impact of symmetry breaking in networks of globally coupled oscillators.* Phys. Rev. E **91**, 052915 (2015).

[322] Bo, S., Del Giudice, M. and Celani, A. *Thermodynamic limits to information harvesting by sensory systems.* J. Stat. Mech. Theor. Exp. **2015**, P01014 (2015).

[323] Hartich, D., Barato, A. C. and Seifert, U. *Sensory capacity: An information theoretical measure of the performance of a sensor.* Phys. Rev. E **93**, 022116 (2016).

[324] Maxwell, J. C. *Maxwell on molecules and gases* (MIT Press, 1986).

[325] Frank, T. D. *Analytical results for fundamental time-delayed feedback systems subjected to multiplicative noise.* Phys. Rev. E **69**, 061104 (2004).

[326] Zheng, Y. and Sun, X. *Governing equations for Probability densities of stochastic differential equations with discrete time delays.* Discrete Continuous Dyn. Syst. Ser. B **22** (2017).

[327] Novikov, E. A. *Functionals and the random-force method in turbulence theory.* Sov. Phys. JETP **20**, 1290–1294 (1965).

[328] Vacchini, B., Smirne, A., Laine, E.-M., Piilo, J. and Breuer, H.-P. *Markovianity and non-Markovianity in quantum and classical systems.* New J. Phys. **13**, 093004 (2011).

[329] of Nature, C. M. *Global Biodiversity.* Bd. 7-8 (The Museum, 1997). URL `https://books.google.de/books?id=O5TzAAAAMAAJ`.

[330] Tichelli, A. and Rovó, A. *A Clever Person Solves a Problem; a Wise Person Avoids It.* Biology of Blood and Marrow Transplantation **19**, 1004–1005 (2013).

[331] Grabert, H., Graham, R. and Green, M. S. *Fluctuations and nonlinear irreversible processes. II.* Phys. Rev. A **21**, 2136 (1980).

[332] De Vries, B. and Príncipe, J. C. *A theory for neural networks with time delays.* In *Advances in neural information processing systems*, 162–168 (San Mateo, CA: Morgan Kaufmann, 1991).

[333] Sartori, P., Granger, L., Lee, C. F. and Horowitz, J. M. *Thermodynamic costs of information processing in sensory adaptation.* PLoS computational biology **10**, e1003974 (2014).

[334] Mandal, D., Quan, H. T. and Jarzynski, C. *Maxwells refrigerator: an exactly solvable model.* Phys. Rev. Lett. **111**, 030602 (2013).

[335] Barato, A. C. and Seifert, U. *An autonomous and reversible Maxwell's demon.* EPL **101**, 60001 (2013).

[336] Hartich, D., Barato, A. C. and Seifert, U. *Stochastic thermodynamics of bipartite systems: transfer entropy inequalities and a Maxwells demon interpretation.* J. Stat. Mech. Theor. Exp. **2014**, P02016 (2014).

[337] Boukobza, E. and Tannor, D. J. *Thermodynamics of bipartite systems: Application to light-matter interactions.* Phys. Rev. A **74**, 063823 (2006).

[338] Munakata, T. and Rosinberg, M. L. *Feedback cooling, measurement errors, and entropy production.* J. Stat. Mech. **13**, 06014 (2013).

[339] Argun, A. and Volpe, G. *Better stability with measurement errors.* J. Stat. Phys. **163**, 1477–1485 (2016).

[340] Kompaneets, R., Vladimirov, S., Ivlev, A. and Morfill, G. *Reciprocal interparticle attraction in complex plasmas with cold ion flows.* New J. Phys. **10**, 063018 (2008).

[341] Ivlev, A. V. *et al. Statistical mechanics where Newtons third law is broken.* Phys. Rev. X **5**, 011035 (2015).

[342] Dauchot, O. and Löwen, H. *Chemical Physics of Active Matter.* J. Chem. Phys. **151**, 114901 (2019).

[343] Roldán, É., Barral, J., Martin, P., Parrondo, J. M. and Jülicher, F. *Arrow of Time in Active Fluctuations.* ArXiv:1803.04743 (2018).

[344] Netz, R. R. *Fluctuation-dissipation relation and stationary distribution of an exactly solvable many-particle model for active biomatter far from equilibrium.* J. Chem. Phys. **148**, 185101 (2018).

[345] Allahverdyan, A. E., Janzing, D. and Mahler, G. *Thermodynamic efficiency of information and heat flow.* JSTAT **2009**, P09011 (2009).

[346] Landauer, R. *et al. Information is physical.* Physics Today **44**, 23–29 (1991).

[347] Steck, D. A., Jacobs, K., Mabuchi, H., Habib, S. and Bhattacharya, T. *Feedback cooling of atomic motion in cavity QED.* Phys. Rev. A **74**, 012322 (2006).

[348] Cohadon, P.-F., Heidmann, A. and Pinard, M. *Cooling of a mirror by radiation pressure.* Phys. Rev. Lett. **83**, 3174 (1999).

[349] Vinante, A. *et al. Feedback cooling of the normal modes of a massive electromechanical system to submillikelvin temperature.* Phys. Rev. Lett. **101**, 033601 (2008).

[350] Wittkowski, R., Löwen, H. and Brand, H. R. *Microscopic approach to entropy production.* J. Phys. A **46**, 355003 (2013).

[351] Lauth, G. and Kowalczyk, J. *Thermodynamik - Eine Einführung* (Springer, 2015).

[352] Lee, D. Y., Um, J., Paneru, G. and Pak, H. K. *An experimentally-achieved information-driven Brownian motor shows maximum power at the relaxation time.* Sci. Rep. **8**, 12121 (2018).

[353] Martínez, I. A., Roldán, É., Dinis, L. and Rica, R. A. *Colloidal heat engines: a review.* Soft Matter **13**, 22–36 (2017).

[354] Zakharova, A., Feoktistov, A., Vadivasova, T. and Schöll, E. *Coherence resonance and stochastic synchronization in a nonlinear circuit near a subcritical Hopf bifurcation.* EPJ ST **222**, 2481–2495 (2013).

[355] Gang, H., Ditzinger, T., Ning, C. and Haken, H. *Stochastic resonance without external periodic force.* Phys. Rev. Lett. **71**, 807 (1993).

[356] Geffert, P. M., Zakharova, A., Vüllings, A., Just, W. and Schöll, E. *Modulating coherence resonance in non-excitable systems by time-delayed feedback.* Eur. Phys. J. B. **87**, 291 (2014).

[357] Strasberg, P., Schaller, G., Brandes, T. and Esposito, M. *Thermodynamics of quantum-jump-conditioned feedback control.* Phys. Rev. E **88**, 062107 (2013).

[358] Kwon, C., Um, J. and Park, H. *Information thermodynamics for a multi-feedback process with time delay.* EPL **117**, 10011 (2017).

[359] Mittler, J. E., Sulzer, B., Neumann, A. U. and Perelson, A. S. *Influence of delayed viral production on viral dynamics in HIV-1 infected patients.* Math. Biosci. **152**, 143–163 (1998).

[360] Johnson, R. W. *Handbook of fluid dynamics* (Crc Press, 2016).

[361] Basset, A. B. *A treatise on hydrodynamics: with numerous examples*, vol. 2 (Deighton, Bell and Company, 1888).

[362] Guo, M. *et al. Probing the stochastic, motor-driven properties of the cytoplasm using force spectrum microscopy.* Cell **158**, 822–832 (2014).

[363] Kempes, C. P., Wolpert, D., Cohen, Z. and Pérez-Mercader, J. *The thermodynamic efficiency of computations made in cells across the range of life.* Philos. Trans. Royal Soc. A **375**, 20160343 (2017).

[364] Larger, L. *et al. High-speed photonic reservoir computing using a time-delay-based architecture: Million words per second classification.* Phys. Rev. X **7**, 011015 (2017).

[365] Li, J., Bai, K., Liu, L. and Yi, Y. *A deep learning based approach for analog hardware implementation of delayed feedback reservoir computing system.* In *2018 19th International Symposium on Quality Electronic Design (ISQED)*, 308–313 (IEEE, 2018).

[366] Röhm, A., Jaurigue, L. and Lüdge, K. *Reservoir Computing Using Laser Networks.* IEEE J. Sel. Top. Quant. **26**, 1–8 (2019).

[367] Röhm, A. and Lüdge, K. *Multiplexed networks: reservoir computing with virtual and real nodes.* J. Phys. Commun. **2**, 085007 (2018).

[368] Snider, G. L. *et al. Minimum energy for computation, theory vs. experiment.* In *2011 11th IEEE International Conference on Nanotechnology*, 478–481 (IEEE, 2011).

[369] Xu, X., Lu, Y. and Liang, Y. *Time-delay recurrent neural networks for dynamic systems control* (Springer, 2004).

[370] Barato, A. C. and Seifert, U. *Cost and precision of Brownian clocks.* Phys. Rev. X **6**, 041053 (2016).

[371] Kaiser, A., Snezhko, A. and Aranson, I. S. *Flocking ferromagnetic colloids.* Sci. Adv. **3**, e1601469 (2017).

[372] (2020). Weblink: `www.math.sci.hiroshima-u.ac.jp/~m-mat/MT/emt.html`.

[373] Hall, M. and Reginatto, M. *Ensembles on Configuration Space.* Fundam. Theor. Phys. **184** (2016).