ROYAL INSTITUTE
OF TECHNOLOGY

Dynamics and Thermodynamics of Translational and Rotational Diffusion Processes Driven out of Equilibrium

RAFFAELE MARINO

Doctoral Thesis
Stockholm, Sweden 2017

TRITA-CSC-A-2017:13
ISNN 1653-5723
ISRN-KTH/CSC/A–16/13-SE
ISBN 978-91-7729-405-4

KTH School of Computer Science and Communication
SE-100 44 Stockholm
SWEDEN

Akademisk avhandling som med tillstånd av Kungl Tekniska högskolan framlägges till offentlig granskning för avläggande av teknologie doktorsexamen i fysik 2017-06-15 klockan 10:00 i AlbaNova Universitetscentrum, Kungl Tekniska högskolan, Roslagstullsbacken 21, Stockholm, FD 5, D5:3008.

Tryck: Universitetsservice US AB

To my grandmother
To Iulia
To my family

Sammanfattning

Diffusionsprocesser är viktiga begrepp inom många vetenskapliga fält såsom fysik, biologi, finans, samhällsvetenskaper. Sådana processer beskriver hur en substans rör sig från områden med hög täthet till områden med låg täthet. Ett av de mest berömda exemplen på diffusionsprocesser är Brownsk rörelse. Brownsk rörelse är den oregelbundna och livliga rörelsen av en mesoskopisk partikel i en vätska.

I denna avhandling undersöks dynamiken och termodynamiken av diffusionsprocesser utanför jämvikt på mesoskopiska skalor. Vad gäller dynamiken så presenteras teorin fär Brownsk rörelse av en partikel som kan röra sig i rummet och rotera i tre dimensioner. Vidare beskrivs hur diffusionsprocesser på en n-dimensionell Riemannsk mångfald kan behandlas med användande av Kolmogorovs framåtekvation. Vad gäller termodynamik så beskrivs i avhandlingen hur termodynamiska storheter kan definieras på mesoskopiska skalor med användning av den Brownska teorin. Teorin för stokastisk energetik samt hur man beräknar entropiproduktionen längs en partikelbana beskrivs, vilket ger en introduktion till den området stokastisk termodynamik.

Vidare införs begreppet anomal entropiproduktion. Denna anomali i entropiproduktionen uppstår när diffusionsprocesser drivs ur jämvikt i ett rumsberoende temperaturfält. Termen är ett uttryck för att den överdämpade approximationen inte ger korrekta värden för termodynamiska kvantiteter. Felet är en korrektionsterm som inte gör någon skillnad för den överdämpade rörelsen i sig men som dominerar i entropiproduktion över lång tid, även i gränsen när partikelns massa är liten.

I den första delen av avhandlingen analyseras en ellipsformad Brownsk partikelns rörelse och rotationer i ett heterogent värmebad med utgångspunkt från stokastisk termodynamik. Det visas hur de Brownska rotationerna bidrar till entropiproduktionen och ger upphov till en ny term i den anomala entropin.

Avhandlingen beskriver också multiskalmetoden och homogeniseringsteori. Multiskalmetoden är ett sätt att med störningsräkning hitta approximativa lösningar till differentialekvationer när problemet har separarade längd och tidsskalor. Homogeniseringsterori är ett sätt att medelvärdesbilda över ett systems snabba processer.

I den senare delen av avhandlingen beskrivs en Brownsk partikels rörelse i ett homogent värmebad på vilken verkar en konstant kraft. Analysen, som använder multiskalmetoden och homogeniseringsteori, utgår från en överdämpad beskrivning av rörelse i rummet och rotationer, och ger som resultat en effektiv beskrivning på stora skalor i rummet endast. Dessutom studeras specialfallet av en Brownsk ellipsoid i en lutande potential.

Abstract

Diffusion processes play an important role in describing systems in many fields of science, as in physics, biology, finance and social science. Diffusion processes formally describe the motion of a quantity from a region of high concentration to another with low concentration. One of the most famous example of diffusion process is the Brownian motion. At mesoscopic scale, Brownian theory describes the very irregular and animated motion of a particle suspended on a fluid.

In this thesis the dynamics and thermodynamics of diffusion processes driven out of equilibrium, at mesoscopic scale, are investigated. For dynamics, theory of Brownian motion for a particle which is able to rotate and translate in three dimensions is presented. Moreover, it is presented how to treat diffusion process on a n-dimensional Riemann manifolds defining the Kolmogorov forward equation on a such manifold.

For thermodynamics, this thesis describes how to define thermodynamics quantities at mesoscopic scale using the tools of Brownian theory. The theory of stochastic energetics and how to compute entropy production along a trajectory are presented introducing the new field of stochastic thermodynamics.

Moreover, the "*anomalous entropy production*" is introduced. This anomaly in the entropy production arises when diffusion processes are driven out of equilibrium by a space dependent temperature field. The presence of this term expresses the fallacy of the overdamped approximation in computing thermodynamic quantities. This fallacy is given by a correction that does not have any impact on trajectories, but controls the entropy production in the long run, even in the small inertia limit.

In the first part of the thesis the translational and rotational motion of an ellipsoidal particle in an heterogeneous thermal environment, with a space-dependent temperature field, is analyzed from the point of view of stochastic thermodynamics. It is shown how the Brownian rotational motion contributes to the entropy production, with also an additional term in the “*anomalous*” entropy.

The thesis also describes the multiscale method and homogenization theory. The multiscale method is a perturbative method to find approximate solutions of differential equations, which describe the motion of a system in presence of separation of length and time scales. The homogenization theory allows to average out fast processes for dynamics of the system.

In the final part of the thesis the motion of a Brownian rigid body three dimensional space in an homogeneous thermal environment under the presence of an external force field is analyzed, using multiscale method and homogenization theory starting from an overdamped description of the translational and rotational motion. As result, effective equations describing only the translational long-term motion of the rigid body are presented. The particular case of a Brownian ellipsoid in a external tilted periodic potential is also studied.

Contents

Acknowledgments

Tempus fugit old Romans said. Four years of efforts end in this Ph.D thesis.

I have been able to do that through the constant and passionating supervision of my supervisor Prof. Dr. Erik Aurell. He has been more than a supervisor, he has been a guide, a friend, an intellectual father, in one word he has been a *magister*, in the old Latin meaning.

He has been like *Socrates* for *Plato*, he used the maieutics method to teach me how to do research. He has been like *Virgilius* for *Dante*, he guided me through the hell of ignorance to see the light of knowledge.

No words exist to explain how much I thank you, dear Erik, for your unique and passionating supervision. You have been always present in this my journey, from the first step to the end. You have always encouraged me to think by myself, to deal with difficult research problems without fear. I learnt a lot from you.

Thank you to leave me free to think, to leave me free to choose my own project. Thank you to show me how to ask the right questions for research topics, and how to seek the right answers. With your example I am a better scientist. You showed me how to speak during workshops, and how to write, in a simple and clear way, scientific papers. I have been like clay in your hands, and you shaped me tirelessly. I am proud to be being one of yours Ph.D. students.

I have been lucky in this Ph.D. journey, I have had a co-supervisor and I want to thank him. He is Prof. Dr. Ralf Eichhorn. He has been a friend for me, I learnt a lot from him. He has taught me the German school of physics. He has also been able to show me how to think, how to make science. He has been always present in my journey and helped me a lot. He used a lot of examples that I will never forget, he has spent a lot of time to teach me everything that he knows. For this reason I want to thank you, dear Ralf, with all my heart.

Thank you to be patient with me and to teach me how to be careful in each calculation. Thank you to leave me to make mistakes and after, like a friend, to show me the right direction. Thank you to teach me how to start form the principles of a problem and how to look at a problem as a whole. Thank you also to show me how to divide a complex problem into simple problems. Thank you for everything, dear Ralf.

I have to thank also who gave me the spark of this scientific life, he has been my supervisor in the Master degree and co-author of a paper. He is Prof. Giorgio

Parisi. He chose me between many, and gave me my first love in theoretical physics, the K-SAT problem. The first love is unforgettable.

Thank you, dear Giorgio, to choose me, to guide me, to be you. Thank you to accept me, to believed in me. You are the example to follow.

A sincerely thanks goes to a co-author of a paper, Prof. Dr. Federico Ricci-Tersenghi. He was the opponent of my Master defense and he has taught me the amazing world of the computational complexity. Thank you, dear Federico.

Two others co-authors I want to thank, Prof. Dr. Marcelo Dias and Dr. Stefano Bo. Thank you to share with me your knowledge and to give me your time. It has been an amazing experience to discuss with you. Thank you Stefano and Marcelo.

I want also to thank Prof. Angelo Vulpiani. We lost together in a mushroom hunting, but he showed me how to find the right answers in physics. I will never forget your teachings and that discussion in all my future life. Thank you, dear Angelo.

A particularly thanks goes to Prof. Dr. Irene Giardina and Prof. Miguel Angel Virasoro, both my supervisors in my Bachelor degree.

During this period here at Nordita I met wonderful people. I want to thank first of all Prof. Dr. Katherine Freese and Prof. Dr. Alexander (Sasha) V. Balatsky for the amazing discussions that we had at Nordita, with the hope to meet you again.

I want also to thank Prof. Paolo Di Vecchia for the scientific and political discussions, Prof. Dr. Konstantin Zarembo for the many issues about tensor algebra and Prof. Dr. John Hertz for the discussions about life and physics.

I want to thank also all the administrative, faculty positions and fellows here at Nordita.

However, a personal thanks goes to Dr. Francesco Mancarella for the discussion about math and finance, I hope all the best for you.

A huge thanks to all my friends here at Nordita, Stanislav, Sergey, Chris, Matthias, Saikat, Martina, Livia for the funny discussions. A special thanks to Gino and Nicolas former Ph.D. students with me in the Prof. Erik Aurell's group, and to Sarah ex officemate.

A special thanks goes to Dr. Francesco Taddia, colleague and friend, with the hope to meet and to eat together again.

I want to thank all my family for the support that they gave me in all this time, and my love Iulia.

A very good luck goes to all my colleagues and friends in Europe and out of Europe.

Finally I want to thank all my friends of Nunziatella and my country, with the hope to see you soon.

Raffaele Marino
May 12, 2017

E quindi uscimmo a riveder le stelle

–(Hell XXXIV, 139), Dante Alighieri

Introduction

> *[...] so I thought, as opposed to him in this small extent I am wiser: that what I do not know, in no way I think I know.*
>
> *–Apology of Socrates, Plato*

Since the earliest days of thermodynamics, processes driven out of equilibrium have always interested scientific community. How to build a theory which describes processes driven out of equilibrium, like thermodynamics does at equilibrium, is still an open question.

Thermodynamics is a theory which describes macroscopically systems at equilibrium. Using the words of J. Kurchan, "*thermodynamics is a 'gift' of mechanics*" [1]. The motion of the microscopic components of a macroscopic system is the fundamental ingredient for thermodynamics. Indeed, by this, it is possible to derive the same thermodynamics quantities using the tools of statistical mechanics.

But if the system is not so large to be considered macroscopic, and not so small to be considered microscopic, how does the thermodynamics describe it? To fill this gap physicists have developed stochastic thermodynamics.

Stochastic thermodynamics deals with systems that are in the "between" scale, i.e. they are too big to be considered microscopic, and too small to be considered macroscopic. This scale is usually referred to the *mesoscopic* scale.

At this scale the dynamics of a system is not reproducible by the usual deterministic law of classical mechanics, but it is described by stochastic theories. The most famous one is the Brownian motion.

This thesis deals with dynamics and thermodynamics of diffusion processes, which perform translational and rotational Brownian motion, driven out of equilibrium. It is divided in two main parts, chapters $1-5$ about the dynamics of diffusion processes and chapters 6 and 7 about thermodynamics.

In the dynamics part, I introduce the mesoscopic scale, which is the physical setting of this thesis, in Ch. 1. I deal with Brownian motion, from the origin to today in Ch. 2. In Ch. 3 and Ch 4, I discuss the rotational motion of a Brownian rigid body. The part of dynamics ends with Ch. 5, which describes homogenization theory and multiscale methods used to perform the analysis in the published papers.

The thermodynamics part starts with Ch. 6. In this chapter, I recall the main results obtained in stochastic energetics with particular emphasis about the First Law of thermodynamics at mesoscopic scale. In Ch. 7, I conclude the thesis with the explanation of the entropy production at mesoscopic scale, and with an introduction of how the anomalous entropy production arises.

The papers on which the thesis is based are included.

Thesis outline

- Chapter 1: **The Mesoscopic Scale**

  The first chapter describes the setting of the thesis. I introduce the mesoscopic scale to give a general presentation of the physical world as treated in this thesis. I also give an overview of the concepts of length and time scales for processes which occur at mesoscopic scale. I end this chapter with an example of length and time scale that appear in the advection-diffusion process on the mesoscopic scale.

- Chapter 2: **Brownian Motion**

  In this chapter I recall the theory of Brownian motion. The chapter starts with a description of the theory developed by Einstein and Smoluchowski. I also introduce the Langevin equation in the underdamped and the overdamped case. I show how to derive the Langevin equation in overdamped case in the small inertia limit, and I end with a discussion about the connection between **Paper I**, **II**, **III** and **IV** and the Brownian theory.

- Chapter 3: **Rotational Brownian Motion**

  Rotational Brownian motion is discussed in this chapter. I discuss the results obtained in the literature, describing how to write down a rotational diffusion equation from physical arguments. I recall the Langevin approach, which also has been used in **Paper I**, when inertia cannot be neglected, to describe the rotational Brownian motion. The discussion ends with examples in physics, chemistry and biology.

- Chapter 4: **Diffusion on Manifolds**

  The theory of diffusion on manifolds is introduced in this chapter. The discussion follows the Russian school of Kolmogorov. I discuss also the works of K. Itô, K. Yosida and H. P. Jr. McKean describing the theory of Brownian motion on Lie groups. I end up writing down a Langevin equation valid on

Lie algebras and "injected" on the group by the McKean-Gangolli's Injection method. This method has been used also in **Paper III**

- Chapter 5: **Homogenization and Multiscale Methods**

  This chapter, principally, deals with the analytical methodology used to perform the research in **Paper I**, **II**, **III** and **IV**. I make a simple introduction to homogenization theory and the multiscale methods, and I apply them to a simple advection-diffusion process. I conclude the chapter with a connection to published papers.

- Chapter 6: **Stochastic Energetics**

  In this chapter I introduce stochastic energetics. I start with a nutshell of thermodynamics and I discuss the First Law of thermodynamics at trajectory level at mesoscopic scale, following the pioneering contribution of Sekimoto [2]. This chapter introduces some concept which are used in **Paper I**.

- Chapter 7: **Stochastic Thermodynamics and Entropy Production**

  In this chapter I discuss entropy production. An historical overview of entropy in thermodynamics and statistical mechanics is given to introduce entropy production in stochastic thermodynamics. A section for the fluctuation theorem is also given. This chapter ends with a discussion about anomalous entropy production, which is the main result obtained in **Paper I**.

List of papers included in the thesis

Three peer-reviewed papers, and a manuscript in preparation are included in this thesis.

Paper I:	Raffaele Marino, Ralf Eichhorn, and Erik Aurell. "Entropy production of a Brownian ellipsoid in the overdamped limit." *Physical Review E 93.1 (2016): 012132.*
Paper II:	Erik Aurell, Stefano Bo, Marcelo Dias, Ralf Eichhorn, and Raffaele Marino. "Diffusion of a Brownian ellipsoid in a force field." *EPL (Europhysics Letters) 114.3 (2016): 30005.*
Paper III:	Raffaele Marino, and Erik Aurell. "Advective-diffusive motion on large scales from small-scale dynamics with an internal symmetry." *Physical Review E 93.6 (2016): 062147.*
Paper IV:	Raffaele Marino, Ralf Eichhorn. "Brownian motion of an ellipsoidal particle in a tilted periodic potential: long-term velocity and diffusion." [In preparation] (June 2017).

Additional paper outside of the scope of this thesis

During my Ph.D. studying period I also have nurtured my passion for the K-SAT problem. This passion has become a paper in October 2016, when I published **Paper V**, see below.

The paper is an extension of the research that I performed during my Master degree. In this paper we have presented a new algorithm, the Backtracking Survey Propagation algorithm, to solve instances for the K-SAT problem, i.e. a constraint satisfaction problem. We show that the Backtracking Survey Propagation algorithm, in a time practically linear in the problem size, is able to find solutions very close to the threshold, in a region unreachable by any other algorithm.

This paper has been mentioned on the web-page "ResearchItaly, the portal of Italian research" of Italian Ministry of Education, University and Research, in the section "Success stories" [3].

Paper V: Raffaele Marino, Giorgio Parisi, and Federico Ricci-Tersenghi. "The backtracking survey propagation algorithm for solving random K-SAT problems." *Nature Communications 7 (2016): 12996.*

Chapter 1

The Mesoscopic Scale

> *The stochastic description in terms of macroscopic variables*
> *will be called mesoscopic. It comprises both*
> *the deterministic laws and the fluctuations about them.*
>
> *–N. G. Van Kampen*

In this chapter we introduce the setting of the thesis. We discuss the mesoscopic scale in general in Sec. 1.1, and introduce the concepts of length scales and time scales in, respectively, Sec. 1.2 and Sec. 1.3, following Gray and Miller in [4]. We conclude this chapter in Sec. 1.4 with an example of length and time scale, that appear in the advection-diffusion process at mesoscopic scale.

1.1 General remarks

The term "mesoscopic", a word coined by Van Kampen in 1981 [5] from the old Greek $\mu\epsilon\sigma o\zeta$ meaning "middle" or "intermediate", refers to intermediate between long and short, and applies to microstructure to be seen in "between" the microscopic and the macroscopic scale [6].

The macroscopic world is composed by objects which can be seen at naked eye. This world is often considered while using the "thermodynamics limit" (system volume, V, and particle numbers, N, tending to infinity with $n = \frac{N}{V}$ kept constant), which is a convenient mathematical device for obtaining bulk properties [7]. It is governed by classical physics.

The microscopic world contains atoms and molecules. We know that they exist, but we cannot see them directly with our eyes. This world is interpreted by quantum physics.

The interest in studying systems at mesoscopic scale, as suggested by Imry [7], is not only in order to understand the macroscopic limit and how it is achieved from the microscopic world, but it is in existence of many novel phenomena which are intrinsic to mesoscopic systems. A mesoscopic system can be defined as a "large molecule", but it is always (at least weakly) coupled to a much larger, essentially infinite, system [7].

The boundaries between microscopic−mesoscopic and between mesoscopic−macroscopic are not sharp. They can be approximately indicated around $1-10$ nm for the first and around 100 μm for the second [8]. As a physical example, a protein is a mesoscopic object with a typical length around 2 nm, while a bacterium, e.g. E. coli, is a mesoscopic object with a typical length around 2 μm. Both these objects cannot be seen at naked eye, because the resolution of a human eye is around 0.1 mm (order of magnitude of the diameter of a human hair).

Mesoscopic and macroscopic objects share the property that they are composed by a huge number of atoms, but they differ substantially. The main difference, according to [9], is that a macroscopic object can be characterized by the average properties of the material from which it is constituted. For example, the free expansion of an ideal gas is well described by the statistical mechanics and by the classical laws of thermodynamics.

In contrast, the mesoscopic scale is the size at which fluctuations have a significant effect on the behaviour of a system and must be taken into account in its analysis. Indeed, a mesoscopic object is so small that deterministic classical physics laws do not reproduce anymore the motion of the system.

A simple and useful example, described also by Zwanzig in [10], is the Brownian motion of a particle suspended on a fluid. Let's imagine to have a very small spherical polystyrene particle, e.g. a spherical particle with radius $a = 0.25$ μm, suspended on water. For a sake of simplicity the motion occurs only along one dimension. When we look at the particle, we observe that it is in a very irregular state of motion, generated by collisions with the water molecules, which never ends.

By Newton's law of the classical physics, the equation of motion of the particle is:

$$m\frac{dv(t)}{dt} = F(t), \tag{1.1}$$

where m identifies the mass of the particle, $v(t)$ its velocity and $F(t)$ is total instantaneous force on the particle at time t. This force is given by the interaction of the Brownian particle with the surrounding medium.

By classical physics, $F(t)$ is, in principle, dominated by a frictional force $-\gamma v$, proportional to the velocity of the particle. The friction constant γ is given by Stokes's law, $\gamma = 6\pi\eta a$ where η is the dynamic viscosity of the medium. The solution of (1.1), under this assumption, is $v(t) = v(0)\exp\{-\gamma t/m\}$. According to this, the velocity of the Brownian particle decays to zero when time goes to infinity.

In contrast, the equipartition theorem of energy shows that the mean squared velocity of the particle at thermal equilibrium is $< v^2 >= k_B T/m$, where k_B is Boltzmann's constant and T is the absolute temperature of the fluid. We infer, therefore, that the actual velocity cannot remain at zero and that $F(t)$ cannot be only dominated by the frictional force.

Microscopic and mesoscopic objects differ by the fact that a mesoscopic system is composed by a number of atoms such that fluctuations in the number of atoms contained in the system are not important. The mesoscopic physics contains a huge number of quantum effects, but these effects are out of the scope of the thesis and will therefore not be treated.

The main message, here, is that the mesoscopic physics treats problems which occur when a physical macroscopic system is miniaturized [7]. At this scale the deterministic laws of classical physics do not reproduce anymore the evolution of the system and stochastic theories to describe physical events must be introduced.

1.2 Length scales

Mesoscopic systems, which we are interested in, can span length scales that range from few nanometers to few micrometers. In deciding what processes are considered in a system of interest, the issue of the scale at which the process will be studied is important. Identification of a length scale is related to which features must be described in a process [4].

The concept of length scale, according to Gray and Miller in [4], is important for physical phenomena where different length scales do not affect each other. If this happens, length scales are said to be separated. In presence of separation of length scales it is possible, therefore, to derive a self-consistent theory that only describes the relevant length scales for a given problem [4]. In philosophy of science this is treated by scientific reductionism. Scientific reductionism claims that the effective description at larger length scales can be described by the physical laws on the shortest length scales [6]. To quantify the idea that one can describe physics at different length scales from one another, physicists have introduced the multiscale analysis.

Separation of scales means that one scale of observation is separated from another by order of magnitude. This quality implies the property that variation in the measurement of a features in a process is not negligible if and only if the length scale changes by several orders of magnitude with respect to changes in the length scale of the measurement [4].

Length scale considerations are also important for measuring other properties of interest such as temperature, velocity, and transport coefficients. When gradient of a property of interest is large, the average value of that property depends strongly on the length scale of measurement [4].

In this thesis, we will consider systems where the length scales are widely separated such that variables and parameters in the equations introduce the concept of hierarchy of length scale. For this reason we identify the following different separated length scales:

$$\ell_{micro} << \underbrace{\ell_{fast} << \ell_{intermediate} << \ell_{slow}}_{\ell_{meso}} << \ell_{macro}. \tag{1.2}$$

These scales, from the smallest to the largest, are referred to, respectively, as microscopic scale, mesoscopic scale and macroscopic scale. The smallest and the largest, which we are not interested in, identify the length scales where, respectively, quantum physics and classical physics hold.

The mesoscopic length scale has been divided in three different scales, respectively *fast, intermediate* and *slow* scale. Although the actual magnitudes of these scales will differ from system to system and are not connected to any specific and universally applicable length [4], this qualification of scales is useful for discussion of mesoscopic systems in this thesis. In next chapters, when the discussion will need to, we will give a quantification of these scales.

As a general point of view the three scales that appear at ℓ_{meso} can be qualitatively justified in the following manner: the *fast* length scale is associated to degrees of freedom of the system that reach an equilibrium distribution faster than other degrees of freedom of the system. The *slow* length scale, instead, refers to degrees of freedom such that are the slowest to reach an equilibrium distribution, if this distribution exists. The *intermediate* length scale is a scale that connects the $fast$ and *slow* length scale. Sometime, this depends on the process under study in the system, the $\ell_{intermediate}$ length scale could not be present. We introduce this scale to be more general in the discussion of length scales.

1.3 Time scales

As in the previous section, here we present, the concept of temporal scales in a mesoscopic system. According to Gray and Miller in [4], the temporal scales appear in the system by the phenomena to be modelled, or, sometimes, by the spatial scales of interest. For example, time scales of a chemical process can be related to reaction rates, the time needed to reach equilibrium. Some reaction occurs before others, and therefore the time needed to achieve a reaction is the time scale of that faster reaction.

Many processes, like biological and chemical reactions, advection, diffusion, and others are the starting point for identification of characteristic time scale of a system [4]. The characteristic time (or "*relaxation time*") can be defined as the amount of time that a system needs to relax under an external "*stimulus*".

A mesoscopic system may have multiple time scales depending on the processes that are occurring. In this thesis three different types of time scales denoted, respectively, as the *fast*, *intermediate* and *slow* scales and designated, respectively, with τ_{fast}, $\tau_{intermediate}$ and τ_{slow} are considered. Also for temporal scales, as for length scales, we have a widely separation between them and they satisfy the conditions:

$$\tau_{fast} << \tau_{intermediate} << \tau_{slow}. \tag{1.3}$$

A *fast* time scale is a measure of the time needed for a degree of freedom of a system to approach an equilibrium configuration while the *slow* time scale refers to a degree of freedom that is the slowest to reach an equilibrium distribution, if this distribution exists. The *intermediate* time scale, as discussed before for the length scale, is a temporal scale that connects the *fast* and *slow* time scale.

1.4 Example: the advection-diffusion process

Here we introduce the advection-diffusion process to help the reader in clarity the previously described concepts about time and length scales.

The advection–diffusion process is a combination of the diffusion and convection (advection) process, and describes physical phenomena where particles, energy, or other physical quantities are transferred inside a physical system due to two processes: diffusion and convection. Convection refers to the movement of a substance within a medium. Diffusion is the movement of the substance from an area of high concentration to an area of low concentration.

This process is one of the most frequently used models in science. The advection-diffusion process is modeled by a partial differential equation that can be derived using the physical law of mass conservation and Gauss' theorem. We first introduce the model and, hence, we show up the different time and length scales of the process.

The model

Let's consider a substance of mass M, e.g. a dissolving tracer, in a moving fluid. We assume also that no chemical reactions occurs during the transport process. Such a motion in a control volume V is governed by the effects of fluid motion (advection) and by the movement from an area of high concentration to an area of low concentration (diffusion).

We define, following Socolofsky and Jirka in [11], $C(\vec{x},t)$ to be the concentration of the substance and D the diffusion coefficient assumed constant in each direction. The relation between mass and concentration is $M = \int_V C(\vec{x},t)\ dV$. From Fick's law we know that the flux given by only the diffusion process is $\vec{J}_{\text{diff}} = -D\vec{\nabla}C(\vec{x},t)$. The flux given by only the advection process, instead, is $\vec{J}_{\text{advec}} = \vec{v}(\vec{x},t)C(\vec{x},t)$. The vector field $\vec{v}(\vec{x},t)$ is the velocity field of the fluid. We assume that the fluid velocity is smooth, steady, and periodic in all direction and that it is incompressible. Thus $\vec{v}(\vec{x},t) = \vec{v}(\vec{x})$ with $\vec{\nabla}\cdot\vec{v}(\vec{x}) = 0$. The total flux of concentration along the directions of motion is, therefore, equal to [11]:

$$\vec{J}(\vec{x},t) = \vec{J}_{\text{advec}} + \vec{J}_{\text{diff}} = \vec{v}(\vec{x})C(\vec{x},t) - D\vec{\nabla}C(\vec{x},t). \tag{1.4}$$

To derive the advection diffusion equation we use the flux law and the conservation principle of mass. From the conservation of mass, i.e. "*the total mass in any volume element of the system can only change if matter flows into (or out of) the volume element*" [12], one knows:

$$\int_V \frac{\partial C(\vec{x},t)}{\partial t}\, dV = \int_\Omega \vec{J}(\vec{x},t)\cdot\vec{n}\, d\Omega, \tag{1.5}$$

where $\vec{n}$ is the vector normal to the surface Ω and counted positive from the inside to the outside. Applying Gauss' theorem to the surface integral we obtain the advection-diffusion equation:

$$\frac{\partial C(\vec{x},t)}{\partial t} + \vec{\nabla}\cdot(\vec{v}(\vec{x})C(\vec{x},t)) = D\vec{\nabla}\cdot\vec{\nabla}C(\vec{x},t). \tag{1.6}$$

Length and time scales

Having developed the model for the advection-diffusion process, we display the different time scales and length scales of the process. To do that we use a dimensional analysis on equation (1.6). This analysis is frequently used in turbulence and in other branches of physics [13–15].

As first step, we split each variable in a dimensionless variable and a constant term that contains all the information about the dimension of the variable. For example, the variable $\vec{x} = L_{\mathrm{ref}}\hat{\vec{x}}$, $t = \tau_{\mathrm{ref}}\hat{t}$, $D = D_{\mathrm{ref}}\hat{D}$, $\vec{v} = v_{\mathrm{ref}}\hat{\vec{v}}$, and $C = C_{\mathrm{ref}}\hat{C}$, where variables address with $\hat{}$ identify dimensionless terms, while the coefficient in front of them has the full dimensional value. The reference values can be considered as local values during the process where τ_{ref} comparisons are being made at L_{ref} length scale. Equation (1.6), under this assumption, becomes:

$$\frac{1}{\tau_{\mathrm{ref}}}\frac{\partial \hat{C}(\hat{\vec{x}},\hat{t})}{\partial \hat{t}} + \frac{v_{\mathrm{ref}}}{L_{\mathrm{ref}}}\hat{\vec{\nabla}}\cdot(\hat{\vec{v}}(\hat{\vec{x}})\hat{C}(\hat{\vec{x}},\hat{t})) = \frac{D_{\mathrm{ref}}}{L_{\mathrm{ref}}^2}\hat{D}\hat{\vec{\nabla}}\cdot\hat{\vec{\nabla}}\hat{C}(\hat{\vec{x}},\hat{t}). \tag{1.7}$$

Comparing dimensional terms, one obtains the characteristic advection and diffusion time scales. The advection (subscript $_a$) and diffusion (subscript $_d$) characteristic temporal scales are [11]:

$$\begin{aligned} &\text{Advection: } \tau_a = \frac{L_{\mathrm{ref}}}{v_{\mathrm{ref}}};\\ &\text{Diffusion: } \tau_d = \frac{(L_{\mathrm{ref}})^2}{D_{\mathrm{ref}}}. \end{aligned} \tag{1.8}$$

Therefore τ_a and τ_d identify, respectively, the mean time that the substance needs to advect and to diffuse over scale L_{ref}.

The ratio between the diffusion and the advection characteristic temporal scale, gives a dimensionless number:

$$\frac{\tau_d}{\tau_a} = \frac{L_{\mathrm{ref}}v_{\mathrm{ref}}}{D_{\mathrm{ref}}} = \mathbf{Pe}. \tag{1.9}$$

In the study of transport phenomena in a continuum **Pe** is called the Péclet number [16]. The Péclet number is defined to be the ratio of advective transport rate of a physical quantity to the diffusive transport rate of the same quantity [17]. In the context of mass transfer, the Péclet number can be also expressed as the product of the Reynolds number and the Schmidt number [17].

Comparing scaling terms for advection and diffusion, we observe that they have different length scale dependence. Thus, each term dominates at a different length scale. For example when **Pe** $<<1$, diffusion is dominant at small scale because of the high molecular velocities relative to the particles of medium. The magnitude of this length scale can be quantified by $\ell_d = \sqrt{D_{\mathrm{ref}}\tau_d}$. At larger scale the advection process becomes relevant, while the diffusion are inefficient. At this scale the mean path that the particles of the substance covers by convection is quantified by $\ell_a = v_{\mathrm{ref}}\tau_a$.

In contrast, when **Pe** $>>1$, the advection process is dominant with respect to the diffusion process in the system, at small scale. In the case where **Pe** $=1$ the two processes in the system are not separated. In this case, however, the scale separation can be imposed between the time scale of the process and the time scale of observation, to which are associated the respectively length scales.

Chapter 2

Brownian Motion

> *I did not believe that it was possible*
> *to study the Brownian motion with a such precision.*
>
> *–A. Einstein*

In this chapter we recall the theory of Brownian motion. We start to describe in Sec. 2.1 the physical aspects of the theory developed by Einstein and Smoluchowski. In Sec. 2.2 we introduce the Langevin approach to the Brownian motion and we end the section introducing the Langevin-Kramers equation for a Brownian particle. In Sec. 2.3 we describe how to write down the Fokker-Planck equation from the Langevin-Kramers equation, and we also describe the main result of the Brownian theory, i.e. the overdamped dynamics. We end this chapter with the connections with the papers published in Sec. 2.4.

2.1 Brownian motion: the origin

The theory of Brownian motion, developed in different formulations by Einstein, Smoluchowski and Langevin between 1905 and 1906 [18–20], and reviewed in many monographs [5, 21–27], describes the "*random*" motion of bodies suspended in a fluid. According to Zwanzig [10], the theory of Brownian motion describes the dynamics of non equilibrium system in the simplest way.

A simple example is a very small polystyrene particle in a glass of water. The particle is in a very irregular state of motion, generated by collisions with the water molecules. This is dubbed *Brownian motion*.

Historically the discovery of this irregular and very animated state of motion is due to the botanist Robert Brownian in 1827. He made this discovery when he was examining pollen grains to elucidate the mechanism which by which the grains moved towards the ova when fertilising flowers [28]. No theory was given to explain this phenomenon, until 1905 when Einstein published his explanation [18]. The same explanation was independently developed by Smoluchowski [19], who was responsible for many experiments checks to verify the Brownian theory [28].

To explain the Brownian motion, Einstein started by observing that the motion of a pollen is induced by the frequent and never ending collisions on the pollen of the molecules of the fluid in which it is suspended, in our case water. He also pointed out, in his paper [18], that the specific motion of the molecules of the fluid is so complicated to be described precisely, and therefore its effect on the particle can be obtained probabilistically, in term of independent and identical impacts with the pollen. These assumptions allowed him to describe the Brownian motion in terms of a diffusive process.

To make the connection with a diffusion process explicit, Einstein used a kinematic argumentations. In his paper of 1905 [18], Einstein noted that in a time scale τ the motions of n grains occur in successive time intervals as independent events. He defined the function $C(\vec{x}, t)$ as the function that describes the number of grains in a volume element $d\vec{x}$ at time t. To simplify our discussion we follow Einstein's paper [18] and, therefore, we confine ourselves in one dimension and in absence of external force.

He observed that in the time scale τ the x-coordinate of the single particle will increase of a distance Δ for each particle. The probability that a given grain will randomly move by Δ along x-axis in the time interval τ is given by the function $\phi(\Delta)$, which satisfies the following properties:

$$\int_{-\infty}^{+\infty} d\Delta \phi(\Delta) = 1\,, \quad \phi(\Delta) = \phi(-\Delta). \tag{2.1}$$

The distribution of the particles at time $t + \tau$ from the distribution at time t will be given by the equation [18]:

$$C(x, t+\tau)dx = dx \int_{-\infty}^{+\infty} d\Delta C(x+\Delta, t)\phi(\Delta). \tag{2.2}$$

Considering experimental observation times that are larger compared to τ, and noticing that $\phi(\Delta)$ mostly differs from zero for small values of Δ, one can Taylor expand both sides of equation (2.2) in t and in x. The diffusion equation (the first derivative averages out to zero) [18, 29] then reads:

$$\frac{\partial}{\partial t} C(x,t) = D \frac{\partial^2}{\partial x^2} C(x,t); \quad D = \frac{1}{\tau} \int_{-\infty}^{+\infty} d\Delta \frac{\Delta^2}{2} \phi(\Delta)\,, \tag{2.3}$$

where D is the diffusion coefficient defined in this way by Einstein.

The solution of the equation (2.3) can be easily found, given $C(x,0) = C_0\delta(x)$ as initial condition, with the characteristic function. The solution is a Gaussian:

$$C(x,t) = C_0 \frac{1}{\sqrt{4\pi D t}} \exp\left[-\frac{x^2}{4Dt}\right]. \tag{2.4}$$

The mean square displacement along the x-axis is linear in t rather than t^2 as in ballistic motion and is given by $\langle x^2 \rangle = 2Dt$, where $\langle \cdot \rangle$ is the average with respect to $C(x,t)$. The diffusion coefficient follows Einstein relation, namely $D = k_B T \gamma^{-1}$ where $k_B T$ is the thermal energy, k_B is Boltzmann's constant and T the temperature, and γ^{-1} is the friction coefficient given by the Stokes' law for a spherical particle, $\gamma = 6\pi\eta a$, with a the radius of the particle and η the viscosity of the fluid. For example the friction coefficient of a polystyrene spherical particle of radius $a = 0.25\mu$m, which performs Brownian motion in one dimension in water, i.e. $\eta = 0.89$ fN s$/\mu$m^2, is equal to $\gamma = 4.2$ fN s$/\mu$m.

It is also possible to introduce an external force field u, produced by a potential U, i.e. $u = -\frac{\partial}{\partial x}U$, in the process. If we also define with $v = u\gamma^{-1}$ the velocity field of the process the diffusion equation becomes the Fokker-Planck equation*, which has the form of a continuity equation:

$$\frac{\partial}{\partial t} C(x,t) = \frac{\partial}{\partial x}\left[\left(-v + D\frac{\partial}{\partial x}\right) C(x,t)\right]; \tag{2.5}$$

and in three dimension looks like equation (1.6) in Ch. 1.

2.2 Langevin approach to Brownian motion

The Langevin equation

In 1906 Langevin [20] gave his alternative formulation about the Brownian motion. As in the previous section we confine us in one dimension to simplify the discussion.

Langevin observed, from statistical mechanics, the mean kinetic energy of a Brownian grain of mass m should achieve at equilibrium the value given by the equipartition theorem, i.e. $\langle \frac{1}{2} m v^2 \rangle = \frac{1}{2} k_B T$. He also pointed out that on the grain two forces act, the viscous drag and the fluctuation force.

The viscous drag is the force proportional to the velocity and acts on the opposite direction of the motion of the particle. He assumed that this force has the same formula of the classical physics, i.e. $-\gamma \frac{dx}{dt}$ where γ has been defined in previous section.

The fluctuation force, indicated by X, was inserted by hand in the equation of motion and represents the "*in perpetuum*" collision, i.e. frequent and incessant impacts, of the molecules of the liquid with the Brownian motion [20]. The knowledge about X is very little. Langevin made few statistical hypotheses to explain this

---

*Also known as Kolmogorov forward equation.

point. First of all, X is a fluctuation about a mean, so, it must, itself, have zero mean $\langle X \rangle = 0$ [21]. Second, X is a stationary process with a very short correlation time. The correlation time is shorter than the only characteristic time of the process $\tau = m/\gamma$ [21]. In the modern scientific terminology $X(t)$ is a "*Gaussian noise*" with zero mean and correlation function given by $\langle X(t)X(t')\rangle = 2B\delta(t-t')$, where B is a "*measure of the strength*" of the fluctuation force [10].

The equation of motion for the position of the particle is given by Newton's law as:

$$m\frac{d^2x}{dt^2} = -\gamma\frac{dx}{dt} + X. \tag{2.6}$$

Multiplying both sides of equation (2.6) by x, observing that $x\frac{dx}{dt} = \frac{1}{2}\frac{dx^2}{dt}$, $x\frac{d^2x}{dt^2} = \frac{1}{2}\frac{d^2x^2}{dt^2} - (\frac{dx}{dt})^2$, recalling $v = \frac{dx}{dt}$, averaging over a large number of different particles, using the equipartition theorem cited before, and the properties of X, we obtain:

$$\frac{d^2\langle x^2\rangle}{dt^2} = 2\frac{k_BT}{m} - \frac{1}{\tau}\frac{d\langle x^2\rangle}{dt}, \tag{2.7}$$

which describe the evolution in time of the average value of the square of the displacement [29].

Recalling $y = \frac{d\langle x^2\rangle}{dt}$ and solving the first order differential equation associated, we obtain, returning to the average value of the square of the displacement, that:

$$\langle x^2(t)\rangle = 2\frac{k_BT}{\gamma}t - Z_0\tau \mathrm{e}^{-\frac{t}{\tau}}, \tag{2.8}$$

where Z_0 is a suitable integration constant.

Equation (2.8) tells us that after an exponential transient with typical time given by the characteristic time of the process, i.e. $\tau = m\gamma^{-1}$, the mean square displacement increases linearly with the time with diffusion coefficient given by $D = k_BT\gamma^{-1}$, in agreement with Einstein's mean square displacement, obtained from the solution of the diffusion equation.

The diffusion coefficient, however, can be computed also from the correlation function of the velocity $v = \frac{dx}{dt}$. From classical mechanics, we know that, the position is given by the equation of motion $x = \int_0^t v(t)$. The displacement square therefore will be given by :

$$\langle x^2(t)\rangle = \int_0^t dt' \int_0^t dt'' \langle v(t')v(t'')\rangle. \tag{2.9}$$

Performing the time derivative and combining two equivalent terms on the right-hand side, we can compute the diffusion coefficient as

$$D = \frac{1}{2}\lim_{t\to\infty}\frac{d}{dt}\langle x^2(t)\rangle = \lim_{t\to\infty}\int_0^t dt' \langle v(t')v(t)\rangle. \tag{2.10}$$

This is given by the observation that the correlation function of the velocity decays to zero in a short time, with an order of magnitude much smaller than the typical characteristic time (for example in liquids this may order of picoseconds [10]). As we have described before, the average value of the square of the displacement agrees with Einstein's mean square displacement after an exponential transient with typical time given by the characteristic time of the process. This allows us to derive the simplest example of the relation of a transport coefficient, i.e. the diffusion coefficient, to a time correlation function:

$$D = \int_0^\infty dt \langle v(t) v(0) \rangle. \tag{2.11}$$

The last equation is obtained under the observation that at equilibrium the velocity correlation function can depend only on the time difference and cannot depend on any arbitrary origin of the time axis [10].

The Fluctuation-dissipation theorem

The value of B as a measure of the strength of the fluctuation force can be easily computed form physical observations.

We know that at equilibrium equipartition theorem tells us that $\langle v^2 \rangle = \frac{k_B T}{m}$. The Langevin equation described in equation (2.6) can be rewritten in the following form:

$$m \frac{dv}{dt} = -\gamma v + X. \tag{2.12}$$

A formal integration of equation (2.12) gives us the following result:

$$v(t) = v_0 \mathrm{e}^{-\frac{t}{\tau}} + \frac{1}{m} \int_0^t dt' \mathrm{e}^{-\frac{t-t'}{\tau}} X(t'), \tag{2.13}$$

where $\tau = m\gamma^{-1}$. The first term of the equation decays to zero in a time much larger than the characteristic time of the process. The second term gives the amount of the velocity produced by the random force X.

We want to compute the average of the square velocity, to be able to apply the equipartition theorem and, therefore, able to quantify the strength of the fluctuation force. To do that, we square the equation (2.13), we average it and, applying the statistical properties of the random force, we end with:

$$\langle v^2(t) \rangle = v_0^2 \mathrm{e}^{-\frac{2t}{\tau}} + \frac{B}{\gamma m}(1 - \mathrm{e}^{-\frac{2t}{\tau}}). \tag{2.14}$$

Finally in the long time limit the l.h.s. of the equation (2.14) approaches the value given by the equipartition theorem, i.e. $\langle v^2 \rangle = \frac{k_B T}{m}$, while the r.h.s becomes equal to $\frac{B}{\gamma m}$. We, finally, obtain:

$$B = \gamma k_B T. \tag{2.15}$$

This result is known as the Fluctuation-dissipation theorem. It relates the strength of random force X to the magnitude of the friction coefficient γ. Using the words of Zwanzig:

"*It expresses the balance between friction, which tends to drive any system to a completely* ***dead*** *state, and noise, which tends to keep the system* ***alive***" [10].

The Langevin-Kramers equation

Let's now take the motion of a particle in a fluctuating environment in one dimension. We introduce, this time, an external force which is non linear. For example, we want to describe the evolution of Brownian particle subjected to an external force field dependent by the position, i.e. $u^{(\mathrm{ext})}(x) = -V(x)'$, where $V'(x)$ is a gradient of a non linear potential. A typical example is Brownian motion of a molecular dipole in a periodic potential [10]. In this case the Langevin equation is called Langevin-Kramers equation and it is given by:

$$\frac{dx}{dt} = v, \tag{2.16a}$$

$$m\frac{dv}{dt} = u^{(\mathrm{ext})}(x) - \gamma v + X. \tag{2.16b}$$

where X is the physical fluctuation force. The drag force is still linear in the velocity, and the noise is related to friction from the Fluctuation-dissipation theorem as described before. This formulation is also called the underdamped dynamics, from which, in the next section, we start to show the main result of the Brownian theory: the *overdamped* approximation.

2.3 From the underdamped to the overdamped dynamics

The dynamics of a Brownian particle can be described by a set of stochastic second-order differential equations, i.e. equations (2.16). This description is usually called the underdamped dynamics. This underdamped mathematical approach can be substituted with the formalism of the overdamped dynamics.

The overdamped dynamics is an approximation, which allows neglecting the inertial term if the mass of the particle is small or if the friction is large, reducing thus the set of stochastic second-order differential equations to a set of first-order differential equations. This reduction from second-order to first-order allows treating the problem in a simpler way.

For example, if we set the mass to zero in the Langevin-Kramers equation, i.e. in equation (2.16b), we obtain the overdamped dynamics and the stochastic second-order differential equations of the motion of the system are mapped into a stochastic first-order differential equation, i.e. equation (2.18). In a such situation, indeed, the velocity disappears and the system is fully described only by the equation of the position.

From this set of equations one can also write down a partial differential equation called Fokker-Planck equation. The Fokker-Planck equation is a second-order parabolic partial differential equation, and for a solution one needs initial conditions and boundary conditions. The stochastic process described by a conditional probability satisfying the Fokker-Planck equation is equivalent to the Langevin integral equation, where the integral is evaluated in the Itô sense.

Formally, to the Langevin-Kramers equation, i.e. equation (2.16), is associated the following Fokker-Planck equation:

$$\frac{\partial}{\partial t}p = -\frac{\partial}{\partial x}(vp) + \frac{1}{m}\frac{\partial}{\partial v}\Big([-u^{(\text{ext})}(x) + \gamma v]p\Big) + \frac{\gamma k_B T}{m^2}\frac{\partial^2}{\partial v^2}p, \tag{2.17}$$

where $p = p(x, v, t)$ is the probability distribution function of the Brownian particle that has a certain position x and velocity v at time t.

If we set the mass to zero in equation (2.16b) we end up with:

$$\gamma\frac{dx}{dt} = u^{(\text{ext})}(x) + X. \tag{2.18}$$

To this Langevin equation is associated the following Fokker-Planck equation:

$$\frac{\partial}{\partial t}\rho = -\frac{1}{\gamma}\frac{\partial}{\partial x}(u^{(\text{ext})}\rho) + \frac{k_B T}{\gamma}\frac{\partial^2}{\partial x^2}\rho, \tag{2.19}$$

where $\rho = \rho(x, t)$ is the probability distribution function of the Brownian particle that has a certain position x at time t.

The formal limit that allows to go from the underdamped dynamics to the overdamped dynamics is called the small inertia limit. It can be explained in the following manner.

One starts to observe the presence of variables in the system, to which are associated two different time scales, i.e. a fast time scale and a slow time scale.

To the fast time scale, which is the characteristic time of the process, is associated the velocity variable. This time scale has an order of magnitude of $\tau_{fast} = \tau = m/\gamma$. In the case that the mass of the particle is small, or the friction coefficient is large, τ_{fast} is smaller than the time of the observation τ_{slow}, which is referred to the time scale of the position x. The difference of order of magnitudes between the time scales tells us that we are in a presence of separation of scales.

In other words, because the fast degree of freedom reaches in a time τ_{fast} an equilibrium configuration, the corresponding probability distribution is statistically independent from the probability distribution that describes the slow degree of freedom, i.e. the position x of the particle.

Using therefore the homogenization theory, described in Ch. 5, it is possible to average out the fast degree of freedom in the equation and to end up with an effective equation, which describes the evolution of the system only with the slow degree of freedom present in the dynamics.

The Fokker-Planck equation in (2.17), describing the evolution of the probability density as a function of the position and the velocity, can be mapped into an effective Fokker-Planck equation describing the evolution of the probability density as a function of the slow degree of freedom, i.e. the position x.

The effective Fokker-Planck equation is called also *overdamped* Fokker-Planck and the stochastic process described by the conditional probability satisfying the overdamped equation in (2.19) is equivalent to the Langevin equation[†], expressed in the form of equation (2.18)

The overdamped dynamics at equilibrium also allows computing thermodynamics quantities as well as the underdamped dynamics. But what happen if the temperature depends by the position? The curiosity of the reader will be satisfied in the last section of the last chapter of this thesis.

2.4 Connection to papers

The theory of Brownian motion has been fully used in Paper I,II, III, and IV.

In Paper I, we have described with the Langevin-Kramers equation the motion of a Brownian ellipsoidal particle, which can translate and rotate in three dimensions. We have applied the homogenization theory, i.e. we have performed the small inertia limit, from the underdamped Fokker-Planck associated and we have ended up with the overdamped Fokker-Planck equation. Moreover, in the same paper, we have also computed the thermodynamics in the underdamped and overdamped case, and we have shown that the overdamped approximation fails in the evaluation of the thermodynamics quantities, if the system is driven out of equilibrium by the presence of a gradient of temperature.

In Paper II, III and IV, we have described the system in the overdamped case for a Brownian particle, which can translate and rotate in the space, as a starting point. We, using the homogenization theory, have averaged out the rotational degrees of freedom from the overdamped Fokker-Planck equation associated describing rotational and translational motion and we have ended up with an effective Fokker-Planck equation describing only the translational effective dynamics. Moreover we have computed analytically the diffusion tensors and we have verified numerically that the average displacement square grows linearly with the time.

[†] The Langevin equation has to be interpreted in Itô sense.

Chapter 3

Rotational Brownian Motion

> *False words are not only evil in themselves,*
> *but they infect the soul with evil.*
>
> *–Socrates*

In this chapter we introduce rotational Brownian motion. We start by deriving in Sec. 3.1 a Fokker-Planck equation, of rotational Brownian motion in non inertial limit, by physical arguments. We introduce, therefore, the inertial case, writing down a Langevin approach for angular velocity and orientational degrees of freedom in Sec. 3.2. In Sec. 3.3 we compute the rotational diffusion tensor using the arguments of Sec. 3.1 and 3.2. We end up in Sec. 3.4 our discussion of rotational Brownian motion with useful examples in physics, chemistry and biology.

3.1 Rotational diffusion

The rotational diffusion studies the change of orientation of a particle due to random torques exerted on it by the surroundings. Rotational diffusion is important, for example, for describing the motion of a non spherical particle during sedimentation and aggregation during the process of sedimentation [30–32].

Historically, rotational Brownian motion of a sphere around a given axis was investigated by Einstein in 1906 [18, 33]. However, only in 1929, Debye [34] gave a theory of rotational diffusion in a constant external field based on the Smoluchowski equation. He assumed that collisions are so frequent that a particle can rotate through only a very small angle before having another reorienting collision.

In this section we consider rotations in three-dimensional space and we show how to develop a theory describing rotational diffusion [35–39].

General remarks

The complications with respect to the translational Brownian motion (see Ch. 2) arise from the non commutativity of rotations and from the finite extent of the space of orientations [35].

The rotational problem is more complicated than the translational one also because the orientation of a rigid body cannot be specified by a vector the time derivative of which is the angular velocity of the body [38]. However, if the orientation is specified by some appropriate coordinates, it is possible to derive a Fokker-Planck equation for the distribution function for the orientation and angular velocity [38, 40].

The orientation of a rigid body is completely determined by fixing one set of coordinates in the body and another set in the laboratory, and specifying the rotation which transforms one set into the other [36].

The orientation of a rigid body can be represented in different ways. For example, using Eulerian angles [41–43], quaternions [35, 44], Cayley-Klein parameters [36, 41], units vectors ("directors") attached to the particle [45, 46], and finally differential geometry and the compact manifold $SO(3)$ [47, 48], see also Appendix A.

Furry in 1957 [35] treated the problem of the rotational Brownian motion of a spherical rigid body representing rotations by quaternions [44], and ended up with a rotational diffusion equation for the random rotational motion of spheres.

Favro in 1960 [36] developed a rotational diffusion theory for a rigid body of arbitrary shape. The description of rotations used by this author is different from the Furry's representation. Indeed, Favro derived the independent variables, for the description of rotation, by the Cayley-Klein parameters [41].

Valiev and Ivanov, instead, in their review on rotational Brownian motion [39] describe a rotation by a rotation vector [49] $\vec{\Omega}$, where the absolute value $|\vec{\Omega}|$ is equal to the angle of rotation, while the direction of the vector $\vec{\Omega}$ is that of the axis of rotation [39].

The Ariadne's thread between the papers is that, in each representation of rotations, any finite rotation can be looked upon as a succession of infinitesimal rotations. The corresponding rotation is, therefore, the product of the corresponding infinitesimal transformations [41, 50].

An infinitesimal rotation is an orthogonal transformation of coordinate axes in which the components of a vector are almost the same in both sets of axes (see Appendix A). Moreover, infinitesimal rotations are a useful tool for describing the evolution in time of a rigid body [41].

According to Goldstein [41], the change in differential time element dt of the total angular momentum, defined with $\vec{L}$, with respect to the body axes will vary only by the effects of the rotation of the body axes from the change with respect to the laboratory axes. This change is in the components of the vector, and it is given by an infinitesimal coordinate rotation.

The connection between the differential total momentum observed in the laboratory frame $(d\vec{L})_{\mathrm{lab}}$, and the differential total momentum observed in the body frame $(d\vec{L})_{\mathrm{body}}$ is given, thus, by:

$$(d\vec{L})_{\mathrm{lab}} = (d\vec{L})_{\mathrm{body}} + d\vec{\alpha} \times \vec{L} \tag{3.1}$$

where $d\vec{\alpha}$ represents a vector composed by the three infinitesimal parameters specifying the representation of rotation (see Appendix A).

Dividing the equation (3.1) by the differential time element dt under consideration, we obtain the time rate of change $\vec{L}$ [41]:

$$\left(\frac{d\vec{L}}{dt}\right)_{\mathrm{lab}} = \left(\frac{d\vec{L}}{dt}\right)_{\mathrm{body}} + \vec{\omega} \times \vec{L}. \tag{3.2}$$

The vector $\vec{\omega}$ is called by Goldstein [41] the instantaneous "*angular velocity*" of the body, i.e. the instantaneous angular rate of rotation of the body, and is defined by the relation $d\vec{\alpha} = \vec{\omega} dt$. The angular velocity lies along the direction of the "*instantaneous axis of rotation*", i.e. the axis of the infinitesimal rotation occurring between the time t and $t + dt$ [41] .

The components of angular momentum along the principal axes are proportional to the components of the angular velocity, the factors of proportionality being the principal moment of inertia, i.e. $\vec{L} = \boldsymbol{I}\vec{\omega}$. $\boldsymbol{I}$ is the moment of inertia of the rigid body. It is a tensor of second rank. We choose the body frame because the inertia tensor $\boldsymbol{I}$ is independent of particle orientation in this frame.

The fundamental equation of motion for the total angular momentum of a rigid body in the body frame, therefore, reads:

$$\left(\boldsymbol{I}\frac{d\vec{\omega}}{dt}\right) + \vec{\omega} \times (\boldsymbol{I}\vec{\omega}) = \vec{N}^{(e)}. \tag{3.3}$$

The total external torque $\vec{N}^{(e)}$ is the sum of the all applied torques on the system. Equations (3.3), in classical mechanics, are also called Euler's rotation equations for the rigid body dynamics.

Rotational diffusion equation

For deriving a rotational diffusion equation we follow the works of Favro [36] and Valiev and Ivanov [39]. A finite rotation can be represented by a rotational operator acting on a probability distribution.

The rotation $\boldsymbol{R}_z(\alpha)$ of an angle α about z-axis transforms a distribution function $\psi(x, y, z)$ into $\boldsymbol{R}_z(\alpha)[\psi(x, y, z)] = \psi(x\cos(\alpha) + y\sin(\alpha), -x\sin(\alpha) + y\cos(\alpha), z)$ [50].

When α is small $\alpha = \epsilon << 1$ we can develop the distribution function in Taylor series:

$$\boldsymbol{R}_z(\epsilon)[\psi(x,y,z)] \sim \psi(x,y,z) + \epsilon\left(y\frac{\partial\psi}{\partial x} - x\frac{\partial\psi}{\partial y}\right), \tag{3.4}$$

and, taking only terms of the first order in ϵ, we obtain the infinitesimal operator $\boldsymbol{R}_z(\epsilon) = 1 + \epsilon\left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\right)$.

Carrying out similar calculation for $\boldsymbol{R}_x(\epsilon)$ and $\boldsymbol{R}_y(\epsilon)$ infinitesimal operators, we note that they can be written as a vector product of $\vec{x}$ and $\vec{\nabla}$, i.e. $\vec{x} \times \vec{\nabla}$. As expressed before, any finite rotation can be looked upon as a succession of infinitesimal rotations (see Appendix A), i.e.

$$\boldsymbol{R}(\vec{\alpha}) = \mathrm{e}^{-i\vec{\alpha}\cdot\vec{\mathcal{J}}}, \tag{3.5}$$

where $\vec{\alpha}$ is the array collecting the angles of rotation, and $\vec{\mathcal{J}} = \vec{x} \times \frac{1}{i}\vec{\nabla}^*$, which has the same form of the angular momentum operators of quantum mechanics when multiplied by $\hbar$, the Plank's constant h divided by 2π.

For obtaining a rotational diffusion equation, hence, we postulate, as made in [35, 36, 39], that in a small enough time Δt the probability that a rigid body performs more than one rotation is zero. In other words, the angle of rotation that a rigid body performs in a time Δt is small because collisions are so frequent that a particle can rotate through only a very small angle before having another reorienting collision.

We now consider an ensemble of similar rigid bodies each of which undergoes a large number of small random rotations in any macroscopic interval of time [36]. It is assumed that the probability of a body's undergoing the rotation, represented by vector $\vec{\epsilon}$, with absolute value $|\vec{\epsilon}|$ such that $\sqrt{\vec{\epsilon}^{\intercal}\vec{\epsilon}} << 1$, Fig. 3.1 (we normalize the angle of rotation with π to be conformed with representation of rotations of a rigid body given by Wigner [51]), in the range $(\vec{\epsilon}, \vec{\epsilon} + d\vec{\epsilon})$, in a time Δt is given in body coordinate system by $p(\vec{\epsilon}, \Delta t)\mathrm{d}\vec{\epsilon}$ [36, 39].

Here body coordinates are used, because important quantities associated with the shape of the body (inertia tensor, diffusion tensor, etc.) are constant when referred to these coordinates [36].

The distribution probability $p(\vec{\epsilon}, \Delta t)\mathrm{d}\vec{\epsilon}$ satisfies following requirements:

1. $p(\vec{\epsilon}, \Delta t)$ does not depend on the orientation which the rigid body had before the beginning of the rotation [39];
2. $p(\vec{\epsilon}, \Delta t) = p(-\vec{\epsilon}, \Delta t)$;
3. $\int \mathrm{d}\vec{\epsilon}\, p(\vec{\epsilon}, \Delta t)\, \epsilon_i\epsilon_j$ is proportional to Δt. Under this assumption the rotational diffusion tensor $D_{ij}^{(\mathrm{rot})}$ can be defined as the second moment of the distribution

---

[*] A word of clarification is needed about the imaginary unit i in the exponential and in the operator $\vec{\mathcal{J}}$. We adopt this convention to express all the rotational operators in the same way along all the thesis (see Appendix A).

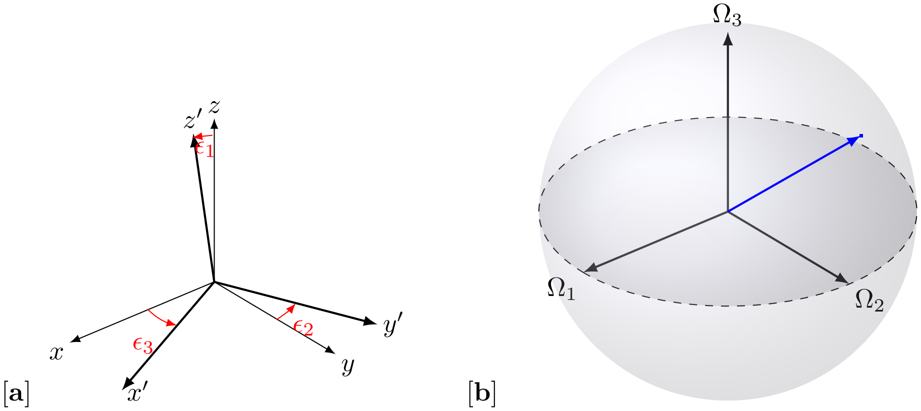

Figure 3.1: [**a**] Schematic picture of a rotation represented by a vector $\vec{\epsilon}^{\intercal} = (\epsilon_1, \epsilon_2, \epsilon_3)$. The reference frame $Oxyz$ is assumed fixed at the center of mass of the rigid body and identifies the orientation of the Brownian particle at $t = 0$. At time $t + \Delta t$ collisions are so frequent that the particle has performed a small rotation $\vec{\epsilon}$ and the orientation of the particle is identified by the primed reference frame $Ox'y'z'$. [**b**] Representation of the possible rotations of a rigid body given by Wigner [51]: "*each rotation is represented by a point inside, or on the surface of, a sphere of unit radius. The distance from the origin to the point is ϕ/π, $\phi < \pi$ being the angle of rotation from the standard or initial orientation. The direction of the vector from origin to representative point gives the direction of the axis of rotation, in agreement with the right hand screw-rule. Surface points at opposite ends of a diameter represent the same rotation (change of orientation), since rotations through π and $-\pi$ give the same result*".

density $p(\vec{\epsilon}, \Delta t)$, divided by Δt, i.e.:

$$D_{ij}^{(\mathrm{rot})} = \frac{1}{2} \int \mathrm{d}\vec{\epsilon}\, p(\vec{\epsilon}, \Delta t) \frac{\epsilon_i \epsilon_j}{\Delta t}. \tag{3.6}$$

According to equation (3.6), the tensor $D_{ij}^{(\mathrm{rot})}$ is symmetric and like every symmetric second-rank tensor, it can be transformed to principal axes [39].

4. Δt is small enough that moments of the distribution $p(\vec{\epsilon}, \Delta t)$ higher than the seconds may be neglected [36].

The first requirement means that the particle does not have any memory about the previous rotation, i.e. the stochastic process is a Markovian process [5,28]. The second one reflects the homogeneity of the thermal bath, i.e. the torque acting on the rigid body are random moments of forces caused by interaction with surrounding molecules in a uniform medium [39]. These two condition are obviously satisfied if there are no any external torques applied to the rigid body tending to orient the body in any specific direction, i.e. the rigid body is free [35]. By the symmetry of the probability distribution function we know the mean and the median occur at the point about which the symmetry occurs, which we assume to be zero.

We define with $P(\vec{\Omega}, t)$ the probability density per unit weight [36] that the orientation of the rigid body at time t belongs to the invariant volume element $\mathrm{d}V = \Gamma^{-1}\mathrm{d}\vec{\Omega}$ [36, 39]. The vector $\vec{\Omega}$ identifies the orientation of the rigid body while $\Gamma^{-1} = 4\sin^2(\Omega/2)/\Omega^2$ [39]. Let $W(\vec{\Omega}, t)d\vec{\Omega}$ be the probability of finding the orientation of a randomly chosen member of the ensemble in the volume $d\vec{\Omega}$ at some time t after the establishment of the ensemble [36].

The obvious connection between $P(\vec{\Omega}, t)$ and $W(\vec{\Omega}, t)$ is $P(\vec{\Omega}, t) = \Gamma W(\vec{\Omega}, t)$ [39].

The assumption of Markovian property for the stochastic process allows us to write the probability of finding the orientation of a rigid body at time $t+\Delta t$ belongs to the distribution $W(\vec{\Omega}, t+\Delta t)d\vec{\Omega}$ as:

$$W(\vec{\Omega}, t+\Delta t)d\vec{\Omega} = \int \mathrm{d}\vec{\epsilon}\, p(\vec{\epsilon}, \Delta t) W(\vec{\Omega}_0, t) d\vec{\Omega}_0, \tag{3.7}$$

i.e. adding the contributions given by all possible elementary rotations $\vec{\epsilon}$ which may take place in the time Δt [36]. The range of $\vec{\Omega}_0$ includes only those orientations which lead to the orientation $\vec{\Omega}$ after the rotation $\vec{\epsilon}$. Thus, by using the inverse of the operator defined in equation (3.5), we obtain:

$$W(\vec{\Omega}_0, t)d\vec{\Omega}_0 = \mathrm{e}^{i\vec{\epsilon}\cdot\vec{\mathcal{J}}} W(\vec{\Omega}, t) d\vec{\Omega}. \tag{3.8}$$

Substituting equation (3.8) into (3.7) and changing from $W(\vec{\Omega}, t)$ to $P(\vec{\Omega}, t)$ we obtain:

$$P(\vec{\Omega}, t+\Delta t)\Gamma^{-1} d\vec{\Omega} = \int \mathrm{d}\vec{\epsilon}\, p(\vec{\epsilon}, \Delta t) \mathrm{e}^{i\vec{\epsilon}\cdot\vec{\mathcal{J}}} P(\vec{\Omega}, t)\Gamma^{-1} d\vec{\Omega}, \tag{3.9}$$

where $\vec{\mathcal{J}}$ plays the role of the angular momentum operator in analogy of quantum mechanics [39].

By Liouville's theorem the equation (3.9) may be valid for each invariant volume element, and therefore may be now written in the form:

$$P(\vec{\Omega}, t+\Delta t) = \int \mathrm{d}\vec{\epsilon}\, p(\vec{\epsilon}, \Delta t) \mathrm{e}^{i\vec{\epsilon}\cdot\vec{\mathcal{J}}} P(\vec{\Omega}, t). \tag{3.10}$$

Expanding $\mathrm{e}^{i\vec{\epsilon}\cdot\vec{\mathcal{J}}} P(\vec{\Omega}, t)$ in powers of $\vec{\epsilon}$, collecting terms up to, and including, the second powers, and expanding $P(\vec{\Omega}, t+\Delta t)$ in powers of Δt [39], we get:

$$\frac{\partial P(\vec{\Omega}, t)}{\partial t} = -\mathcal{J}_j D_{jk}^{(\mathrm{rot})} \mathcal{J}_k P(\vec{\Omega}, t). \tag{3.11}$$

We interpret equation (3.11) as the diffusion equation, in analogy of the ordinary translational diffusion equation [35,36,39], with $D_{ij}^{(\mathrm{rot})}$ the rotational diffusion tensor.

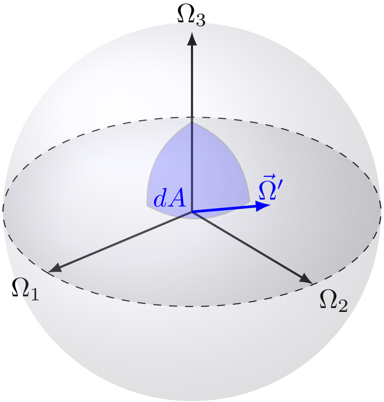

Figure 3.2: In the picture an element of surface dA, normal to the direction of Ω_l at the point $\vec{\Omega}'$ in the orientations space, is shown. The general element of surface dA is plotted in blue. The vector $\vec{\Omega}'$ passes through the surface dA and it contributes to the current through the surface.

The equation of rotational diffusion (3.11) for the orientation distribution $P(\vec{\Omega}, t)$ is independent by the representation used to describe the rotation. Indeed, if the representation of rotation is not specified by the vector $\vec{\Omega}$ but by Euler angles ϕ, θ, ψ the rotational diffusion equation looks like as in the form of (3.11) but with different infinitesimal operators $\vec{\mathcal{J}}$ [39].

The equation written in (3.11) can be rewritten, in analogy with the translational case, also in term of a divergence of a current density in the space of orientations.

To do that, we consider an element of surface dA, normal to the direction of increasing of a general axis Ω_l in the space of orientations at the point $\vec{\Omega}$ [36] (Fig. 3.2). Let's assume that the rigid body performs small rotations. All such rotations that allow the point in the space of orientations to go through the surface dA will contribute to the current through the surface, and we collect all these rotations into the set $\{\vec{\epsilon}\}$ [36]. In other words, the set $\{\vec{\epsilon}\}$ contains only small rotations, of some fixed size $|\vec{\epsilon}|$, that reach or pass the surface dA, i.e. it can be described as the union of two groups of small rotations $\{\vec{\epsilon}\} = \{\vec{\epsilon}_{\text{on}}\} \cup \{\vec{\epsilon}_{\text{out}}\}$. The set $\{\vec{\epsilon}_{\text{on}}\} \in \{\vec{\epsilon}\}$ contains all such rotations $\vec{\epsilon}$ which stop themselves on the surface dA, while $\{\vec{\epsilon}_{\text{out}}\} \in \{\vec{\epsilon}\}$ contains all such rotations $\vec{\epsilon}$ which pass through the surface and beyond. The fraction of rotations in $\{\vec{\epsilon}_{\text{on}}\}$ can be identified with β_1, while the fraction of rotations in $\{\vec{\epsilon}_{\text{out}}\}$ with β_2. Observing that $\beta_1 + \beta_2 = 1$, we infer the simple relation that $\beta_2 = 1 - \beta_1$. In other words, the range of β_1 goes from 0 to 1.

The probability transferred through the surface is $P(\vec{\Omega}', t)\Gamma^{-1}d\vec{\Omega}'$ where $\vec{\Omega}'$ is the point of origin of the rotation [36]. The rotation of $\vec{\Omega}'$, as discussed previously, includes those orientations which comes from the orientation $\vec{\Omega}$ after a small rotation

$\vec{\epsilon} \in \{\vec{\epsilon}_{\text{on}}\}$. Modifying ad hoc the rotational operator defined in equation (3.5), we may write:

$$P(\vec{\Omega}', t)\Gamma^{-1} d\vec{\Omega}' = \mathrm{e}^{i\beta_1 \vec{\epsilon}\cdot\vec{\mathcal{J}}} P(\vec{\Omega}, t)\Gamma^{-1} d\vec{\Omega}. \tag{3.12}$$

The invariant volume element is on $\vec{\Omega}$ and not anymore in $\vec{\Omega}'$ [36]. It is identified by a small parallelepiped placed on the surface dA and oriented in the direction of the change in $\vec{\Omega}$ under the rotation $d\beta_1\vec{\epsilon}$ [36, 39], i.e. $\Gamma^{-1}d\vec{\Omega} = \Gamma^{-1}dA\Delta\Omega_l = \Gamma^{-1}dAd\beta_1(-i\vec{\epsilon}\cdot\vec{\mathcal{J}}\Omega_l)$, where the last equivalence is given by the observation that the rotational operator must be infinitesimal. The probability passed through the surface given by all small rotations $\vec{\epsilon} \in \{\vec{\epsilon}_{\text{on}}\}$ in the time Δt is given by $p(\vec{\epsilon}, \Delta t)d\vec{\epsilon}\int_0^1 d\beta_1\Gamma^{-1}dA(-i\vec{\epsilon}\cdot\vec{\mathcal{J}}\Omega_l)\mathrm{e}^{i\beta_1\vec{\epsilon}\cdot\vec{\mathcal{J}}}P(\vec{\Omega}, t)$ and must be integrated over all possible values of $\vec{\epsilon}$ to get the total probability passed through the surface dA [36]. Dividing the result of integration by $\Delta t dA$, we can write the rotational diffusion equation, i.e. in Einstein's summation convention, as:

$$\frac{\partial P}{\partial t} = -\Gamma\partial_l J_l, \tag{3.13}$$

where $\partial_l J_l$ is the divergence of the current density. The current density is equal to:

$$J_l = \Gamma^{-1}(\mathcal{J}_j\Omega_l)D^{(\text{rot})}_{jk}\mathcal{J}_k P(\vec{\Omega}, t). \tag{3.14}$$

In Ch. 4, where the works about diffusion on manifolds of Kolmogorov [52, 53] are discussed, we will derive a Fokker-Planck equation on manifolds, which will also describe rotational diffusion processes.

3.2 The Langevin approach

In this section we describe rotational Brownian motion using a Langevin approach. In Ch. 2 we have derived the translational Brownian motion of a particle introducing, by hand, a random force in the Newton's second law of motion, because of the equipartition theorem. Here we will follow the same idea, i.e. we will derive a Langevin equation for rotational degrees of freedom from the fundamental equation of motion for the total angular momentum of a rigid body, and we will introduce by hand, a rotational random torque, to model the thermal fluctuations.

Our setting is always the mesoscopic scale, and for simplicity, without loss of generality, we assume that our rigid body is a particle that performs only random rotations around its principal axes, i.e. the particle center of mass motion is absent. A classical rotation is represented by a succession of infinitesimal rotations, see Appendix A.

The fundamental equation of motion for the total angular momentum of a rigid body in the body frame is given by Euler's equations in (3.3). The total external torque $N^{(e)}$, in this case, is composed by three terms: the viscous friction, the external applied and the thermal noise one.

The viscous friction torque is directed in the opposite verse of the motion of the rigid body and it is proportional to the angular velocity $\vec{\omega}$. It is given by $-\boldsymbol{\eta}\vec{\omega}$, where $\boldsymbol{\eta}$ is the rotational friction tensor. The rotational friction tensor is a 3×3 symmetric and positive definite matrix with a unique symmetric and positive defined square root $\boldsymbol{\eta}^{1/2}$ [54], meaning $(\boldsymbol{\eta}^{1/2})^{\mathsf{T}}\boldsymbol{\eta}^{1/2} = \boldsymbol{\eta}$ (the symbol $^{\mathsf{T}}$ labels the matrix transpose).

The external applied torque represents the deterministic torques acting on the particle. For a sake of simplicity we restrict ourselves in the case where no deterministic torques are applied on the rigid body. Only the viscous friction torque and random one are used to balance rotational motion.

The thermal fluctuations represent the incessant collisions of the molecules of the surrounding medium on the Brownian rigid body [28] which allow the particle to perform random rotations. The strength of these thermal fluctuations is given by $\sqrt{2k_BT}\boldsymbol{\eta}^{1/2}$, where k_B is the Boltzmann constant and T the temperature of the bath. The thermal bath is assumed to be homogeneous. The fluctuations are modeled by a vector $\vec{\xi}(t)$, which are unbiased Gaussian noise sources, i.e. zero mean and correlation function equal to $\langle \xi_i(t)\xi_j(t') \rangle = \delta_{ij}\delta(t-t')$. The fundamental equation of motion for the total angular momentum of a Brownian rigid body, therefore, looks like:

$$\left(\boldsymbol{I}\frac{d\vec{\omega}}{dt}\right) + \vec{\omega} \times (\boldsymbol{I}\vec{\omega}) = -\boldsymbol{\eta}\vec{\omega} + \sqrt{2k_BT}\boldsymbol{\eta}^{1/2}\vec{\xi}(t), \tag{3.15}$$

which is the inertial Langevin equation for rotational degrees of freedom.

In the case where the moment of inertia $\boldsymbol{I}$ is small or in the case of large friction, the inertial terms can be neglected, i.e. the r.h.s. of equation (3.15) is zero. In this case the approximate equation is called overdamped Langevin equation and has the following form:

$$\vec{\omega} = \sqrt{2k_BT}\boldsymbol{\eta}^{-1/2}\vec{\xi}(t), \tag{3.16}$$

where $\vec{\omega}$ is the angular velocity.

Fixing on the body a unit vector $\vec{n}$, to express the orientation of a particle, assumed spherical for simplicity, the following relation holds:

$$\frac{d\vec{n}}{dt} = \vec{\omega} \times \vec{n}\,. \tag{3.17}$$

This relation, equation (3.17), is purely kinematic relation that holds for any unit vector $\vec{n}$ fixed in the body of a particle and has nothing to do with whether or not the particle is undergoing Brownian motion [21]. Therefore the overdamped Langevin equation in (3.16) can be written as:

$$\frac{d\vec{n}}{dt} = \sqrt{2k_BT}\boldsymbol{\eta}^{-1/2}\vec{\xi}(t) \times \vec{n}, \tag{3.18}$$

so, we have a Langevin equation in $\vec{n}$ alone, not involving $\vec{\omega}$. The same result can be obtained performing the formal limit for small inertia using the multiscale analysis (see Ch. 5), as computed in Paper I for an ellipsoidal Brownian particle.

The Fokker-Planck associated to the process has the same form of equation (3.11) with a rotational diffusion tensor equal to $\boldsymbol{D}^{(\mathrm{rot})} = k_B T \boldsymbol{\eta}^{-1}$. The rotational diffusion tensor has dimension of $[\mathrm{time}]^{-1}$, and is not the same as that of the translational diffusion tensor, which is $[\mathrm{length}]^2[\mathrm{time}]^{-1}$ [55]. In the next section, i.e. Sec. 3.3, we will describe the Einstein's relation for the rotational Brownian motion.

Rotational diffusion of an object can be characterized by the time it takes for the object to forget its initial orientation. This time is known as the rotational relaxation time, and is related to the rotational friction tensor. By dimension of the rotational diffusion coefficient, one can infer that the order of magnitude of rotational relaxation time is inversely proportional to rotational diffusion coefficient. Characteristic rotational relaxation times for small molecules are very fast, from picoseconds to nanoseconds, while for large macromolecules the order of magnitude of milliseconds or longer. The rotational diffusion tensor may be determined experimentally through methods sensitive to nanoseconds, e.g. fluorescence anisotropy [56–58]

3.3 Rotational diffusion tensor

In this section, we want to describe the Einstein's relation for the rotational Brownian motion, i.e. the connection between the rotational diffusion tensor and the rotational friction tensor.

The discussion about the current density in Sec. 3.1 has been treated to explain how the rotational diffusion tensor, defined by equation (3.6), is connected with the rotational friction tensor. As Favro in [36], and Valiev and Ivanov in [39], we follow the Einstein's method [33] to establish this connection.

As first step, we assume that a weak external torque given by a scalar potential $\psi(\vec{\Omega})$ acts on the rigid body and causes the deterministic changing of the orientation. The fundamental equation of motion for the total angular momentum of a rigid body in body coordinate system, given in (3.3), therefore is:

$$\left(\boldsymbol{I} \frac{d\vec{\omega}}{dt} \right) + \vec{\omega} \times (\boldsymbol{I}\vec{\omega}) = i\vec{\mathcal{J}}\psi(\vec{\Omega}) - \boldsymbol{\eta} \cdot \vec{\omega}, \tag{3.19}$$

where the r.h.s. is composed by the torque on a typical body $i\vec{\mathcal{J}}\psi(\vec{\Omega})^\dagger$, given by the scalar potential $\psi(\vec{\Omega})$ multiplied the differential operator $\vec{\mathcal{J}}$, and the viscous drag on the body, given by $-\boldsymbol{\eta}\cdot\vec{\omega}$, where $\boldsymbol{\eta}$ is the rotational friction tensor and $\vec{\omega}$ the angular velocity. A stationary distribution of orientation will be given by the Boltzmann law $P^{(\mathrm{stat})}(\vec{\Omega}) = c\exp(-\psi(\vec{\Omega})/k_B T)$, where c is the integration constant.

Following Favro [36], and Valiev and Ivanov [39], we make several simplifying assumptions on the equation (3.19) to compute an easily solution of the equation (3.19).

[†] We recall that the inverse of the imaginary unit i is in the operator $\vec{\mathcal{J}}$.

Firstly, we assume that the nonlinear term in the l.h.s, $\vec{\omega}\times(\boldsymbol{I}\vec{\omega})$ ($\boldsymbol{I}$ is the moment of inertia), may be neglected because the angular velocity $\vec{\omega}$ is assumed to be small enough. Secondly, ψ varies with respect to $\vec{\Omega}$ slowly, such that the torque may be treated as constant over a small interval [36].

Under these two simple assumptions the solution of equation (3.19) becomes:

$$\vec{\omega} = \vec{\omega}_0 \mathrm{e}^{(-\boldsymbol{I}^{-1}\boldsymbol{\eta}t)} + i\left[1 - \mathrm{e}^{(-\boldsymbol{I}^{-1}\boldsymbol{\eta}t)}\right]\boldsymbol{\eta}^{-1}\vec{\mathcal{J}}\psi(\vec{\Omega}). \tag{3.20}$$

We are interested in effects that take place in a time interval much bigger than the typical time scale of the process, i.e. $t >> \boldsymbol{I}\boldsymbol{\eta}^{-1}$, such that the angular velocity of the rigid body can be assumed equal to $\vec{\omega}_\infty = i\boldsymbol{\eta}^{-1}\vec{\mathcal{J}}\psi(\vec{\Omega})$. The rate of change of the vector $\vec{\Omega}$ [36, 39, 41] is then given by the following relation:

$$d\vec{\Omega} = -i\vec{\omega}_\infty dt \cdot \vec{\mathcal{J}}\vec{\Omega}. \tag{3.21}$$

The current density caused by the presence of the scalar potential $\psi(\vec{\Omega})$, according to Valiev and Ivanov [39], is given by $\vec{J}^{(\psi)} = W(\vec{\Omega})\vec{\omega}_\infty\cdot\vec{\mathcal{J}}\vec{\Omega} = \Gamma^{-1}P(\vec{\Omega})\vec{\omega}_\infty\cdot\vec{\mathcal{J}}\vec{\Omega}$. Therefore the total current, for a rotational Brownian rigid body subjected to an external potential, can be described by the sum of two terms, $\vec{J}^{(\psi)}$ and $\vec{J}$, where $\vec{J}$ is given by the equation in (3.14), i.e. in Einstein's summation convention :

$$J_l^{(\text{total})} = J_l^{(\psi)} + J_l = \Gamma^{-1}(\mathcal{J}_j\Omega_l)(D_{jk}^{(\text{rot})}\mathcal{J}_k P(\vec{\Omega},t) + P(\vec{\Omega},t)\eta_{jk}^{-1}\mathcal{J}_k\psi(\vec{\Omega})). \tag{3.22}$$

Under the assumption of stationary conditions, the probability distribution $P(\vec{\Omega},t)$ must be substituted by $P^{(\text{stat})}(\vec{\Omega}) = c\exp(-\psi(\vec{\Omega})/k_B T)$ and the equation (3.22) written above becomes:

$$J_{(\text{stat})l}^{(\text{total})} = \Gamma^{-1}(\mathcal{J}_j\Omega_l)(D_{jk}^{(\text{rot})} - k_B T\eta_{jk}^{-1})\mathcal{J}_k P^{(\text{stat})}(\vec{\Omega}). \tag{3.23}$$

By equation (3.13) and by stationary conditions, i.e. [39]:

$$0 = \partial_l J_{(\text{stat})l}^{(\text{total})} = \mathcal{J}_j(D_{jk}^{(\text{rot})} - k_B T\eta_{jk}^{-1})\mathcal{J}_k P^{(\text{stat})}(\vec{\Omega}), \tag{3.24}$$

we finally obtain the desired relationship between the rotational diffusion tensor $\boldsymbol{D}^{(\text{rot})}$ and the rotational friction tensor $\boldsymbol{\eta}$:

$$\boldsymbol{D}^{(\text{rot})} = k_B T\boldsymbol{\eta}^{-1}. \tag{3.25}$$

The calculation of the rotational friction tensor $\boldsymbol{\eta}$ can be performed analytically for spherical and ellipsoidal particles. The rotational friction tensor for a spherical particle obeys to Stokes' law, i.e. $\boldsymbol{\eta} = 8\pi\nu a^3\mathbf{I}$ where ν is the viscosity of the medium where the Brownian particle is suspended, a is the radius of the particle and $\mathbf{I}$ the identity matrix, while the general calculation of rotational friction tensor for an ellipsoidal particle was performed in 1892 by D. Edwards in [59]. The results of Edwards has been used explicitly in Paper I.

3.4 Rotational diffusion in physics, chemistry and biology

Rotational diffusion is an important feature in the study of many physical, chemical and biological topics. In this section we recall some important result obtained in the last decades for these topics.

In physics, using digital video microscopy [42], fluorescence anisotropy [56–58, 60–62], or other techniques [63, 64], one is able to compute, for example, the translational and rotational diffusion coefficients of particles and thereby determine their form. As a simple tutorial example, we discuss how to compute the effective radius of a virus with a spherical shape. Let's assume that an experiment has studied the time-dependence of interaction with light of a spherical virus immersed in a water solution at 300 K. We also assume that the orientation of the virus can be distinguished by its electrical properties. The result of the experiment is that the rotational diffusion time is measured to be $\tau_r = 0.18$ ms. The Stokes' law and the connection between the rotational diffusion coefficient and the characteristic rotational relaxation time allows us to compute the effective radius of a spherical virus. Indeed, the radius $r = (3k_B T\tau_r/16\pi\nu)^{1/3}$, where ν is the viscosity of the water, is equal to 35 nm.

Although it is useful to apply rotational diffusion for extrapolating the shape and size of a particle, another interesting application of rotational diffusion exists. This application consists of studying the equilibrium mechanical shear response of viscoelastic soft materials by the size of the particle [65]. In other words, probe particles having a known shape and size are introduced into a non-Newtonian viscoelastic soft material and by measuring their rotational diffusion one can deduce the viscoelastic shear modulus of that soft material [65].

In chemistry rotational diffusion is used to describe the kinetics of diffusion-controlled reaction between chemically asymmetric molecules [66], while in biology it is used to describe rotational motion of biomacromolecules in membranes, like rhodospsin chromophores [67] and protein in the human erythrocyte membrane [68]. For example, rhodospsin chromophores are light-sensitive receptor proteins involved in visual phototransduction. The orientational and rotational motions of them are directly observed with polarized light in normal physiological condition, and therefore are useful cases for investigating rotational diffusion [67].

Chapter 4

Diffusion on Manifolds

That which does not kill us makes us stronger.

–F. Nietzsche

In this chapter we introduce the general theory of Brownian motion on manifolds. This theory was developed principally by A. N. Kolmogorov [52, 53, 69–71] around 1930's. Successively, K. Itô [72] in 1950 , K. Yosida [73] in 1952 and H. P. Jr. McKean [74] in 1960, developed theories for Brownian motion on Lie Groups, which are described also in many monographs [75–78]. We start, therefore, with recalling the Kolmogorov's theory in Sec. 4.1. In this section we discuss and develop the forward Fokker-Planck equation on manifolds. In Sec. 4.2 we discuss Itô [72], Yosida [73,79] and McKean [74] works, and we end up describing the stochastic evolution over the Lie groups writing down a Langevin equation valid on Lie algebras and "injected" on the group by the McKean-Gangolli's Injection method.

4.1 Fokker-Planck equation on manifolds

Sometimes physical systems are subjected to holonomic constraint. For example, the motion of a particle constrained to lie on a curved surface is subject to a holonomic constraint (see Fig.1). The dynamics of these systems cannot be described by orthogonal curvilinear coordinates. Consequently, these system must be analyzed through the general framework of analytical mechanics invoking the introduction of generalized coordinate [80].

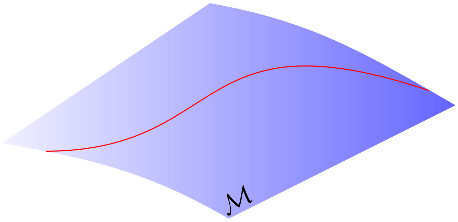

Figure 4.1: The figure displays the motion of a particle constrained to lie on a curved surface subject to some external force. The curved surface, i.e. the smooth manifold with metric tensor $g_{\mu\nu}$ plotted in blue, is the holonomic constraint. The trajectory performed by the particle is plotted in red.

At mesoscopic scales, a rigid body suspended in a fluid displays a continuous and irregular motion given by the incessant collisions with the surrounding molecules. If this irregular motion occurs on a curved space, how can this Brownian motion be described? In other words, which modifications does a Fokker-Planck need to describe a motion of a particle subjected to holonomic constraint?

For answering these questions we develop, in this section, the forward Fokker-Planck equation on a closed n-dimensional Riemannian manifold $\mathcal{M}$ following Kolmogorov's works [52, 53, 69–71].

A closed n-dimensional Riemannian manifold is a smooth compact manifold of dimensions n equipped with an inner product on the tangent space and without boundary [81]. In this section the Einstein's summation convention is used.

The admissible states $\vec{x}$ of a physical system $\mathcal{G}$ with n degrees of freedom form a Riemannian manifold $\mathcal{M}$ of dimensions n [70]. We assume that the Riemannian manifold $\mathcal{M}$ is closed. The process of variation of the physical system $\mathcal{G}$ is said, using the words of Kolmogorov [70], to be "*stochastically determined*" if "*under an arbitrary choice of $\vec{x}$, the region $\mathcal{U} \subset \mathcal{M}$ and time s and t, with $s < t$, the probability $P(s, \vec{x}, t, \mathcal{U}) = \int_{\mathcal{U}} f(s, \vec{x}, t, \vec{y}) dV_{\vec{y}}$ that the system in state $\vec{x}$ at time s will be in one of the state of the $\mathcal{U}$ at time t is defined*". The function $f(s, \vec{x}, t, \vec{y})$ has to satisfy, therefore, the following requirements [70]:

$$\int_{\mathcal{M}} f(s, \vec{x}, t, \vec{y}) dV_{\vec{y}} = 1; \tag{4.1a}$$

$$f(s, \vec{x}, t, \vec{y}) = \int_{\mathcal{M}} f(s, \vec{x}, t', \vec{z}) f(t', \vec{z}, t, \vec{y}) dV_{\vec{z}} \qquad s < t' < t, \tag{4.1b}$$

where the invariant volume element has been denoted with $dV_{\vec{y}}$.

Let's suppose now to take a point $\vec{y}_0$ on the manifold $\mathcal{M}$ and let's call the neighborhood of this point U. We choose a coordinate system $\vec{z} = (z_1, ..., z_n)$ on the neighborhood U in $\mathcal{M}$ [70], i.e. $U \in \mathcal{M}$. Moreover, we assume that in the open

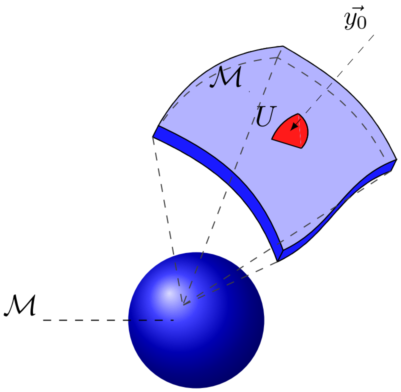

Figure 4.2: Schematic picture of a portion of a closed Riemannian manifold $\mathcal{M}$. The closed manifold $\mathcal{M}$ is identified by the shaded dark blue sphere. The portion of the closed Riemannian manifold is plotted in light blue, where in red is plotted neighborhood U of point $\vec{y}_0$. The point $\vec{y}_0 \in U$ is indicated by the dashed arrow.

set U of the point $\vec{y}_0$, for a given t, the following limits exist uniformly:

$$A_i(t,\vec{y}) = \lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_U f(t,\vec{z},t+\Delta t,\vec{y})(z_i-y_i)dV_{\vec{z}} = \lim_{\Delta t\to 0}\frac{a_i(t,\vec{y},\Delta t)}{\Delta t}; \tag{4.2a}$$

$$B_{ij}(t,\vec{y}) = \lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_U f(t,\vec{z},t+\Delta t,\vec{y})(z_i-y_i)(z_j-y_j)dV_{\vec{z}} = \lim_{\Delta t\to 0}\frac{b_{ij}(t,\vec{y},\Delta t)}{\Delta t}; \tag{4.2b}$$

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_U f(t,\vec{z},t+\Delta t,\vec{y})\rho^3(\vec{y},\vec{z})dV_{\vec{z}} = \lim_{\Delta t\to 0}\frac{\nu(t,\vec{y},\Delta t)}{\Delta t} = 0, \tag{4.2c}$$

where the geodesic distance between $\vec{y}$ and $\vec{z}$ has been denoted with $\rho(\vec{y},\vec{z})$ [69,70].

Equation (4.2a) can be physically interpreted as a general velocity field (also called drift term) of the physical system $\mathcal{G}$, while (4.2b) a general diffusion tensor of the process with the matrix $B_{ij}(t,\vec{y})$ symmetric and positive defined. We also assume the property that the drift term $A_i(t,\vec{y})$ and the diffusion tensor $B_{ij}(t,\vec{y})$ are twice continuously differentiable in the open set U [70].

Let's now define an arbitrary non-negative function $Q(\vec{y})$ of the class C^3 on the set U, such that $Q(\vec{y})$ is zero outside U. For two general point in the set U, i.e. $\vec{y}$, $\vec{z} \in U$, the function $Q(\vec{y})$ can be thus expressed as:

$$Q(\vec{y}) = Q(\vec{z}) + (y_i-z_i)\frac{\partial Q(\vec{z})}{\partial z_i} + \frac{1}{2}(y_i-z_i)(y_j-z_j)\frac{\partial^2 Q(\vec{z})}{\partial z_i \partial z_j} + \theta'\rho^3(\vec{y},\vec{z}), \quad |\theta'| \le C', \tag{4.3}$$

where C' is a constant independent of $\vec{z}$.

If the point $\vec{y} \in \mathcal{M} - U$ and the point $\vec{z} \in U$ then the function

$$Q(\vec{y}) = Q(z) + \theta'' \rho^3(\vec{y}, \vec{z}) \quad |\theta''| \le C'', \tag{4.4}$$

where C'' is a constant independent of $\vec{z}$. Finally if $\vec{y}, \vec{z} \in \mathcal{M} - U$ then

$$Q(\vec{y}) = 0. \tag{4.5}$$

Under these assumptions we observe that [70]:

$$\int_U Q(\vec{y}) \frac{\partial f(s, \vec{x}, t, \vec{y})}{\partial t} dV_{\vec{y}} = \frac{\partial}{\partial t} \int_{\mathcal{M}} Q(\vec{y}) f(s, \vec{x}, t, \vec{y}) dV_{\vec{y}}. \tag{4.6}$$

Applying the definition of derivative and the property described in equation (4.1b), we obtain:

$$\begin{aligned} \frac{\partial}{\partial t} \int_{\mathcal{M}} Q(\vec{y}) f(s, \vec{x}, t, \vec{y}) dV_{\vec{y}} &= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{\mathcal{M}} f(s, \vec{x}, t, \vec{z}) \\ \int_{\mathcal{M}} Q(\vec{y}) f(s, \vec{z}, t + \Delta t, \vec{y}) dV_{\vec{y}} dV_{\vec{z}} &- \int_{\mathcal{M}} Q(\vec{z}) f(s, \vec{x}, t + \Delta t, \vec{z}) dV_{\vec{z}}. \end{aligned} \tag{4.7}$$

Using the results for $Q(\vec{y})$ in (4.3), (4.4), (4.5) we can easily find:

$$\begin{aligned} \int_U Q(\vec{y}) \frac{\partial}{\partial t} f(s, \vec{x}, t, \vec{y}) dV_{\vec{y}} &= \int_U f(s, \vec{x}, t, \vec{y}) \\ \left[A_i(t, \vec{y}) \frac{\partial}{\partial y_i} Q(\vec{y}) + B_{ij}(t, \vec{y}) \frac{\partial^2}{\partial y_i \partial y_j} Q(\vec{y}) \right] dV_{\vec{y}}. \end{aligned} \tag{4.8}$$

The volume element $dV_{\vec{y}}$ is invariant. Following now the formalism of S. Weinberg [82], to be consistent also with the monographs [77, 78], we claim that the volume element can be written as $dV_{\vec{y}} = \sqrt{g(\vec{y})} d^n y$, where $g(\vec{y})$ is the determinant of the metric tensor $g_{\mu\nu}(\vec{y})$ of the Riemannian manifold $\mathcal{M}$ [82]. Indeed, because $\sqrt{g(\vec{y})}$ transforms as a scalar density is easily to check that $dV_{\vec{y}} = \sqrt{g(\vec{y})} d^n y$ is invariant under general coordinate transformations. Recall that the function $Q(\vec{y})$ is arbitrary, after one integration by part for the integral containing the drift term $A_i(t, \vec{y})$ and after two integration by part of the integral containing the diffusion tensor $B_{ij}(t, \vec{y})$ we obtain finally the *second differential equation*, using the words of Kolmogorov in [69, 70], at interior points of U:

$$\frac{\partial}{\partial t} f(s, \vec{x}, t, \vec{y}) = - \frac{1}{\sqrt{g(\vec{y})}} \left[\frac{\partial}{\partial y_i} \sqrt{g(\vec{y})} A_i(t, \vec{y}) - \frac{\partial^2}{\partial y_i \partial y_j} \sqrt{g(\vec{y})} B_{ij}(t, \vec{y}) \right] f(s, \vec{x}, t, \vec{y}). \tag{4.9}$$

It is very useful to make the following observation. The function $f(s, \vec{x}, t, \vec{y})$ is a scalar function and transforms as a scalar. Indeed, it is easy to show that, under general coordinate transformations, $f(s, \vec{x}, t, \vec{y}) dV_{\vec{y}} = f(s, \vec{x}, t, \vec{y}') dV_{\vec{y}'}$ because $dV_{\vec{y}}$

is the invariant volume element. If at time t_0 is also given a non negative function $h(t_0, \vec{y})$, such that:

$$\int_{\mathcal{M}} h(t_0, \vec{y}) dV_{\vec{y}} = 1; \tag{4.10a}$$

$$h(t, \vec{y}) = \int_{\mathcal{M}} h(t_0, \vec{x}) f(t_0, \vec{x}, t, \vec{y}) dV_{\vec{x}} \quad \text{with} \quad t > t_0. \tag{4.10b}$$

This simple observation, allows us to define the quantity $p(t, \vec{y}) = \sqrt{g(\vec{y})} h(t, \vec{y})$, which by construction transforms as a scalar density. This density distribution function satisfies the Fokker-Planck equation valid on manifold $\mathcal{M}$ [52, 53]:

$$\frac{\partial}{\partial t} p(t, \vec{y}) = -\left[\frac{\partial}{\partial y_i} (A_i(t, \vec{y}) p(t, \vec{y})) - \frac{\partial^2}{\partial y_i \partial y_j} (B_{ij}(t, \vec{y}) p(t, \vec{y})) \right]. \tag{4.11}$$

Equations (4.9) and (4.11) were used in Paper III, respectively in equation (3) and equation (10).

Let's discuss in detail equation (4.9). The l.h.s. of equation (4.9) is composed by a scalar function $f(s, \vec{x}, t, \vec{y})$ and the time derivative. Our setting is the mesoscopic scale and we can assume that the time is non-relativistic, i.e. the time is the same in any reference frame. For these reasons, the l.h.s of equation (4.9) is a scalar because transforms as a scalar. The r.h.s., thus, has to follow the same nature of transformation.

We observe that equation (4.9) can be rewritten in the following way:

$$\frac{\partial}{\partial t} f(s, \vec{x}, t, \vec{y}) = -\frac{1}{\sqrt{g(\vec{y})}} \frac{\partial}{\partial y^i} \sqrt{g(\vec{y})} J^i(s, \vec{x}, t, \vec{y}), \tag{4.12}$$

where J^i shall be a contravariant vector current defined as

$$J^i(s, \vec{x}, t, \vec{y}) = \left[A^i(t, \vec{y}) - \frac{1}{\sqrt{g(\vec{y})}} \frac{\partial}{\partial y^j} \sqrt{g(\vec{y})} B^{ij}(t, \vec{y}) \right] f(s, \vec{x}, t, \vec{y}). \tag{4.13}$$

This current vector has to transform as a contravariant vector, because if $J^i(s, \vec{x}, t, \vec{y})$ is a contravariant vector the r.h.s. of equation (4.12) is the covariant divergence. Because the covariant divergence transforms as a scalar [82], we obtain that both sides of equation (4.9) transform as a scalar and are therefore invariant under general coordinate transformations. The indices of the equation (4.12) have been raised following the convention of Weinberg [82].

Let's define with the symbol $V^{\nu}_{;\nu}$ the covariant divergence of a contravariant vector V^{ν} [82] as:

$$V^{\nu}_{;\nu} = \frac{\partial V^{\nu}}{\partial y^{\nu}} + \Gamma^{\nu}_{\nu\lambda} V^{\lambda} = \frac{1}{\sqrt{g(\vec{y})}} \frac{\partial}{\partial y^{\nu}} \sqrt{g(\vec{y})} V^{\nu}, \tag{4.14}$$

where $\Gamma^{\nu}_{\mu\lambda}$ is the affine connection and $\Gamma^{\nu}_{\nu\lambda} = \frac{1}{2}\frac{\partial}{\partial y^{\lambda}}\ln g = \frac{1}{\sqrt{g(\vec{y})}}\frac{\partial}{\partial y^{\lambda}}\sqrt{g(\vec{y})}$. We can also use the previous formula about the contraction of two indices in the affine connection to simplify the formula for the covariant divergence of a tensor $T^{\mu\nu}_{;\mu}$, for example:

$$T^{\mu\nu}_{;\mu} = \frac{\partial T^{\mu\nu}}{\partial y^{\mu}} + \Gamma^{\mu}_{\mu\lambda}T^{\lambda\nu} + \Gamma^{\nu}_{\mu\lambda}T^{\mu\lambda} = \frac{1}{\sqrt{g(\vec{y})}}\frac{\partial}{\partial y^{\mu}}(\sqrt{g(\vec{y})}T^{\mu\nu}) + \Gamma^{\nu}_{\mu\lambda}T^{\mu\lambda}. \tag{4.15}$$

Because we want the equation (4.9) to be invariant under general coordinate transformations, we assume that $\boldsymbol{B}$ transforms as a tensor of rank two, and therefore we may write:

$$\frac{1}{\sqrt{g(\vec{y})}}\frac{\partial}{\partial y^{\mu}}(\sqrt{g(\vec{y})}B^{\mu\nu}) = B^{\mu\nu}_{;\mu} - \Gamma^{\nu}_{\mu\lambda}B^{\mu\lambda}. \tag{4.16}$$

For obtaining an invariant equation, therefore, the drift term may not transform as a tensor of rank one, but must transform in a different way. Indeed, the law of transformation for $A^{i}(t,\vec{y})$, to make equation (4.9) invariant under general coordinate transformations, is:

$$A^{i'}(t,\vec{y}) = \frac{\partial y^{i'}}{\partial y^{k}}A^{k}(t,\vec{y}) + \frac{\partial^2 y^{i'}}{\partial y^{k}\partial y^{l}}B^{kl}(t,\vec{y}), \tag{4.17}$$

which is the usual transformation law for the the drift term. Therefore plugging in equation (4.16) into (4.13), and using the formalism of covariant derivative, we obtain:

$$\frac{\partial}{\partial t}f(s,\vec{x},t,\vec{y}) = -J^{\mu}_{;\mu}(s,\vec{x},t,\vec{y})\,, \tag{4.18a}$$

$$J^{\mu}(s,\vec{x},t,\vec{y}) = \left[KH^{\mu} - B^{\nu\mu}_{;\nu}\right]f(s,\vec{x},t,\vec{y})\,, \tag{4.18b}$$

where

$$KH^{\mu} = A^{\mu}(t,\vec{y}) + \Gamma^{\mu}_{\nu\lambda}B^{\nu\lambda}(t,\vec{y})\,. \tag{4.18c}$$

The new contravariant vector KH^{μ} transforms under general coordinate transformations as a contravariant vector, and equation (4.18a) in each interior point U of $\mathcal{M}$ holds and it is invariant. KH^{μ} is often called Hänggi-Klimontovich drift [83–85].

4.2 Stochastic processes on Lie groups

Lie groups have many applications in real world. For example, elastic models of DNA [86], steering of flexible needles [87] and nonholonomic kinematic systems such as wheeled vehicles that are subject to rolling constraints [88–90] are applications that are described by diffusion processes on Lie groups [78]. According to Chirikjian [86], Lie groups can also be made use of like a configuration space over which ensembles of conformations evolve.

In this section we discuss stochastic processes on Lie groups recalling the pioneering works of Itô [72], Yosida [73, 79] and McKean [74]. Let's start with a definition of Lie group. We refer to Appendix A for more details about Lie groups.

Lie groups [91, 92] provide a way to express the concept of a continuous family of symmetries for geometric objects. By differentiating the Lie group action, you get a Lie algebra action, which is a linearization of the group action. As a linear object, a Lie algebra is often a lot easier to work with than working directly with the corresponding Lie group. A formal definition given by Bishop and Crittenden in [93] is: "*a Lie group G is a set which is both a group and a manifold and such that the group operation are C^∞ , that is, the maps:*"

$$G \times G \to G \;\; \textit{given by} \;\; (g, h) \to gh; \tag{4.19a}$$

$$G \to G \;\; \textit{given by} \;\; g \to g^{-1} \;\; \textit{are} \; C^\infty. \tag{4.19b}$$

The whole information in a Lie group is almost contained in its Lie algebra [92].

A Lie algebra, according to Bishop and Crittenden [93], " *is a vector space L for which is given a bilinear function from $L \times L$ to L, called bracket, and denoted by $[,]$, which satisfies:*

- *$[\vec{x}, \vec{x}] = 0$ for every $\vec{x} \in L$, and that $[\vec{x}, \vec{y}] = -[\vec{y}, \vec{x}]$ for every $\vec{x}, \vec{y} \in L$;*
- *the Jacobi identity: for every $\vec{x}, \vec{y}, \vec{z} \in L$ $[\vec{x}, [\vec{y}, \vec{z}]] + [\vec{z}, [\vec{x}, \vec{y}]] + [\vec{y}, [\vec{z}, \vec{x}]] = 0$".*

The Lie algebra is the cornerstone of the Lie group theory. It allows to replace complicated non-linear conditions of invariance by relatively simple linear infinitesimal conditions.

Given a Lie group G, it is possible to define the left invariant vector field of G. A left invariant vector field of G is, by definition given in [93], "*a vector field which is fixed under differentials of left translations, that is, if $L_g : G \to G$ is defined by $L_g(h) = gh$, then X is a left invariant vector field if $dL_g X(h) = X(gh)$ for every $g, h \in G$. Therefore, the Lie algebra of G, denoted by $\boldsymbol{m}$, is the Lie algebra of the left invariant vector field*".

Lie group and its Lie algebra are connected via exponential map. Formally, the definition of the exponential map is [93]: "*Let's $X \in \boldsymbol{m}$ be a point on the algebra. Let's ζ_X be the integral curve of X starting at the identity. Then the exponential map $exp : \boldsymbol{m} \to G$ is the map which assigns ζ_X to X*".

In other words, the exponential map is a map from the Lie algebra $\boldsymbol{m}$ of a Lie group G to the group G and it is able to recapture the local group structure of the Lie algebra on the group, see Fig. 4.3. The Lie algebra $\boldsymbol{m}$, which is the tangent space at the identity element e of the group G, is colored in red in Fig. 4.3. The Lie Group G is colored in blue in Fig. 4.3. In Fig. 4.3 the exponential map is indicated by e^X, which is equal to ζ_X.

In this thesis we are interested in only matrix Lie groups, then the exponential map coincides with the matrix exponential and is given by the ordinary series expansion, i.e. $\exp(X) = \sum_{n=0}^{\infty} X^n/n!$ where X is a square matrix.

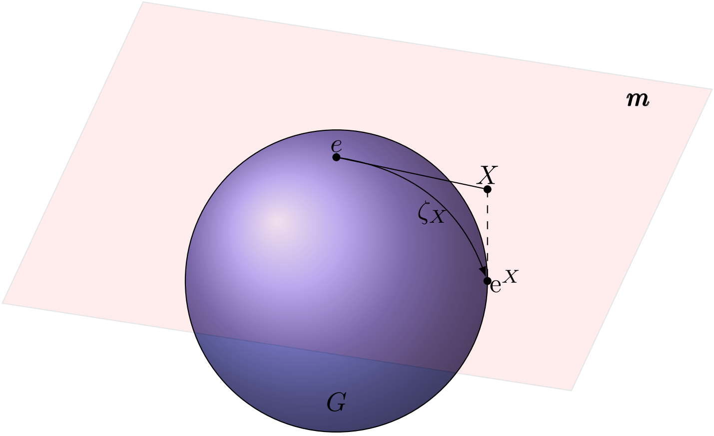

Figure 4.3: The picture displays how the exponential map works. The Lie algebra $\boldsymbol{m}$ of a Lie Group G, colored in blue, is the tangent space at the identity element of the group, indicated with e. In the figure the Lie algebra $\boldsymbol{m}$ is the plane colored in red. Let's $X \in \boldsymbol{m}$ be a point on the algebra. Let's ζ_X be the integral curve of X starting at the identity. Then the exponential map is the map which assigns ζ_X to X. In the picture the exponential map is indicated by e^X, which is equal to ζ_X [93].

The first author to introduce Brownian motion on a Lie group was Yosida in [79]. He analyzed the Brownian motion on the surface of a 3-sphere, studying a homogeneous Markov process on a n-dimensional compact Riemannian space S with a metric given by $ds^2 = g_{ij}(x)dx^i dx^j$ such that the totality G of the isometric transformations of S onto S constitutes a Lie group transitive on S [79]. Using the property given by Kolmogorov [70] and using the infinitesimal operators of the Lie group G, he shows that exists only and only one Brownian motion on the surface of the 3-sphere S^3 [79].

It is useful recall that every simply connected, closed 3-manifold is homeomorphic to the 3-sphere. Indeed, the Poincaré conjecture, the only problem solved of the Seven Millennium Problems [94] proved in 2003 by Grigori Perelman [95–97], provides that the 3-sphere is the only three-dimensional manifold (up to homeomorphism) with these properties. When thought as a Lie group S^3 is often denoted $Sp(1)$ or $U(1, \boldsymbol{H})$. It is a nonabelian, compact Lie group of dimension 3.

Itô in [72] generalized the result of Yosida to all Lie groups. For a n-dimensional Lie group G, Itô [72] shows that a Markovian random process $\pi(t)$ on G is called to be a right (left) invariant Brownian motion on G if it satisfies the Kolmogorov-Feller's continuity condition [70, 98], temporal homogeneity, spatial homogeneity and continuity.

The case of Brownian motion on $SO(3)$ group manifold was explicitly constructed by McKean in 1960 [74,75]. In this paper, the author introduce the method of injection and with it the first stochastic differential equation on the Lie algebra $\boldsymbol{so}(3)$. The McKean-Gangolli injection method "injects" a stochastic process from the Lie algebra into the group using the exponential map and product integral formula. In the following we explain the mathematical formalism of the method.

A stochastic differential equation with constant or time-varying coefficient has the following form:

$$dx_i = a_i(t)dt + b_{ij}(t)dW^j(t). \tag{4.20}$$

This equation describes a stochastic sample trajectories that evolve in a Euclidean space $\mathbb{R}^n$ [78]. The term $a_i(t)$ transforms following the Itô's lemma, while the matrix $\boldsymbol{b}(t)$ transforms as a tensor of rank two under general coordinate transformations. The term $dW^j(t)$ are Wiener processes with zero mean and correlation function equal to $\langle dW^i(t)dW^j(t')\rangle = \delta_{ij}\delta(t-t')$. The fact that the tensor $\boldsymbol{b}(t)$ does not depend by coordinates $\vec{x}$, the equation (4.20) can be interpreted as an Itô or a Stratonovich equation (see Appendix B).

Identifying this Euclidean space $\mathbb{R}^n$ with an n-dimensional Lie algebra $\boldsymbol{m}$, associated to a generic Lie group G, then through the exponential map, it is possible to transfer, or "inject" with the words of McKean and Gangolli, this sample path into the corresponding Lie group [78].

Each time that we use the exponential map, the path obtained brings us away from the identity e of the group G. Since a Lie algebra is defined as the tangent space at the identity, McKean and Gangolli [74,75,78,99,100] use the idea to reset things at the identity at each step. To explain this resetting it is useful to introduce an instructive example given by Gregory S. Chirikjian [78]. Let's $g \in G$ an element of the Lie group G. At time $t = 0$ we assume that the element $g(t)$ is at the identity $g(0) = e$. At time $t = dt$, the evolution on the the Lie algebra $\boldsymbol{m}$ can be exponentiated on the Lie Group using the exponential map, and therefore on the Lie group the element is $g(dt) = g(0) \circ \mathrm{e}^{(dx_i J^i)}$, where the symbol $\circ$ identifies the composition of two elements of the group, i.e. the multiplication, dx_i is evaluated at time dt, and J^i are the generator of the Lie algebra, i.e. a basis for G (see Appendix A). At the next step $2dt$, $g(2dt) = g(dt) \circ \mathrm{e}^{(dx_i J^i)}$, where dx_i is evaluated at time $2dt$ and so on [78]. Finally, we can compute the general path $g(t + dt) \in G$ defined recursively as:

$$g(t+dt) = g(t) \circ \mathrm{e}^{(dx_i J^i)}, \text{ where } g(0) = e, \tag{4.21}$$

where $g(t)$ is generated by the product integral:

$$g(t) = \bigcap_{0\leq\tau\leq t} \mathrm{e}^{(dx_i J^i)}. \tag{4.22}$$

Equation (4.22) is an infinite product of exponential maps. It results from (4.21) with infinitesimal dt [78].

This method has been used in Paper III, where the stochastic evolution of the elements of Lie groups $SO(3)$ and $SU(2)$ has played a central role for describing the Brownian rotations of a Brownian rigid body subject to an external force field in the $\mathbb{R}^3$ Euclidean space and $\mathbb{C}^2$ complex space, respectively.

Chapter 5

Homogenization and Multiscale Methods

Science is organized knowledge.
Wisdom is organized life.

–I. Kant

In this chapter we introduce the homogenization theory and the multiscale methods. We do a simple introduction of these methods and we apply them to a simple advection-diffusion process. We show that effective equations arise from the solvability conditions.

The chapter is divided in four sections. In Sec. 5.1, following a pioneering contribution published in the mid-1990s by M. Vergassola and M. Avellaneda [101], we introduce the homogenization theory and the multiscale analysis. We show how to define quantitatively the different time and length scales in an advection-diffusion process and how to write the hierarchy of equations for the perturbative analysis. In Sec. 5.2 we explain how solvability conditions work by using Fredholm alternative, and in Sec. 5.3 we show how effective equations arise from solvability conditions. We conclude this chapter introducing the connection with calculations performed in the published papers in Sec. 5.4.

5.1 Homogenization and multiscale methods

Many problems in physics, according to Mazzino et al. [102], show the separation of different length and time scales between the degrees of freedom from which they are described. Simulations with modern supercomputers are not able to describe all the relevant scales involved in such difficult problems. A suitable technique, for the treatment of such difficult problems with length and/or time scales separation, is the Multiscale Methods. [102, 103].

The *multiscale analysis* is an asymptotic method for solving problems which have important features at multiple scales of time and/or scale. In mathematics and physics, multiscale analysis consist of techniques used to construct uniformly valid approximation to solution of perturbation problems, both for small as well as large value of independent variables. In a regular perturbation expansion some non-uniformity can appear as a result of resonant interactions between consecutive orders of perturbation theory. This method, identified small and large scale of the model under study, eliminates these non-uniformities and builds a solution as an asymptotic sum [104].

Our aim is to show how the multiscale method works when applied to the analysis of a general problem containing fast and slow degrees of freedom. We want to describe how effective equations, where the fast degrees of freedom are averaged out, arise from the analysis. The procedure to average out such degrees of freedom is called homogenization. Homogenization theory is concerned with the derivation of equations for averages of solutions of equations with rapidly varying coefficients. This procedure allows us to write effective equations for systems with different temporal and spatial scales using perturbation theory. A first order perturbation theory is called averaging, while a second order perturbation theory is called homogenization [105].

To show how homogenization and multiple analysis works, we apply the multiscale techniques to an advection-diffusion process with scale separation, following the work of M. Vergassola and M. Avellaneda [101].

The following Langevin equation describes the motion of particles advected by a velocity field $\vec{v}(\vec{x}, t)$ and subjected to molecular diffusion:

$$\frac{d\vec{x}(t)}{dt} = \vec{v}(\vec{x}, t) + \vec{W}(t). \tag{5.1}$$

With function $\vec{x}(t)$ has been denoted the position at time t of the particle which was initially in $\vec{x}(0) = \vec{x}_0$ [101]. The random movement of the particle is modeled by random process $\vec{W}(t)$, with zero mean and white Gaussian noise in time $\langle W_i(t) W_j(t') \rangle = 2D\delta_{ij}\delta(t - t')$, where the molecular diffusivity is denoted by $\boldsymbol{D} = D\mathbf{I}$, symmetric and definite positive tensor. With the symbol $\mathbf{I}$ is identified the identity matrix.

The velocity field $\vec{v}(\vec{x}, t)$, as pointed out by Vergassola et al. [101], must belong to one of the following classes:

- the first class contains fields which are deterministic and periodic in space and time, or fields which are periodic in space and time independent. The period can be not the same in the various directions;
- the second class contains fields which are homogeneous and time independent, or random, homogeneous and stationary fields.

The symbol $\langle \cdot \rangle$ defines the average over periodicities and the ensemble average in the random case. It is also assumed, here, that the velocity field mean value is equal to zero [101].

As described in Ch. 2, the Fokker-Planck equation associated to the process described by (5.1) is:

$$\frac{\partial p(\vec{x},t)}{\partial t} + \vec{\nabla} \cdot (\vec{v}(\vec{x},t)p(\vec{x},t)) = D\vec{\nabla} \cdot \vec{\nabla} p(\vec{x},t). \tag{5.2}$$

where the probability density function that the particle has a certain position $\vec{x}$ at time t is $p(\vec{x},t)$ [101].

We want to apply homogenization procedure to (5.2) to build an effective diffusion equation at large scales. Specifically, as in [101], we are looking at the dynamics of the density field $p(\vec{x},t)$ on large scales. These scales are larger than the periodicities of $\vec{v}(\vec{x},t)$. To do that we start to identify the fast and slow degrees of freedom of the process.

Time and length scales

The strategy, that here we use, is to describe each variable occurring in equation (5.2) as a dimensionless variable multiplying a dimensionful term (see Ch. 1). In this way, it will be easier to separate the different scale of the process.

We start by assuming that the equation (5.2) describes the evolution of a Brownian particle on a large scale in space, i.e. the variable $\vec{x} = L\hat{\vec{x}}$ where L is the dimensionful term of the variable, while $\hat{\vec{x}}$ is the associated dimensionless variable. The term L identifies size of the long length scale and it can be identified as the ℓ_{slow} (see Ch. 1 Sec. 1.2). This long length scale is the length scale which we are interested in.

The molecular diffusivity tensor, using the same strategy, can be rewritten as $\boldsymbol{D} = D_0\hat{\boldsymbol{D}}$. D_0 by order of magnitude equals $\frac{k_B T}{6\pi\eta a}$ is given by Stokes' law and Einstein's relation, η the viscosity of the fluid and a is the radius of the particle, assumed spherical.

The velocity field can be interpreted as $\vec{v} = v_0\hat{\vec{v}}$, where v_0 has an order of magnitude equal to $\frac{k_B T}{6\pi\eta a l_c}$, where l_c, by order of magnitude $\frac{D_0}{v_0}$, is the periodicity of the velocity field. By assumption $l_c << L$, therefore l_c can be interpreted as the fast length scale of the process, ℓ_{fast}.

At each length scale, l_c and L, we can associate a time scale, respectively τ_c (fast time scale) and τ_L (slow time scale). τ_c refers to the time of the periodicity of velocity field and it has an order of magnitude equal to $\tau_c = \frac{l_c}{v_0}$, while τ_L refers to the time of the effective process that we are looking for. The ratio between $\frac{l_c}{L} << 1$ denotes a small parameter ϵ for the multiscale analysis.

We are able now to define new independent variables for each scale:

$$\tilde{\vec{x}} = \epsilon^{-1}\hat{\vec{x}}\,, \quad \vec{X} = \hat{\vec{x}}\,, \tag{5.3}$$

and derivatives in the new variables become:

$$\frac{\partial}{\partial \hat{x}_i} = \epsilon^{-1}\frac{\partial}{\partial \tilde{x}_i} + \frac{\partial}{\partial X_i}\,, \tag{5.4}$$

The time in the Fokker-Planck equation (5.2) is the time that describes the evolution of a Brownian particle on a large scale in space, i.e. $t = \tau_L \hat{t}$. We, therefore, define the relative time scales, which appear in the process, as:

$$\theta = \epsilon^{-2}\hat{t}\,, \quad \vartheta = \epsilon^{-1}\hat{t}\,, \quad \tau = \epsilon^{0}\hat{t}\,, \tag{5.5}$$

and derivatives become:

$$\frac{\partial}{\partial \hat{t}} = \epsilon^{-2}\frac{\partial}{\partial \theta} + \epsilon^{-1}\frac{\partial}{\partial \vartheta} + \frac{\partial}{\partial \tau} \tag{5.6}$$

We are interested in a perturbation analysis, thus, we require that the probability distribution function p must be dependent by all new independent variables:

$$p = p(\theta, \vartheta, \tau, \tilde{\vec{x}}, \vec{X})\,. \tag{5.7}$$

We, therefore, write down the probability distribution function as an asymptotic sum, with $\epsilon \to 0$, i.e.:

$$p = p^{(0)} + \epsilon p^{(1)} + \epsilon^2 p^{(2)} + \dots\,, \tag{5.8}$$

where the functional dependence in (5.7) on the various variables is inherited to all $p^{(i)}$. The field $p^{(0)}$ is the only normalized to one. All other field $p^{(i)}$ with $i > 0$ may be normalized to zero.

The fast length scale composed by $\tilde{\vec{x}}$ is essentially the fast spatial degree of freedom, with the restriction that p, i.e. equation (5.7), has periodic boundary condition in these variables.

The spatial periodicity is assumed to be l_c. The small length scale $\tilde{\vec{x}} = \epsilon^{-1}\hat{\vec{x}}$, is of order one only for very small $\hat{\vec{x}}$. The variable $\vec{X}$ defines the slow length scale, i.e. the one we are interested in. This variable can be identified as the effective coordinates of the center of the Brownian particle and it describes the effective evolution of the particle's center, where the fast degrees of freedom are integrated away.

The fast time scale is identified with the variable θ. Such variable is essentially the characteristic time of the process. The time ϑ can be identified as an intermediate time scale, at which the motion of the process is governed by a ballistic motion. τ, instead, identifies the slow time where the particle performs an effective motion.

Hierarchy of equations

After all these formal descriptions of small and large scales variables, we can plug (5.4), (5.6) and (5.8) into (5.2) and, collecting terms of equal power in ϵ in the resulting expression, we obtain the following hierarchy of inhomogeneous equations (order ϵ^{-2}, ϵ^{-1} and ϵ^{0}):

$$\mathcal{O}(\epsilon^{-2}): \quad \frac{\partial p^{(0)}}{\partial \theta} + \frac{\partial}{\partial \vec{\tilde{x}}} \cdot (\vec{\hat{v}} p^{(0)}) - \hat{D} \frac{\partial}{\partial \vec{\tilde{x}}} \cdot \frac{\partial}{\partial \vec{\tilde{x}}} p^{(0)} = 0, \tag{5.9a}$$

$$\begin{aligned} \mathcal{O}(\epsilon^{-1}): \quad \frac{\partial p^{(1)}}{\partial \theta} + \frac{\partial}{\partial \vec{\tilde{x}}} \cdot (\vec{\hat{v}} p^{(1)}) - \hat{D} \frac{\partial}{\partial \vec{\tilde{x}}} \cdot \frac{\partial}{\partial \vec{\tilde{x}}} p^{(1)} = -\frac{\partial p^{(0)}}{\partial \vartheta} + \\ -\frac{\partial}{\partial \vec{X}} \cdot (\vec{\hat{v}} p^{(0)}) + 2\hat{D} \frac{\partial}{\partial \vec{\tilde{x}}} \cdot \frac{\partial}{\partial \vec{X}} p^{(0)}, \end{aligned} \tag{5.9b}$$

$$\begin{aligned} \mathcal{O}(\epsilon^{0}): \quad \frac{\partial p^{(2)}}{\partial \theta} + \frac{\partial}{\partial \vec{\tilde{x}}} \cdot (\vec{\hat{v}} p^{(2)}) - \hat{D} \frac{\partial}{\partial \vec{\tilde{x}}} \cdot \frac{\partial}{\partial \vec{\tilde{x}}} p^{(2)} = -\frac{\partial p^{(1)}}{\partial \vartheta} + \\ -\frac{\partial}{\partial \vec{X}} \cdot (\vec{\hat{v}} p^{(1)}) + 2\hat{D} \frac{\partial}{\partial \vec{\tilde{x}}} \cdot \frac{\partial}{\partial \vec{X}} p^{(1)} - \frac{\partial p^{(0)}}{\partial \tau} + \hat{D} \frac{\partial}{\partial \vec{X}} \cdot \frac{\partial}{\partial \vec{X}} p^{(0)} \end{aligned} \tag{5.9c}$$

In the variables $\vec{\tilde{x}}$, the density $p^{(0)}$ obeys to periodic boundary condition.

To fix the dependence of the $p^{(i)}$ functions in (5.7) on the slow variables over fast scale degrees of freedom we need to use solvability conditions, from which effective equations for dynamics of large scales arise.

5.2 Solvability condition: the Fredholm alternative

In this section we describe in a simple way the concept of solvability condition.

Solvability condition of a differential equation is, with a trivial definition, the condition that we need to impose for finding a non trivial solution of a non homogeneous differential equation. Mathematically speaking, solvability conditions are described by Fredholm theory. Fast and slow independent variables and solvability conditions are used to remove secular terms appearing into perturbation analysis.

For a sake of simplicity, following an example given by G. Pavliotis and A. Stuart in [105], we consider the equation:

$$\mathcal{A} u(\vec{x}, t) = f(\vec{x}, t), \tag{5.10}$$

where $\mathcal{A}$ is a linear differential operator $\mathcal{A} : H \to H$ with H an Hilbert space with inner product $(\cdot, \cdot)$. The function $u(\vec{x}, t)$ is unknown and lives on the Hilbert space H and the function $f(\vec{x}, t)$ is the known source function on H of the non homogeneous differential equation. Let $\mathcal{A}^\dagger : H \to H$ be the adjoint of $\mathcal{A}$ via the identity:

$$(\mathcal{A}u(\vec{x},t), v(\vec{x},t)) = (u(\vec{x},t), \mathcal{A}^\dagger v(\vec{x},t)) \;\; \forall u(\vec{x},t), v(\vec{x},t) \in H. \tag{5.11}$$

Suppose now that $v(\vec{x},t) \in H$ denote any element of the kernel $\mathbf{ker}(\mathcal{A}^\dagger)$ of $\mathcal{A}^\dagger$, i.e. $\mathbf{ker}(\mathcal{A}^\dagger) = \{v(\vec{x},t) \in H \,|\, \mathcal{A}^\dagger v(\vec{x},t) = 0\}$. From equation (5.10) we deduce that $f(\vec{x},t)$ is in the image of the linear differential operator $\mathcal{A}$, i.e. $f(\vec{x},t) \in \mathbf{Im}(\mathcal{A})$, and substituting into (5.11) we infer that:

$$(f(\vec{x},t), v(\vec{x},t)) = 0 \;\; \forall v(\vec{x},t) \in H. \tag{5.12}$$

A necessary condition for the existence of a non trivial solution for (5.10), therefore, is that the function $f(\vec{x},t)$ has to be orthogonal to the kernel of the operator $\mathcal{A}^\dagger$.

5.3 Effective equations

In this section we solve the equations in (5.9), to build an uniformly accurate approximation of p.

We start with equation in (5.9a). We consider the case of a compressible velocity field, as in [101]. For the case of incompressible velocity field $\frac{\partial}{\partial \vec{\tilde{x}}} \cdot \vec{\tilde{v}} = 0$ we remind to [106] and to Paper II and Paper III.

For a compressible velocity field, the solution of (5.9a) can be computed using linearity and expressed as:

$$p^{(0)} = m(\vec{\tilde{x}}, \theta)\rho^{(0)}(\vec{X}, \vartheta, \tau), \tag{5.13}$$

where $m(\vec{\tilde{x}}, \theta)$ is normalized to 1 over the fast spatial degrees of freedom and $\rho^{(0)}(\vec{X}, \vartheta, \tau)$ is normalized to 1 over slow spatial degrees of freedom [101].

Ballistic transport

Plugging the solution (5.13) in (5.9b) and applying solvability condition (see Sec. 5.2), we obtain the first effective equation. Averaging out the fast spatial degrees of freedom, indeed, brings to have the following equation:

$$\frac{\partial \rho^{(0)}}{\partial \vartheta} + \vec{\mathbf{V}}_b \cdot \frac{\partial \rho^{(0)}}{\partial \vec{X}} = 0. \tag{5.14}$$

Equation (5.14) describes a ballistic transport. The effective velocity, or ballistic velocity, is equal to:

$$\vec{\mathbf{V}}_b = \langle \vec{\tilde{v}}(\vec{\tilde{x}}, \theta) \rangle = \int \mathrm{d}\vec{\tilde{x}} \;\; \vec{\tilde{v}}(\vec{\tilde{x}}, \theta) \;\; m(\vec{\tilde{x}}, \theta). \tag{5.15}$$

The Brownian particle, therefore, in the intermediate time scale, performs a ballistic motion with a constant velocity given by (5.15).

Diffusive transport

For computing the effective diffusion equation to which we are interested in, we need to find a solution for equation order $\mathcal{O}(\epsilon^{-1})$. To seek this solution, we plug (5.14) into (5.9b) and, using the linearity of Fokker-Planck, we obtain:

$$p^{(1)} = \vec{\chi}(\vec{\tilde{x}}, \theta) \cdot \vec{\nabla} \rho^{(0)}(\vec{X}, \vartheta, \tau), \tag{5.16}$$

where $\vec{\chi}$ is a vector field, which satisfies:

$$\frac{\partial \vec{\chi}}{\partial \theta} + \frac{\partial}{\partial \underline{\vec{\tilde{x}}}} \cdot (\underline{\vec{\hat{v}}} \vec{\chi}) - \hat{D} \left(\frac{\partial}{\partial \vec{\tilde{x}}} \cdot \frac{\partial}{\partial \vec{\tilde{x}}} \right) \vec{\chi} = -m(\vec{\hat{v}} - \vec{\mathbf{V}}_b) + 2D \frac{\partial m}{\partial \vec{\tilde{x}}}. \tag{5.17}$$

On the above equation, the underline in the l.h.s. indicates that the derivative is contracted only with the velocity field, and not with the auxiliary vector field $\vec{\chi}$. Using Einstein's summation convention, equation (5.17) becomes:

$$\frac{\partial \chi_k}{\partial \theta} + \frac{\partial}{\partial \tilde{x}_i} (\vec{\hat{v}}_i \chi_k) - \hat{D} \delta_{ij} \left(\frac{\partial}{\partial \tilde{x}_i} \frac{\partial}{\partial \tilde{x}_j} \right) \chi_k = -m(\hat{v}_k - \mathbf{V}_{bk}) + 2D \delta_{ki} \frac{\partial m}{\partial \tilde{x}_i}. \tag{5.18}$$

Using the same procedure described before, i.e. plugging in equation (5.9c) the solution obtained in (5.16), and applying solvability condition to the equation (5.9c), we finally obtain the effective equation governing the large scales dynamics, which will be expressed in Einstein's summation convention for simplicity:

$$\frac{\partial \rho^{(0)}}{\partial \tau} = \hat{D}^{\text{eff}}_{ij} \frac{\partial}{\partial X_i} \frac{\partial}{\partial X_j} \rho^{(0)}, \tag{5.19}$$

where

$$\hat{D}^{\text{eff}}_{ij} = \hat{D} \delta_{ij} - \langle (\hat{v}_i - \mathbf{V}_{bi}) \chi_j \rangle \tag{5.20}$$

In order to compute an effective equation of motion, the dimensionless variables must be replaced to the original variables t and $\vec{x}$, and the marginal probability density $\rho(t, x) = \int \mathrm{d}\vec{\tilde{x}} p$, i.e. the initial probability density, where the fast degrees of freedom are integrated out, must be used. For our assumption on the probability density function, equation (5.7), we conclude that $\rho = \rho^{(0)}$ and the effective long term equation looks like:

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_i} \left(\mathbf{V}_{bi} - D^{\text{eff}}_{ij} \frac{\partial}{\partial x_j} \right) \rho = 0\,. \tag{5.21}$$

Although the original variables t and $\vec{x}$ appear in equation (5.21), we know that this effective Kolmogorov forward equation is valid only in the long-term regime [107], and can be used to describe the dynamics of the process in (5.1) in a such regime.

Under arbitrary rotations the flow could be non-invariant. This property may affect the effective diffusion tensor $\boldsymbol{D}^{\text{eff}}$, which might be anisotropic [45,101,107]. By a physical intuition, we infer that this effective diffusivity tensor is symmetric and positive defined. It can be computed, however, only solving the auxiliary equation in (5.17).

5.4 Connection to papers

The multiscale analysis has been applied in Paper I, II, III and IV. For each paper, a proper scale separation for each process has been introduced and a hierarchy of equations, as in (5.9), has been developed.

In Paper I, we have developed an effective equation for an underdamped dynamics of a Brownian ellipsoidal particle, which performs translational and rotational motion in three dimensions. We have derived, therefore, the overdamped dynamics of a Brownian ellipsoidal particles, where fast degrees of freedom of velocity and angular velocity are averaged out. Moreover we have applied the multiscale analysis to thermodynamics functionals using the Feynman-Kac formula, see Appendix C of Paper I.

In Paper II, Paper III and Paper IV, we have developed an effective diffusion equation for an ovrdamped dynamics of a Brownian body, which performs translational and rotational motion in three dimensions, and subjected to an external force field. We have derived, thus, an effective equation where fast degrees of freedom of orientation are averaged out. In these case the calculation shown above, is similar to the one in the papers. The main difference with the hierarchy of equations (5.9) is the presence of the orientational degrees of freedom, which introduce new terms in the multiscale analysis. In these papers, we have recalled methodology of the literature for computing the effective diffusion tensor, which arises from the solvability conditions. Equation (5.21) correspond to equation (18) in Paper II and equation (52) in Paper III, but with the difference that the ballistic velocity and the effective diffusion tensor have been computed analytically.

Chapter 6

Stochastic Energetics

We are what we repeatedly do.
Excellence, then, is not an act, but a habit.

–Aristotle

In this chapter we give a short introduction of the theory of stochastic energetics. We start with a nutshell of thermodynamics in Sec. 6.1. In Sec. 6.2 we describe the stochastic energetics at trajectory level at mesoscopic scale. We end this chapter with Sec. 6.3, where we explain the connection of stochastic energetics with Paper I.

6.1 General remarks

In this section we briefly review classical thermodynamics, assuming that the reader is familiar with this field, in order to compare it with stochastic energetics [108].

Classical thermodynamics deals with the First and Second fundamental Laws governing transformations of a systems [*]. The First Law is also called "*the law of conservation of energy*" [12] and treats the exchange of heat, work and matter

---

[*] The Zeroth Law, which states, according to Sekimoto [2], that "*if a system is left isolated for sufficiently long time from any environment and from any external system, it will reach a state with no further macroscopic changes*", is assumed always valid, while the Third Law is out of scope of this thesis.

with the environment. The Second Law is about the variation of entropy. It is also called "*the entropy law*" [12] and states that the total entropy never decreases.

In this chapter we focus our attention on the physical quantities heat and work, leaving the discussion about the entropy to the next chapter, i.e. Ch. 7.

Classical thermodynamics is built upon empirical basis [109], where each macroscopic system is assumed as thermodynamics system. Thermodynamics variables, as temperature T, volume V etc., identify a thermodynamics state of the system. Energy E, volume V are extensive variables, whereas pressure P, temperature T and magnetic field are intensive.

The thermodynamic equilibrium is defined by a thermodynamic macroscopic state which does not change during the time.

A thermodynamics transformation is defined as a change of a state. If we start from an equilibrium state the transformation is given by an external control parameter. Thermodynamic transformations can be divided in two sets: reversible and irreversible transformations.

A transformation is called reversible if, inverting the time arrow, the transformation does the same path back in the history. If a transformation is not reversible, then is called irreversible.

If we are able to compute the mechanical energy of a system, then it is possible to identify this energy as internal energy of the macroscopical thermodynamics system. If we have an exchange of energy given by an external control parameter, this exchange of energy is called work. Heat, therefore, is the form of exchange of energy that is not given by work, or mass exchange [2].

The First Law of thermodynamics states that "*the balance of energy is established by changes in the internal energy of the system ΔE, the work W and the heat Q during any process*" [29]:

$$\Delta E = Q - W. \tag{6.1}$$

We have used the following sign convention: we define $W > 0$ when the system makes work on the environment and $Q > 0$ when the system receives heat from the environment[†]. In other words in any thermodynamic process the balance of exchanged work and heat is equal to the variation of the internal energy E [2].

The First Law of thermodynamics, therefore, distinguishes between two different forms of energy, which can be exchanged by a system with the environment. The addition of heat to a body is not due to the application of a generalized force, but by conduction, convection or radiation. In all these cases the body changes its thermodynamic state as a result of a change of motion at microscopic level [29].

In the next section we explain how Sekimoto [2,110] generalizes the First Law of thermodynamics at mesoscopic scale using the Langevin equation, opening, thus, the field of stochastic energetics.

[†] We inform the reader that we are following the usual sign convention for thermodynamics, in contrast to the convention used by Sekimoto in [2]

6.2 Stochastic energetics

The main idea of stochastic energetics is to describe the energy exchanges and statistical properties for mesoscopic systems, in analogy with classical thermodynamics for macroscopic systems. Mesoscopic internal energy, heat and work need to be, therefore, introduced.

In this section we introduce the field of stochastic energetics following the works of Sekimoto [2,110]. We present in detail the concept of heat along a trajectory, i.e. physics of the Langevin equation, and the First Law of thermodynamics for system governed by Langevin dynamics.

Langevin equations

We are interested in how to describe the exchanging of the energy between the particle and the thermal environment along an individual realization of such process.

We want to show explicitly the transferred energy along a single trajectory and for this reason we are not interested in describing our dynamics with the Fokker-Planck equation, because this equation deals with an ensemble of stochastic processes [2], but we are interested in using the Langevin equation, described in Ch. 2, because more appropriate.

The underdamped and overdamped Langevin equation describing such motion read:

$$\frac{d\vec{x}}{dt} = \vec{v}, \quad m\frac{d\vec{v}}{dt} = \vec{f}^{\text{ext}}(\vec{x}, \lambda) - \boldsymbol{\gamma} \cdot \vec{v} + \vec{\xi}(t); \tag{6.2a}$$

and

$$\boldsymbol{\gamma} \cdot \frac{d\vec{x}}{dt} = \vec{f}^{\text{ext}}(\vec{x}, \lambda) + \vec{\xi}(t). \tag{6.2b}$$

Equation (6.2a) describes the underdamped dynamics of the Brownian particle of mass m. We recall that $\vec{x}$ describes the position of the center of mass of the particle, while $\vec{v}$ the velocity of the particle. The friction tensor is here denoted by the symbol $\boldsymbol{\gamma} = \gamma\mathbf{I}$, where $\mathbf{I}$ is the identity matrix in three dimensions, while $\vec{f}^{\text{ext}}(\vec{x}, \lambda)$ is, in agreement with Sekimoto [2], an external force given by a potential $U(\vec{x}, \lambda)$, where λ is the control external parameter.

The external control parameter λ is governed by an agent which is able to control macroscopically the system through it. This agent is also called external system. The term "external" indicates that the evolution of the parameter λ is not determined by equation (6.2a) and (6.2b) [2].

Equation (6.2b) describes the overdamped dynamics of the Brownian particle, i.e. the mass of the particle is assumed small or the friction large. We recall that the overdamped approximation is obtained by the formal small inertia limit.

In both equations the random fluctuation force is modeled by Gaussian white noise sources, i.e. zero mean and correlation function $\langle \xi_i(t)\xi_j(t')\rangle = 2B\delta_{ij}\delta(t-t')$ with $t < t'$, where B is a measure of the strength of the fluctuation force satisfying the Fluctuation-dissipation theorem discussed in Sec. 2.2.

Definition of heat

By definition, the heat is an energy exchanged between the system and the thermal environment [2, 110]. We know that the system is described by the Langevin equations in (6.2), but how is the thermal environment defined?

The fluctuation force connects the environment and the particle with the strength of the fluctuation force given by the Fluctuation-dissipation theorem. The environment can be defined as the background of the system with which the system is in a perpetual contact. It is characterized by the thermodynamic variable temperature T. In other words, the environment is the thermal reservoir in equilibrium at constant temperature T. The time scale of the thermal bath is assumed always faster than the characteristic time of the process.

As described in the previous section, thermodynamics was developed at equilibrium as a theory of the macroscopic world. The evolution of the humanity, driven by the science, has brought to the description of the macroscopic world by the movement of its microscopic components, using the statistical mechanics. Statistical mechanics describes physical phenomena in terms of a statistical treatment of the behaviour of large numbers of atoms or molecules, i.e. microscopic components, especially as regards the distribution of energy among them. More precisely the primary goal of equilibrium statistical mechanics is to derive the classical thermodynamics of materials in terms of the properties of their constituent particles and the interactions between them. At mesoscopic grains the dynamics of a particle can be described by stochastic theories and the thermodynamics must be defined.

Following Sekimoto [2], we suppose that the law of action and reaction holds always between the Brownian particle and its environment: when the environment exerts a force $-\boldsymbol{\gamma} \cdot \frac{d\vec{x}}{dt} + \vec{\xi}(t)$ on the particle, given by the drag force $-\boldsymbol{\gamma} \cdot \frac{d\vec{x}}{dt}$ and the fluctuation force with zero mean, i.e. $\vec{\xi}(t)$, the particle exerts the reaction force, $-(-\boldsymbol{\gamma} \cdot \frac{d\vec{x}}{dt} + \vec{\xi}(t))$, on the environment [2].

The concept of heat, therefore, spontaneously arises as the form of exchange of energy from the environment to the system [2]. Along a trajectory described by the evolution of the center of mass of the particle, the work done by the Brownian body on the environment in a infinitesimal displacement $d\vec{x}(t)$, given by Langevin equation, is given by the product between the infinitesimal displacement $d\vec{x}(t)$ and the reaction force that the particle exerts on the environment, that is, $-(-\boldsymbol{\gamma} \cdot \frac{d\vec{x}}{dt} + \vec{\xi}(t)) \circ d\vec{x}(t)$ [2]. We have adopted the Stratonovich-type product $\circ$, to use the usual rules of calculus about the differentials.

A physical motivation about the adoption of the Stratonovich-type product is given by the fact that, as claimed by Wong [111], white noise does not exist in any real world system and, therefore, the continuous motion of Brownian particles needs to be considered within the framework of normal calculus [112, 113]. In contrast, in finance, Itô calculus is preferred because stochastic processes are truly discontinuous at short time scale [114].

The work can be positive or negative depending if the energy is lost by the

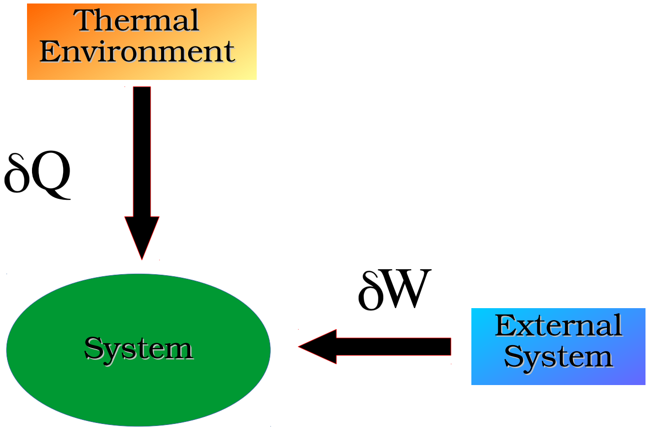

Figure 6.1: Schematic picture of the energy balance of a system, for example a particle, with its thermal environment and an external system [2].

system or by the environment, respectively. In both case there is a flux of energy between the environment and the Brownian body. This energy transfer has been defined as heat. In other words, the law of action and reaction, defined above from the Langevin equation, allows to quantify the amount of energy that has been transfered between the system and the environment.

Following the sign convention of Sekimoto [2], we assign the positive sign to the heat when the system receives energy from the environment. At mesoscopic scale, for instance, when the Brownian particle makes a positive work on the environment releases energy in the environment and the heat is negative. Then, we can define the equation when the system receives heat from the environment with:

$$\delta Q \equiv \left(-\boldsymbol{\gamma} \cdot \frac{d\vec{x}}{dt} + \vec{\xi}(t) \right) \circ d\vec{x}(t). \tag{6.3}$$

The δ is introduced to remind that the heat is not differentials.

The First Law

The definition of the heat at the mesoscopic scale given in equation (6.3) and the Langevin equations in (6.2) are sufficient to describe the First Law of thermodynamics at mesoscopic scale [2].

When the inertia is not negligible the formalism used to describe the evolution of a Brownian particle is the underdamped dynamics. In a such situation equation

(6.3) becomes $\delta Q = (m\frac{d\vec{v}}{dt} + \vec{\nabla}U) \circ d\vec{x}$. Using the Stratonovich calculus, we can obtain easily, from the definition of the heat for the underdamped dynamics, the following expression:

$$d\left(m\frac{\vec{v}\cdot\vec{v}}{2} + U(\vec{x},\lambda)\right) = \delta Q + \frac{\partial U}{\partial \lambda} \circ d\lambda. \tag{6.4}$$

The l.h.s. of equation (6.4) identifies the differential of the total energy $E = m\frac{\vec{v}\cdot\vec{v}}{2} + U(\vec{x},\lambda)$ of the system. In the r.h.s. of the equation we identify the work done by the external system on the particle system with $\delta W = \frac{\partial U}{\partial \lambda} \circ d\lambda$ and the balance of energy for the system becomes:

$$dE = \delta Q + \delta W, \tag{6.5}$$

which is the analogous of the First Law of thermodynamics, see Fig. 6.1. Equation (6.5) holds for each realization of the stochastic process [2].

If the variation of the external control parameter is abrupt, i.e. very fast respect to the characteristic time of the process $\tau = m/\gamma$, but slow respect to the time resolution of the Langevin model [2], the purely mechanical energy relation in (6.5) holds.

The relation in (6.5) can be also obtained when the inertia can be neglected. The main difference, as pointed out by Sekimoto [2], is that the term of the kinetic energy is not present in the complete differential form. In the case the temperature T does depend on the position $\vec{x}$, the balance of energy for the system can be obtained only by the underdamped dynamics.

In the last decades experiments on systems at mesoscopic scales have been performed. Typical experiments can be done recording the motion, controlled by an optical trap [42, 115, 116], of a micron sized particle, like microscopic colloidal particle suspended in a solution, with different shapes. From these experiments it is possible computing the amount of energy discharged in the solution (heat) and also have information on the work applied on the particle from the intensity and the focus of the optical trap [115].

One experiment noteworthy was performed by Blicke and Bechinger [115]. They were able to explore the limitations of classical thermodynamics, arising on scales which thermal fluctuations are important, i.e. mesoscopic scale. They demonstrated the experimental realization of a microscopic heat engine, comprising a single colloidal particle subjected to a time dependent optical trap. Moreover, they showed that the work associated with the system is a fluctuating quantity [115].

6.3 Connection to Paper I

In Paper I, we have used the definition of heat along the trajectory given by Sekimoto [2], see equation (8). Moreover, in this paper we have generalized the concept of heat to a Brownian particle which is able to translate and rotate in three dimensions. In equation (8) of Paper I is possible to note this.

The stochastic energetics has been useful to compute the heat performed by the particle on the environment. This has been helpful for define the entropy production of the system and for the environment, helping us to identify the anomalous term in the entropy production. Ch. 7 will explain in detail the anomalous entropy production term.

Chapter 7

Stochastic Thermodynamics and Entropy Production

There is nothing impossible to him who will try.

–Alexander the Great

In this chapter we discuss entropy production in stochastic thermodynamics. We start in Sec. 7.1 with an historical overview of entropy in thermodynamics and statistical mechanics. We therefore introduce Shannon's entropy and its representation for continuous distributions, and we also discuss criticisms about entropy for probability distributions in Information theory. In Sec. 7.2 we introduce the fluctuation theorem. In Sec. 7.3 we discuss entropy production in stochastic thermodynamics, introducing also Seifert's work and we end this chapter with a discussion about anomalous entropy production in Sec. 7.4.

7.1 Historical overview

In this section we recall the main basic concepts about entropy in thermodynamics and statistical mechanics, discussing also the Second Law of thermodynamics. We then introduce Shannon's entropy and its representation for continuous distributions. The last discussion is about criticisms about Shannon's entropy for continuous distributions.

Entropy in thermodynamics

Between 1850s and 1860s, the German physicist Rudolf Clausius coined an artificial word for describing a new quantity related to the energy of a thermodynamical system and also to its capability to change its thermodynamic state [117]. This word is entropy. It is composed by a root "*en*", which stands for energy, and the old Greek word $\tau\rho\acute{o}\pi o\varsigma$, which stands for transformation, response.

Clausius, in his work "Abhandlungen über die mechanische Wärmetheorie" published in 1864, described entropy as a transformation during a change of state, given by dissipative energy, of a thermodynamic system. According to him, the change of entropy of a system in classical thermodynamics is given by the ratio of heat received by the system to the absolute temperature during the heat absorption [118]. More precisely, the mathematical statement of entropy in classical thermodynamics reads:

$$\Delta S = \int_{a_{\text{state}}}^{b^{\text{state}}} \frac{\mathrm{d}Q_{rev}}{T} \,, \tag{7.1}$$

where the entropy change ΔS is described by integrating the quotient of the reversible heat addition and the absolute temperature over the change from the state a_{state} to state b^{state} of a given closed system [119].

Entropy is defined as a state function in classical thermodynamics, from which we infer that each infinitesimal reversible transformation is given by $\mathrm{d}S = \mathrm{d}Q_{rev}/T$.

In general the equality in (7.1) may be substituted with an inequality, indeed for irreversible processes Clausius postulated that in a closed system entropy cannot decrease, which also known as Second Law of thermodynamics.

The Second Law of thermodynamics states that "*the total entropy of any system cannot decrease other than by increasing the entropy of some other system. Hence, in a system isolated from its environment, the entropy of that system does not decrease*" [29].

Entropy in statistical mechanics

Boltzmann, Gibbs, and Maxwell gave to entropy a probabilistic interpretation around 1870s and later.

Boltzmann proposed a statistical formulation given by his famous formula:

$$S_{(B)} = k_B \ln W \,, \tag{7.2a}$$

where

$$W = \frac{N!}{N_1!N_2!....N_s!} , \tag{7.2b}$$

with the N's representing possible states, configurations, of the system and k_B the Boltzmann constant [119]. In words Boltzmann's formula in combinations with the Second Law says that systems tend to go from order to disorder, i.e. low entropy stands for high order, while high entropy stands for low order.

Gibbs generalized the Boltzmann formula using probability theory. The formula reads:

$$S_{(G)} = -k_B \sum_i P_i \ln P_i \tag{7.3}$$

where k_B is Boltzmann's constant. P_i identifies the probability of a microstate i with energy E_i. A microstate is defined as microscopic configuration of a thermodynamic system that the system, in the course of its thermal fluctuations, could occupy with a certain probability. At each microstate is associated the energy E_i.

Maxwell helped the development of statistical mechanics with the formulation of the kinetic theory [120]. Indeed, the Maxwell–Boltzmann statistics describes the average distribution of non-interacting material particles over various energy states in thermal equilibrium, and is applicable when the temperature is high enough or the particle density is low enough to render quantum effects negligible.

Maxwell is also famous for his ideal experiment, in which he suggested how the Second Law of thermodynamics might hypothetically be violated, i.e. the Maxwell's demon which is:

" *... if we conceive of a being whose faculties are so sharpened that he can follow every molecule in its course, such a being, whose attributes are as essentially finite as our own, would be able to do what is impossible to us. For we have seen that molecules in a vessel full of air at uniform temperature are moving with velocities by no means uniform, though the mean velocity of any great number of them, arbitrarily selected, is almost exactly uniform. Now let us suppose that such a vessel is divided into two portions, A and B, by a division in which there is a small hole, and that a being, who can see the individual molecules, opens and closes this hole, so as to allow only the swifter molecules to pass from A to B, and only the slower molecules to pass from B to A. He will thus, without expenditure of work, raise the temperature of B and lower that of A, in contradiction to the second law of thermodynamics.*" [121–123].

Entropy and information: Shannon's entropy

Entropy was slowly accepted and was not easy to understand in contrast to concepts as energy, work, temperature etc. As a quantity that cannot be measured directly, entropy has attracted some kind of mystic aurea [117]. For example, when Shannon rediscovered entropy in 1948, he turned to von Neumann and asked him an advice about how to call this quantities. Von Neumann told him:

"... call it entropy. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, no one knows what entropy really is, so in a debate you will always have the advantage." [119]

In Information theory Shannon's entropy describes the minimal information capacity of the transmission channel that is required to transmit the signal. In other words the information or entropy of a system is the minimal number of bits needed to specify it configuration [117]. The formal mathematical definition of the Shannon's entropy is:

$$S_{(S)} = -\sum_{\omega \in \Omega} P_\omega \ln P_\omega \,, \tag{7.4}$$

where ω is a configuration of the physical system and Ω is the whole space of the configuration, i.e. $\sum_{\omega \in \Omega} P_\omega = 1^*$. Equation (7.4) is Gibbs entropy, i.e. equation (7.3), when each configuration ω is a microstate of the system.

In this thesis we are interested in continuous random processes and therefore we identify with ω the outcome of a random experiment. A random variable is a map X which associates with each outcome ω an element of a different set, usually a real number, i.e. $X : \omega \to x = X(\omega)$. For a sake of simplicity we fix to one the dimension of the experiment. Shannon's entropy in this case is called in textbook continuous or differential entropy [117].

To do that we first start to discretized the abscissa into of the width Δx and enumerate them with the index i. To each of these intervals we attach a column of the height $p(x_i)$ centered horizontally at x_i. In this way we generate a discretized distribution, where the area of each column $p(x_i)\Delta x$ represents the probability of x_i. Making the limit of $\Delta x \to 0$ the sum must be replaced by an integral and the probability becomes a density distribution [117]. The formula for the differential entropy therefore reads:

$$S_{(S)_{\text{diff}}} = -\int_{-\infty}^{+\infty} dx\, p(x) \ln p(x), \tag{7.5}$$

where the equation (7.5) is defined formally up to a divergent constant. Indeed, during the process of limit $\Delta x \to 0$, we have subtracted a constant dependent by the resolution of the discretization but not on the shape of the distribution. In the next subsection some criticism will be shown about the definition of the differential entropy.

Criticisms on the differential entropy

Kolmogorov in the papers "On the Shannon Theory of Information Transmission in the Case of Continuous Signals" [124] and "The theory of transmission of information" [125], the first one presented in September 1956 and the second one presented in October 1956 [53, 125], claims that "differential entropy", here called $S_{(S)_{\text{diff}}}$ in equation (7.5), does not have a straightforward meaningful interpretation. Indeed, it has been defined up to a divergent constant dependent by the resolution of the discretization.

Moreover, Kolmogorov [53, 124, 125] observed that equation (7.5) is even non-invariant with respect to transformations of coordinates in the x-space. Indeed,

---

[*] The entropy in this case is dimensionless, i.e. k_B is set to one, to be uniform with information theory formalism.

a distribution function transforms as a density. Under general coordinate transformations in the logarithm of equation (7.5) the Jacobian appears, which does not have any physical meaning. Only for which transformations where the Jacobi determinant is equal to 1, the differential entropy is invariant, because the local volume is preserved [117].

7.2 Fluctuation theorem

The fluctuation theorem was developed independently by Evans et al. [126] and Gallavotti et al. [127] and after studied and applied in non equilibrium thermodynamics and stochastic thermodynamics by Crooks, Evans, Kurchan, Jarzynski, Seifert and others [126, 128–147] and experimentally studied by Ciliberto, Collin, Wang and others [148–150]

Fluctuation theorem refers to the Second Law of thermodynamics, but being an equation rather than an inequality it is much stronger [117].

We will give now the formal definition of fluctuation theorem and integral fluctuation theorem following Hinrichsen [117]. We refer to the scientific literature, cited above, for a formal derivation of these theorems as this is out of the scope of this thesis.

Let's consider, as described in previous section, a real random variable X and the corresponding probability density $p(X)$. The fluctuation theorem is a relation of the following form:

$$p(X) = \mathrm{e}^{X} p(-X). \tag{7.6}$$

This equation tells us that the left part of distribution, the one indicated with $p(-X)$, goes to zero exponentially. This expression restricts the shape of the function $p(X)$ to distributions which relate the left and right half of themselves. In other words, when the the function of the positive random number $X > 0$ is given, immediately we can infer the function for all negative random number $X < 0$.

By the observation that the probability density is an even function, i.e. $p(X) = p(-X)$, normalized to one, we can immediately infer the integral fluctuation theorem:

$$\langle \mathrm{e}^{-X} \rangle = 1, \tag{7.7}$$

where $\langle \cdot \rangle$ identifies the integral over whole space where the density distribution function is defined. Equation (7.7) is also known as integral fluctuation theorem.

If we use the Jensen's inequality for convex functions, which by definition relates the value of a convex function of an integral to the integral of the convex function [117], we infer from the integral fluctuation theorem that the first moment of the distribution is positive, i.e. $\langle X \rangle \geq 0$. Plugging in X the ΔS we obtain the conventional Second Law of thermodynamics.

Crooks in 1998 [128] generalized the formulation of the fluctuation theorem to stochastic microscopically reversible dynamics. This theorem, as made by Crooks in 1999 [129], can be also applied to the entropy production $\Delta S / k_B$. The theorem states:

$$\frac{p_F(\Delta S)}{p_R(-\Delta S)} = \mathrm{e}^{\Delta S/k_B}, \tag{7.8}$$

where $p_F(\Delta S)$ is the probability distribution of this entropy production [129], and $p_R(-\Delta S)$ is the probability distribution of the entropy production when the system goes on time-reversal path [129]. The distinction between forward entropy and reversed entropy is mandatory. Indeed, systems driven by a time process are considered.

7.3 Entropy production in stochastic thermodynamics

Stochastic thermodynamics

Stochastic thermodynamics has emerged as a new branch of thermodynamics specially aimed at describing the time evolution of thermodynamic quantities such as work, heat, and entropy [151].

The origin of this field can be found in works of Vulpiani, Sekimoto, Sasa, Van der Broeck, Seifert, Jarzynski etc. [152–160].

The system considered by stochastic thermodynamics are composed by few components (e.g. molecules). The amount of energy of such system is small, of order of magnitude of $1 - 10^2\ k_BT$ [161], and to emphasize the smallness of the energy scale, these systems are also referred to as "small systems" [161]. For example, small systems obeying to stochastic thermodynamics include: few molecules experiments [161], Brownian motion, etc. In this thesis, however, we are interested in only diffusion processes.

At mesoscopic scales, as introduced in Ch. 1, fluctuations play an important role. In contrast to the macroscopic scale, where fluctuations in a system vanish with one over the square root of N, with N the number of components, because of the enormous number of components which it is made, at mesoscopic scale fluctuations are clearly not negligible. As a consequence, a realization of a mesoscopic process can differ strongly from its average behavior.

Important results in the field of stochastic thermodynamics are the Jarzynski equality [146, 162–164] and the Crooks fluctuation theorem [128, 129, 143]. The Jarzynski equality relates the difference of Helmholtz's free energy between two states to the irreversible work along a ensemble of trajectories joining the same states [164]. The Crooks fluctuation theorem, instead, relates the work done on a system during a non-equilibrium transformation to the difference of Helmholtz's free energy between the final and the initial state of the transformation [129]. However, in this thesis, we are interested in the entropy production and not in the work. For this reason we leave the entire chapter to the entropy production, while we refer to the literature for these two results concerning work.

Entropy production

Entropy is produced in irreversible processes (see Ch. 6). It is not only an ensemble property as described before, but also can be applicable to a single trajectory. We will explain in this section how it is possible.

Let's assume an overdamped motion of a Brownian spherical particle in three dimensions, generalizing the work of Seifert in [133], in a time interval $0 \leq t \leq t''$. Let's assume also that the particle is connected to a homogeneous thermal bath with constant temperature T and friction coefficient γ. The friction tensor, therefore, will be $\boldsymbol{\gamma} = \gamma\mathbf{I}$, where $\mathbf{I}$ is the identity matrix. The friction tensor is by definition symmetric and positive definite.

Einstein relation holds at equilibrium, i.e. $\boldsymbol{D} = \boldsymbol{\mu} k_B T$, where $\boldsymbol{D}$ is the diffusion tensor given by relation $\boldsymbol{D} = D\mathbf{I}$, and $\boldsymbol{\mu} = \boldsymbol{\gamma}^{-1}$ is the mobility of the particle. We assume also that such relation persists even in a non equilibrium situation [133].

The particle is subject to an external force which is given by $\vec{F}(\vec{x}, \lambda) = -\partial_{\vec{x}} V(\vec{x}, \lambda) + f(\vec{x}, \lambda)$, where $V(\vec{x}, \lambda)$ is a conservative potential and $\vec{f}(\vec{x}, \lambda)$ is a force that can be applied directly on the particle.

The variable λ is the external control parameter varied by some experimental protocol and may be dependent on time. The three dimensional motion of the Brownian particle is therefore described by the overdamped Langevin equation:

$$\dot{\vec{x}} = \boldsymbol{\mu}\vec{F}(\vec{x}, \lambda) + \vec{\xi}(t). \tag{7.9}$$

The random fluctuations are modeled by Gaussian white sources with correlation function given by $\langle \xi_i(t)\xi_j(t') \rangle = 2D\delta_{ij}\delta(t - t')$ and zero mean.

As described in Ch. 2 Sec. 2.2 we can define a probability density $\rho(\vec{x}, t)$, which describes the probability to find the Brownian particle in the position $\vec{x}$ at time t, that satisfies the Fokker-Planck equation. The Fokker-Planck associated to equation (7.9):

$$\frac{\partial \rho(\vec{x}, t)}{\partial t} = -\vec{\nabla} \cdot \vec{J}(\vec{x}, t) \tag{7.10a}$$

where

$$\vec{J}(\vec{x}, t) = (\boldsymbol{\mu}\vec{F}(\vec{x}, \lambda) - \boldsymbol{D} \cdot \vec{\nabla})\rho(\vec{x}, t). \tag{7.10b}$$

From here, Seifert defined in [133] the following entropy for the particle along a trajectory $\vec{x}(t)$ for the system:

$$s(\vec{x}(t)) = -k_B \ln \rho(\vec{x}(t), t). \tag{7.11}$$

where $\rho(\vec{x}(t), t)$ is the solution of the Kolmogorov forward equation evaluated along the trajectory $\vec{x}(t)$ performed by the Brownian particle. Obviously, the entropy $s(\vec{x}(t))$ depends on initial boundary conditions for any given trajectory [133]. Equation (7.11) has been suggested by the definition of a non equilibrium Gibbs entropy for continuous processes, which is in the form of equation (7.5) when it is multiplied with k_B.

In equilibrium, i.e. for $f(\vec{x},\lambda) = 0$ and constant λ [133, 165], the definition of entropy production given in equation (7.11) assigns a stochastic entropy given by $s(\vec{x}(t)) = (V(\vec{x}(t),\lambda) - \mathcal{F}(\lambda))/T$, where $\mathcal{F}(\lambda)$ is the free energy given by the usual formula in statistical mechanics. Therefore at any time this thermodynamic relation holds along the trajectory.

Along the trajectory $\vec{x}(t)$ the average rate of change of the total entropy is given by $\dot{S}_{\text{tot}}(t) \equiv \langle \dot{s}_{\text{tot}}(t)\rangle = k_B \int d\vec{x}(\vec{J}(\vec{x},t) \cdot \vec{J}(\vec{x},t))/(D\rho(\vec{x},t)) \geq 0$, where equality holds only at equilibrium. The average rate of change of the entropy of environment is given by $\dot{S}_{\text{env}}(t) \equiv \langle \dot{s}_{\text{env}}(t)\rangle = \int d\vec{x}\vec{F}(\vec{x},\lambda) \cdot \vec{J}(\vec{x},t)/T$. The two entropy can be computed by the equation (7.11), we recall the reader interested in this to [133].

The rate of change for the system therefore can be easily computed as $\dot{S}(t) \equiv \langle \dot{s}(t)\rangle = \dot{S}_{\text{tot}}(t) - \dot{S}_{\text{env}}(t)$.

The entropy production in the environment can be computed using the Sekimoto's stochastic energetics approach [2] (see also Ch. 6). The heat received to the thermal bath during the motion of the particle is identified with the work done by the particle on the environment. The infinitesimal entropy production in the environment dS_{env} is equal to $\delta Q/T$. Using therefore the action reaction force to define the heat, see Ch. 6, we obtain that the ΔS_{env} can be defined as:

$$\Delta S_{\text{env}} = \int_{t_0}^{t} \frac{1}{T} \left[\vec{F}(\vec{x},\lambda)\right] \circ d\vec{x}(t') \tag{7.12}$$

where we have adopted the Stratonovich-type product $\circ$, to use the usual rules of calculus about the differentials. The term $d\vec{x}(t')$ identifies the infinitesimal displacement of the trajectory given from the Langevin equation (7.9). In Paper I is presented also the calculation for the entropy production in the environment given for a system described by underdamped dynamics.

The most significant result given by Seifert [133, 134] is that the total entropy production change along a trajectory $\Delta S_{\text{tot}} = \Delta S_{\text{env}} + \Delta S$, with ΔS given by the difference between the logarithm of the probability density from the initial point to the final point, obeys an integral fluctuation theorem [133]:

$$\langle \mathrm{e}^{-\Delta S_{\text{tot}}/k_B} \rangle = 1, \tag{7.13}$$

which is equal to the Second Law of thermodynamics $\langle \Delta S_{\text{tot}} \rangle \geq 0$. According to Seifert [165], this integral fluctuation theorem for the total entropy production ΔS_{tot} may be universal. It may hold for any kind of initial condition, any force which are time dependent, any potential and any length of trajectory.

As described in the in Sec. 7.1 the main criticism that one can observe in the differential entropy and in equation (7.11) is that the entropy is not invariant under general coordinate transformations. To solve this problem Seifert in [165, 166] introduces, explicitly, the correction with the Jacobian of the transformation $x \to y$. The equation (7.11), thus, becomes in a new set of coordinates $s(\vec{y}(t)) = -k_B \ln[\rho(\vec{y}(t),t)\text{det}(\frac{\partial y}{\partial x})]$, (formally the term in the logarithm is assumed multiplied to the correspondent thermal de Broglie length to make it dimensionless).

Criticisms

However, we think that introducing the correction with the Jacobian of the transformation $x \to y$ into the equation, i.e. as Seifert did in [165,166], does not answer to criticisms rose about the non-invariance under general coordinate transformations. We think, instead, that an answer may be sought following the "way" treaded by Kolmogorov over seventy years ago.

We have described continuous processes, i.e. Brownian motion of particle, on a n-dimensional closed manifolds in Ch.4. We have started to define $\mathcal{M}$ as a closed Riemannian manifold and to assume that a particle performs Brownian motion on manifolds. The density probability that the particle is in a certain position $x \in \mathcal{U} \subset \mathcal{M}$ at time t is given by $\rho(t,\vec{x}) = \sqrt{g(\vec{x})} h(t,\vec{x})$, where $g(\vec{x})$ is the determinant of the metric tensor of the manifold $\mathcal{M}$, while $h(t,\vec{x})$ is a scalar function in $\vec{x}$ that satisfies the requirements in equations (4.10a) and (4.10b) in Ch.4, i.e. the scalar function $h(t,\vec{x})$ is normalized to 1 on the manifold and that the stochastic process describing the Brownian motion on manifold is Markovian.

From classical thermodynamics and statistical mechanics, we know that the entropy at equilibrium reaches the maximum. Let's assume, therefore, that always exists an hypothetical equilibrium state that will be reached if the source terms, which drive the system out of equilibrium, are switched off. The entropy of this state will be given by $S_{(eq)} = k_B \ln W$, where W determines the value of the entropy of total system (Brownian particle and fluid) at equilibrium [12]. For example, assume that the particle, which performs Brownian motion on manifold, is subject to an external force field. The presence of this force will be of course a source of entropy production in the system. When the external force field will be shut down, the particle will reach an equilibrium state. This equilibrium state is the state that we are looking for. Indeed, we assume that at this equilibrium state is associated a density probability given by $\rho^{(\mathrm{eq})}(t,\vec{x}) = \sqrt{g(\vec{x})} h^{(\mathrm{eq})}(\vec{x})$.

Using the Kullback–Leibler divergence [167] for density probability we obtain:

$$S_{(\mathcal{M})_{\mathrm{diff}}} = -k_B \int_{\mathcal{M}} dV_{\vec{x}}\, h(t,\vec{x}) \ln \frac{h(t,\vec{x})}{h^{(\mathrm{eq})}(\vec{x}) W}, \tag{7.14}$$

where $dV_{\vec{x}}$ is the invariant volume on manifold $\mathcal{M}$.

Let's discuss in detail equation (7.14). First of all, we observe that the formula in (7.14) is invariant with respect to transformations of coordinates in the x-space and that it is defined up to a constant. Indeed, it is easily to see that by construction the equation (7.14) is invariant under general coordinate transformations, because $dV_{\vec{x}}$ is the invariant volume on manifold $\mathcal{M}$ and the ratio in the logarithm is dimensionless. Moreover, this equation does not show anymore the presence of a divergent constant dependent by the the resolution of discretization because that term goes away in the ratio of the logarithm.

When the system reaches the hypothetical state of equilibrium the equation (7.14) becomes equal to $S_{(\mathrm{eq})} > 0$. The equation is also additive and definite positive, and therefore can be identified as entropy.

This equation is not completely new in the scientific literature. Already Van Kampen, in "The definition of entropy in non-equilibrium states" published in 1959 [168], and de Groot and Mazur in "Non-equilibrium thermodynamics" [12] defined in a similar way the entropy from the "Gibbs entropy postulate" [12].

In the equation of these authors it is present, therefore, the Boltzmann's constant k_B, which does not play a relevant role by observations of Jaynes in 1957 [169, 170]. Our equation matches the equation of Van Kampen [168] and de Groot and Mazur [12] in the case of flat space. However it generalizes their results for Brownian motions on manifolds.

It is simple to check that equation (7.14) does not affect in anyway the fluctuation theorem. Indeed, we have assumed that the metric tensor does not depend by the time.

One, therefore, can use the following definition of entropy production for the system along a trajectory on a manifold:

$$s(\vec{x}(t)) = -k_B \ln \frac{h(t,\vec{x})}{h^{(\mathrm{eq})}(\vec{x})W}. \tag{7.15}$$

This equation avoids the problem of the change of coordinates because the term defined inside the logarithm is by construction a scalar function. The physical meaning must be the same of Seifert's entropy.

7.4 Anomalous entropy production

We have seen in the previous chapter how to define for a Brownian particle thermodynamics quantities as energy, heat, entropy, work etc. In this chapter, we have given an introduction for the entropy production along a single trajectory. This section is devoted to describe the arising of anomalous entropy production proposed in [47, 171, 172] and also analyzed in Paper I.

Definition of anomalous entropy production

Along this thesis we have also shown how compute effective equations for the dynamics of Brownian particle averaging out the fast degrees of freedom. We want to ask now the following question: may this procedure of integration of fast degrees of freedom of the process affect the thermodynamics and in particular the entropy production?

We show a simple example. Let's take the motion of a Brownian sphere described by underdamped Langevin equations. Let's imagine, for simplicity, that the particle performs only translational Brownian motion on a fluid in a vessel A where the temperature T of the fluid is constant. Let's assume that the mass of the particle is small and the friction is large. Under this assumption the system can be also expressed using the overdamped approximation. In both formulations, i.e. underdamped and overdamped dynamics, it is possible compute the thermodynamics quantities. We note that the entropy production described by the effective equation, i.e. by the overdamped equation, has the same structure of the entropy production described from the underdamped equations. In this case the effective equation describes the dynamics and thermodynamics of the system as the underdamped formalism does, and, therefore, the answer of our question seems to be negative.

Let's now take the same Brownian spherical particle. This time, we allow the particle to perform a Brownian motion in a different vessel, for simplicity we call this vessel B. This vessel has the peculiarity that the fluid, which its contains, displays a temperature that is not anymore constant, but it is dependent by the position, i.e. $T(\vec{x})$[†]. Going to compute the thermodynamics, i.e. the entropy production, for the underdamped case and the overdamped case, we obtain two different expressions for the entropy production. In this case the answer of our question is positive. We claim, thus, that, the overdamped approximation does not estimate completely the thermodynamics of the Brownian particle, although correctly, it describes its evolution in space [171].

The term that is missing in the expression of entropy production in the overdamped approximation is called "*anomalous entropy production*". It was dubbed with this name because it has some property similar to other anomalies in physics [46, 47, 171]. As noted by Celani et al. [171], the most famous example of anomaly encountered in classical physics is the "*viscous dissipative anomaly*" [13]. This viscous dissipative anomaly is related to the energy dissipation in a fluid flow. A such dissipation of energy does not vanish even in the small viscosity limit. When the viscosity is exactly zero, then this viscous dissipative anomaly vanishes [13]. This "*viscous anomaly*" identifies the loss of time-reversal symmetry of fluid dynamics when it goes from the non-viscous case to the viscous one. The entropic anomaly shows the same behavior. Indeed, in this entropic case the broken symmetry is the time-reversal symmetry implied by detailed balance of the fast state.

Physical interpretation of anomalous entropy production

To explain better this phenomenon we refer to the example of the Brownian spherical particle that performs translational motion in the vessel B, i.e. the fluid, where that particles is suspended, displays a gradient of temperature.

The overdamped dynamics is obtained averaging out the velocities, i.e. the fast degrees of freedom. It is assumed, fairly, that the velocities instantaneously equilibrate to the local Maxwell Boltzmann distribution. In three dimensions this distribution is given by $w(\vec{v}|\vec{x}) = (m/(2\pi k_B T(\vec{x})))^{3/2}\exp[-m\vec{v}\cdot\vec{v}/2k_B T(\vec{x})]$, where m is the mass of the particle.

This distribution is symmetric under time inversion, indeed the position $\vec{x}$ does not change under time reversal because even variable, while the fast degrees of freedom, although it is a odd variable, i.e. changes under time inversion, in the Maxwell Boltzmann distribution appears multiplied by itself.

[†] We assume a smooth temperature profile, e.g. a linear one [171]

The underdamped dynamics, instead does not show this symmetry. As computed by Celani et al. [171] and by us in Paper I, the velocity statistic is given by:

$$p(\vec{v}|\vec{x}) = w(\vec{v}|\vec{x}) \left(1 + \frac{5k_BT(\vec{x}) - m\vec{v}\cdot\vec{v}}{6\gamma k_B^2 T^2(\vec{x})} m\vec{v}\cdot\vec{\nabla}k_BT(\vec{x})\right), \tag{7.16}$$

where terms of order $\boldsymbol{\gamma}^{-2}$ or higher (we recall that the particle is spherical and the friction tensor is given by $\boldsymbol{\gamma} = \gamma\mathbf{I}$) are negligible. The small correction, which shows the broken symmetry in the time inversion for velocity, plays an important role for the production of entropy. Indeed, the term that effectively breaks the time-reversal symmetry is the momentum of the particle. The momentum of the particle, therefore, is a source of entropy production, because is not balanced when we invert the time arrow.

By the fact that the temperature is space dependent, during the motion the particle visits positions at different temperatures and exchanges heat at different temperatures. This process produces entropy and it takes place with a frequency of order of magnitude γ/m. This exchanging of heat, together with the asymmetry in the velocity, yields to a finite contribution of entropy. However, when we start with the overdamped approximation our analysis for the thermodynamics quantities, we forget, i.e. we neglect, this anomalous term.

The equation (7.16) shows, as discussed by Celani et al. [171], that trajectories which follow a down gradient motion with speeds larger than the thermal velocity, occur with higher probability (positive sign of the correction in the equation) than trajectories which follow a slow up gradient motion. This correction in equation (7.16) does not take part in the spatial flux of particles. Indeed, we can directly check this by multiplying by $\vec{v}$ and integrating over velocity. Moreover, the mean kinetic heat exchange is not affected by this small correction. Indeed, this correction is odd in velocity coordinates.

The exchange of heat occurs during a cycling. At higher temperature the heat is absorbed by the particle and at lower temperature the heat is released into the environment by the particle. The duration of a cycle is of order of magnitude m/γ, i.e. the characteristic time of the process. The order of magnitude of the entropy produced is, therefore, given by the difference of inverse temperatures at extremes of the trajectory times heat exchanged, i.e. $\sim k_BT^{-1}|\nabla T|\gamma^{-1}m\sqrt{k_BT/m}$. A similar contribution with opposite sign is given, instead, by the cycles traveled in the reversal way. However, the cycles that are characterized by a travel forward in time occur with higher probability than cycles that are characterized by a travel backward in time. This probability is given by $\sim T^{-1}|\nabla T|\gamma^{-1}m\sqrt{k_BT/m}$ obtained by equation (7.16). The last observation allows to compute the average entropy production rate. This rate has an order of magnitude equal to $\sim \gamma^{-1}T^{-1}|\nabla k_BT|^2$ [171, 172], which is the anomalous term rate.

Using Seifert's definition for entropy production for the particle, equation (7.11), the source of the anomaly can be isolated from the total entropy production of the original Langevin-Kramers dynamics, i.e. underdamped dynamics. The total entropy production will be given by two contribution: the regular entropy production and the anomalous one. The regular entropy production is identified as the entropy produced in the process that is the same of the entropy production produced in the overdamped approximation and, in the case where no external force fields are applied, reads:

$$\Delta S_{\mathrm{reg}} = k_B \int_{t'}^{t} \frac{\vec{v}(\tau) \cdot \vec{\nabla} T(\vec{x}(\tau))}{T(\vec{x}(\tau))} d\tau, \tag{7.17}$$

where $t' < t$.

The anomalous entropy production term, instead, is given by :

$$\Delta S_{\mathrm{anom}} = k_B \int_{t'}^{t} \frac{5k_B T(\vec{x}(\tau)) - m\vec{v}(\tau) \cdot \vec{v}(\tau)}{2k_B^2 T^2(\vec{x}(\tau))} \vec{v}(\tau) \cdot \vec{\nabla} k_B T(\vec{x}(\tau)) d\tau, \tag{7.18}$$

where $t' < t$.

In the small inertia limit, the anomalous entropy production obeys to the integral fluctuation theorem (7.13):

$$\langle \mathrm{e}^{-\Delta S_{\mathrm{anom}}/k_B} \rangle = 1. \tag{7.19}$$

Moreover, using the multiscale analysis, has been proved in [47, 171] that does not exists any sequential functional of the overdamped trajectories that gives limiting statistic of ΔS_{anom} (see Supplemental Material [171]). In other words, no redefinition of the entropy production, obtained from the overdamped dynamics, can give this "anomalous" term [171].

The average rate of the anomalous entropy production, obtained by the multiscale analysis of functionals [171, 172], is given by :

$$\frac{d\langle \Delta S_{\mathrm{anom}} \rangle}{dt} = \frac{5}{6} \int \rho(\vec{x}, t) \frac{(k_B \vec{\nabla} T) \cdot (k_B \vec{\nabla} T)}{\gamma T} d\vec{x}. \tag{7.20}$$

As we have introduce in the previous sections, entropy is a measure of the irreversibility of a process. In other words, the ratio between the probability of the reversed path and the forward one allows to compute the entropy produced in the environment along the trajectory [140].

Following Celani et al. [171], we can use the fluctuation theorem to quantify this irreversibility. Indeed, in the underdamped dynamics, we know that the fluctuation theorem is satisfied by entropy production in the environment. More precisely, reversal process is a process where the time arrow is inverted, which implies the inversion also in the direction of the velocity. Mathematically one gets:

$$\frac{P(\vec{x}_{[t,t']}, -\vec{v}_{[t,t']})}{P(\vec{x}_{[t',t]}, \vec{v}_{[t',t]})} = \mathrm{e}^{-\Delta S_{\mathrm{env}}/k_B}. \tag{7.21}$$

Similarly, in the overdamped case is possible compute the entropy production in the following way:

$$\frac{P^{(\mathrm{over})}(\vec{x}_{[t,t']})}{P^{(\mathrm{over})}(\vec{x}_{[t',t]})} = \mathrm{e}^{-\Delta S_{\mathrm{env}}^{(\mathrm{over})}/k_B}. \tag{7.22}$$

From the ratio of equation (7.21) and (7.22), recalling the presence of scale separation, it is possible to have that, in the small inertia limit, the anomalous entropy production satisfies the following relation:

$$\frac{P(-\vec{v}_{[t,t']}|\vec{x}_{[t,t']})}{P(\vec{v}_{[t',t]}|\vec{x}_{[t',t]})} = \mathrm{e}^{-\Delta S_{\mathrm{anom}}/k_B}. \tag{7.23}$$

Equation (7.23) shows that the breaking of time and velocity reversal symmetry at a given spatial position is associated to anomalous term [171].

Connection to Paper I

In Paper I, we have computed anomalous entropy production for an ellipsoidal Brownian particle which performs rotational and translational motion. We have generalized the result of Celani et al. [171] and Lan et al. [47].

We have started defining the entropy production of the particle and of the environment and we ends, after a long and tedious calculation with the multiscale analysis, with the anomalous entropy production rate for this specific system.

The equation (7.18) becomes more complicated because of the presence of the rotational degrees of freedom, see equation (16.c) Paper I, but satisfying always the integral fluctuation theorem, see equation (49) Paper I. Also the rate of the anomalous entropy production found becomes more complicated, because of the presence of the rotational degrees of freedom, see equation (51.c) Paper I.

From the result obtained in equation (51.c) Paper I, we have also pointed out that without translational motion anomalous entropy production cannot arise.

The result obtained by us in this paper is valid for any particle with three mutually perpendicular symmetry planes [173], including, in particular, the large class of spheroids, but also rods and rod-like shaped objects.

Appendix A

In this appendix we describe the connection between the Rallison's method [174] used in Paper II and the differential geometry method used in Ch. 4 and Paper III.

We start by describing how to derive an expression for the rotational differential operator in terms of derivatives with respect to a generic rotation matrix $\boldsymbol{R}$ [174]. We then recall the main mathematical definitions of $SO(3)$ and $SU(2)$ group manifolds. We end this appendix describing the rotational differential operator in local coordinates on manifolds.

Rotational differential operator and infinitesimal rotations

We start by defining, following the formalism of J. M. Rallison [174], two frames of reference for the particle orientation:

- the body frame S', which is free to rotate:
- an initial reference frame S, which allows to define a rotating frame fixed in the particle S' a time $t = 0$.

As well described in many books of classical mechanics [41, 175, 176], proper rotations in three dimensions are identified by the set of special orthogonal matrices, with determinant equal to $\det(\boldsymbol{R}) = 1$ and such that $\boldsymbol{R}\boldsymbol{R}^{\intercal} = \boldsymbol{R}\boldsymbol{R}^{-1} = \mathbf{I}$, where $\mathbf{I}$ identifies the identity matrix and where $^{\intercal}$ identifies the transpose. $\boldsymbol{R}$ is a matrix 3×3. Such rotations allow a vector to rotate with the transformation law:

$$x'_i = R_{i\omega} x_\omega, \tag{24}$$

where x_ω identifies the vector in the reference frame S at time $t = 0$ and x'_i the resultant rotating vector in the new reference frame S'. We adopt Einstein summation convention and the convention that tensors to be evaluated in the reference state shall have Greek suffix, while Latin suffix are used in the current state [174].

Assume now to perform an infinitesimal rotation $\boldsymbol{\phi}$. An infinitesimal rotation is a special orthogonal transformation of coordinate axes in which the components of a vector are almost the same in both sets of axes, i.e. the change is infinitesimal [41].

Here we recall some property about infinitesimal rotations following H. Goldstein in [41].

For any general component x'_i the equations of infinitesimal transformation can be written in matrix notation as:

$$\vec{x}' = \boldsymbol{\phi}\vec{x} = (\mathbf{I} + d\boldsymbol{R})\vec{x}\,. \tag{25}$$

Equation (25) claims that the form of an infinitesimal transformation is $\boldsymbol{\phi} = (\mathbf{I} + d\boldsymbol{R})$, i.e. it is almost the identity transformation, differing at most by an infinitesimal operator $d\boldsymbol{R}$ [41].

The main difference between finite rotations and infinitesimal rotations is that the first group is not commutative under the operation of multiplication of two rotations, because matrix multiplication is not commutative, while the second one is commutative under the operation of multiplication.

Given two infinitesimal rotations $\boldsymbol{\phi}_1$ and $\boldsymbol{\phi}_2$, their product is:

$$\boldsymbol{\phi}_1\boldsymbol{\phi}_2 = \mathbf{I} + d\boldsymbol{R}_1 + d\boldsymbol{R}_2\,, \tag{26}$$

neglecting higher order infinitesimals. The product in the reverse order interchanges only $d\boldsymbol{R}_1$ and $d\boldsymbol{R}_2$, which has no effect on the result, because matrix addition is always commutative [41].

An infinitesimal rotation has the inverse matrix $\boldsymbol{\phi}^{-1} = (\mathbf{I} - d\boldsymbol{R})$, and knowing that the property that for an orthogonal transformation the inverse matrix is identical with the transpose, we easily infer the property that

$$d\boldsymbol{R} = -d\boldsymbol{R}^{\intercal}\,, \tag{27}$$

which is the definition of an antisymmetric matrix.

An infinitesimal rotation transforms, under orthogonal transformations given by a matrix $\boldsymbol{B}$, as $\boldsymbol{\phi}' = \boldsymbol{B}\boldsymbol{\phi}\boldsymbol{B}^{\intercal}$ [41]. Orthogonal matrices are all those matrices such that $\boldsymbol{B}\boldsymbol{B}^{\intercal} = \boldsymbol{B}\boldsymbol{B}^{-1} = \mathbf{I}$, but with determinant equal to $\det(\boldsymbol{B}) = \pm 1$. The group of orthogonal transformations in three dimensions is called $O(3)$, and contains the spacial case of special orthogonal transformations .

$d\boldsymbol{R}$ is a skew-symmetric matrix[‡] composed by the three independent parameters specifying the rotation. An explicit form of $d\boldsymbol{R}$, without losing in generality, is:

$$d\boldsymbol{R} = \begin{pmatrix} 0 & d\alpha_3 & -d\alpha_2 \\ -d\alpha_3 & 0 & d\alpha_1 \\ d\alpha_2 & -d\alpha_1 & 0 \end{pmatrix}. \tag{28}$$

The three independent parameters specifying the rotation, i.e. $d\alpha_1$, $d\alpha_2$, $d\alpha_3$ can be collected in an array of elements, i.e. $d\vec{\alpha} = (d\alpha_1, d\alpha_2, d\alpha_3)$. This array of three elements, we claim, transforms as a vector if the transformation is a special orthogonal transformation.

[‡] A skew-symmetric matrix is a square matrix with diagonal elements are equal to zero and the non-vanishing off-diagonal elements are the negative of elements situated symmetrically across the diagonal.

If $d\vec{\alpha}$ is truly a vector, its components must transform under orthogonal matrix $\boldsymbol{B}$ according to the equations:

$$d\alpha'_i = B_{i\omega} d\alpha_\omega. \tag{29}$$

The quantity $d\vec{\alpha}$ has been introduced as an array of three independent parameters specifying the rotation, given as elements of an antisymmetric matrix. This does not assure that the array $d\vec{\alpha}$ transforms as a vector [41].

To show that we formally derive such transformation, following the derivation given by H. Goldstein in [41].

The quantities $d\alpha_\omega$ are related to the elements of $d\boldsymbol{R}$ by the equation:

$$d\alpha_\omega = \frac{1}{2}\epsilon_{\omega\mu\nu} dR_{\mu\nu}. \tag{30}$$

The quantity $\epsilon_{\omega\mu\nu}$ is known as Levi-Civita symbol. It is by definition equal to zero if any two of the indices $\omega\mu\nu$ are equal, and either $+1$ or -1 otherwise, according as $\omega\mu\nu$ is an even or odd permutation of 123. The Levi-Civita symbol is a tensor density whose components are the same in all coordinate systems [82].

A tensor is defined, respect to a square matrix, only in terms of its transformation properties under orthogonal coordinate transformations [41].

By definition, a tensor with upper indices μ, $\nu \ldots$ and lower indices κ, $\lambda \ldots$ transform like the product of contravariant vectors $U^\mu W^\nu \ldots$ and covariant vectors $V_\kappa V_\lambda \ldots$ [82].

A quantity that transforms as a tensor except for extra factors of the Jacobian determinant is called a tensor density [82].

To see that the Levi-Civita symbol is a tensor density we consider the following quantity:

$$\frac{\partial x'_\rho}{\partial x_\mu}\frac{\partial x'_\sigma}{\partial x_\nu}\frac{\partial x'_\eta}{\partial x_\lambda}\epsilon_{\mu\nu\lambda}. \tag{31}$$

This quantity is completely antisymmetric in the indexes ρ, σ, η and therefore proportional to $\epsilon_{\rho\sigma\eta}$ [82]. The equation (31) is the determinant of $|\partial x'/\partial x|$, and thus is easily to infer that :

$$\frac{\partial x'_\rho}{\partial x_\mu}\frac{\partial x'_\sigma}{\partial x_\nu}\frac{\partial x'_\eta}{\partial x_\lambda}\epsilon_{\mu\nu\lambda} = \left|\frac{\partial x'}{\partial x}\right|\epsilon_{\rho\sigma\eta}. \tag{32}$$

From the property that under orthogonal transformations $\boldsymbol{\phi}' = \boldsymbol{B}\boldsymbol{\phi}\boldsymbol{B}^\intercal$, we infer the transformation law of the infinitesimal operator $d\boldsymbol{R}$, i.e. $d\boldsymbol{R}' = \boldsymbol{B}d\boldsymbol{R}\boldsymbol{B}^\intercal$. Using the relation in (30), we can write:

$$d\alpha'_i = \frac{1}{2}\epsilon_{ijk} dR'_{jk} = \frac{1}{2}\epsilon_{ijk} B_{j\mu} dR_{\mu\nu} B_{k\nu}. \tag{33}$$

By making use of Levi-Civita tensor density properties, one can also express the skew-symmetric matrix $dR_{\mu\nu}$ in terms of $d\alpha_\omega$ as:

$$dR_{\mu\nu} = \epsilon_{\omega\mu\nu} d\alpha_\omega. \tag{34}$$

Plugging in (34) into (33), and using the identity $\frac{1}{2}\epsilon_{ijk}\epsilon_{\omega\mu\nu}B_{j\mu}B_{k\nu} = B_{i\omega}\det(\boldsymbol{B})$ [41], we obtain the transformation law for $d\vec{\alpha}$:

$$d\alpha'_i = \det(\boldsymbol{B})B_{i\omega}d\alpha_\omega. \tag{35}$$

The array $d\vec{\alpha}$, therefore, under orthogonal transformations transforms as a *pseudovector*. However, under special orthogonal transformations, i.e. when the $\det(\boldsymbol{B}) = 1$, $d\vec{\alpha}$ transforms as a vector. For proper rotations the equations (35) and (29) agree [41].

The property that $d\vec{\alpha}$ transforms as a vector, allows us to write:

$$d\vec{x}' = \vec{x}' \times d\vec{\alpha} = d\boldsymbol{R}\,\vec{x} = d\boldsymbol{R}\,\boldsymbol{R}^T\,\vec{x}'. \tag{36}$$

By definition of cross product and by the fact that $\vec{x}'$ is arbitrary, we find $d\alpha_i = \frac{1}{2}\epsilon_{ijk}R_{k\alpha}dR_{j\alpha}$. Finally one gets the important relation:

$$\frac{\partial R_{mn}}{\partial \alpha_k} = \epsilon_{mkl}R_{ln}. \tag{37}$$

The relation in (37) allows us to quantify the derivative of an infinitesimal rotation with respect to parameters of rotations' representation. Hence if f is any scalar function of orientation,

$$\frac{\partial f}{\partial \alpha_k} = \frac{\partial R_{j\omega}}{\partial \alpha_k}\frac{\partial f}{\partial R_{j\omega}} = \epsilon_{jki}R_{i\omega}\frac{\partial f}{\partial R_{j\omega}}. \tag{38}$$

Moreover, this approach can be also generalized for the whole space of possible orthogonal transformations, as made by J. M. Rallison in [174]. We do not present this generalization because out of the scope of this thesis.

Group theory for SO(3) and SU(2) group manifolds

Rotations are linear transformations which evolve continuously from the identity transformation [41]. As described in the previous section, infinitesimal rotations are special orthogonal transformations of coordinate axes in which the components of a vector are almost the same in both sets of axes, i.e. the change is infinitesimal [41]. In $\mathbb{R}^3$ the cross product describes the infinitesimal generators of rotations. These infinitesimal generators form the Lie algebra $\boldsymbol{so}(3)$ of the rotation group $SO(3)$. To define Lie algebra and generators we start by the definition of a Lie group.

A Lie group, by definition, is a group which is also a manifold [93]. To make sense of this definition, we should describe the concept of *group* and the concept of *manifold* and how they are related. Groups, as described by Olver [91], are an "*algebraic abstraction*" of the notion of symmetry. The rotation groups in two or three-dimensional space are the most famous examples in physics. Manifolds, which form the main objects in differential geometry, instead, are a generalization of concepts of curves and surfaces in space.

In general, a manifold locally looks like Euclidean space, but globally is quite different to the Euclidean space. The matching of this two different fields of mathematics synthesises a new theory that can be applied to many physical and mathematical problems [91].

Let's now recall the definition of groups and Lie groups given in literature: "*A group G is a set of elements g with the following properties:*

- *the set is closed under operation of multiplication. Multiplying any two elements g_1 and g_2 in G, their product $g_1 g_2$ is also an element of G. Multiplication is associative, so for any three element of $(g_1 g_2)g_3 = g_1(g_2 g_3)$;*
- *exists the identity element e of G such that $eg = ge = g$ for any element $g \in G$;*
- *for every element g, exists a unique inverse element g^{-1} such that $gg^{-1} = g^{-1}g = e$.*

Multiplication is not necessarily commutative. That is, $g_1 g_2$ is not necessarily the same of $g_2 g_1$. If $g_1 g_2 = g_2 g_1$ for every pair of elements then the group is said to be abelian, otherwise is nonabelian" [92].

A Lie group, according to Bishop et al., " *is a set which is a both a group and a manifold and such that the group operations are C^∞, that is, the maps:*

$$\begin{aligned} &G \times G \to G \;\; \textit{given by} \;\; (g,h) \to gh \\ &G \to G \;\; \textit{given by} \;\; g \to g^{-1} \;\; \textit{are} \; C^\infty \end{aligned} \tag{39}$$

" [93].

The Lie group G is characterised by the invariance of certain vector fields under the group multiplication. These invariant vector fields form a finite-dimensional vector space, called the Lie algebra of G, which is in a precise sense the infinitesimal generator of G. The whole information in the group is almost contained in its Lie algebra [92].

A Lie algebra, according to Bishop et al., "*is a vector space L for which is given a bilinear function from $L \times L$ to L, called bracket, and denoted by $[,]$, which satisfies:*

- $[\vec{x},\vec{x}] = 0$ *for every* $\vec{x} \in L$, *and that* $[\vec{x},\vec{y}] = -[\vec{y},\vec{x}]$ *for every* $\vec{x},\vec{y} \in L$;
- *the Jacobi identity: for every* $\vec{x},\vec{y},\vec{z} \in L$ $[\vec{x},[\vec{y},\vec{z}]] + [\vec{z},[\vec{x},\vec{y}]] + [\vec{y},[\vec{z},\vec{x}]] = 0$

" [93].

The Lie algebra is the cornerstone of the Lie group theory. It allows to replace complicated non-linear conditions of invariance by relatively simple linear infinitesimal conditions.

Here we recall the main properties of $SO(3)$ and $SU(2)$ group in their fundamental representation, respectively, the special orthogonal group and the special unitary group, which represent rotations in three dimensions.

The special orthogonal group $SO(3)$ is composed by 3×3 orthogonal matrices whose determinant are equal to 1. This group describes the rotation in three dimensional space. The special unitary group $SU(2)$ is composed by 2×2 unitary matrices whose determinant are equal 1. These two groups are homomorphic because the mapping between them is not one to one. Indeed the $SU(2)$ group covers the $SO(3)$ group twice, i.e. the map is $2:1$.

The Lie groups $SU(2)$ and $SO(3)$ have the same dimension. The dimension of a Lie group, by definition of Bade et al., can be obtained counting the number of real parameters needed to specify a member of the group. This dimension is, then, equal to the number of dimensions of the manifold of the group members [177].

The rotation group has dimension equal to 3 because three parameters are needed to specify a rotation. Two parameters are needed to pick an axis of rotation, while one parameter is needed to give the angle of rotation. The matrix group $SO(3)$ is three dimensional because a general 3×3 matrix has nine parameters, but the orthogonality and unit determinant condition together set six constraints. The matrix group $SU(2)$ is three dimensional because a general 2×2 unitary matrix can be described by 4 parameters and the determinant condition gives one constraint [41].

SO(3) group manifold

The $SO(3)$ group is the group of all rotations about the origin of the three-dimensional Euclidean space $\mathbb{R}^3$, under the operation of composition [178, 179]. $SO(3)$ additionally preserve the volume element and the Levi-Civita tensor density ϵ_{ijk}, also called the *constant structure* of the group. By definition, a transformation that preserve the origin, Euclidean distance and orientation is a rotation about the origin [41]. The composition of two rotation is another rotation and each rotation has a unique inverse. The identity map satisfies the definition of rotation.

The Lie group, according to Choquet-Bruhat et al. [81], $SO(3)$ is a *compact manifolds* of dimension equal 3, where a compact manifold is defined as a manifold that is compact as a topological space [81].

As described before, rotations are linear transformations in the Euclidean space $\mathbb{R}^3$ [81], and can be represented by matrices, once a basis of $\mathbb{R}^3$ has been chosen. When we choose an orthonormal basis of $\mathbb{R}^3$, every rotation is represented by an orthogonal 3×3 matrix with determinant equal to 1. This representation is called the *fundamental representation* [81].

As described in the previous section, it is often helpful considering a finite small rotation as a sequence of infinitesimal rotations. If it is so, we can build up a rotation through a finite angle α, as a succession N times through the angle α/N for N very large. For $N \to \infty$ and for a general rotation in three dimensions we obtain:

$$\boldsymbol{R}(\vec{\alpha}) = \lim_{N\to\infty} \left(\mathbf{I} + i \frac{\sum_k \alpha_k \boldsymbol{J}_k}{N} \right)^N = \mathrm{e}^{i \sum_k \alpha_k \boldsymbol{J}_k}, \tag{40}$$

where $\vec{\alpha}$, $\boldsymbol{J}_k$ are respectively the vector of $SO(3)$ group parameters (angles in our case), the generators of $SO(3)$. $\boldsymbol{J}_i$ are the generators of the rotations in three dimensions[§]:

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \; J_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}, \; J_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{41}$$

The sum in the exponential in equation (40), $i\sum_k \alpha_k \boldsymbol{J}_k$, is an element of the Lie algebra $\boldsymbol{so}(3)$.

The Lie algebra $\boldsymbol{so}(3)$ is defined as the tangent space of $SO(3)$ at the identity element e. The peculiarity of $\boldsymbol{so}(3)$ is that it is isomorphic to $\mathbb{R}^3$ [93].

The group is covered by the exponential map. Each element of $g \in SO(3)$, therefore, has the form as in equation (40) [81, 93]. The commutator between two generators of the Lie algebra $\boldsymbol{so}(3)$ is given by $[\boldsymbol{J}_k, \boldsymbol{J}_l] = i\epsilon_{klh} J_h$, therefore, the sum $i\sum_k \alpha_k \boldsymbol{J}_k = \epsilon_{k\mu\nu}\alpha_k$ is the same infinitesimal operator in equation (28).

SU(2) group manifold

The group $SU(2)$ is the special unitary group of degree 2. This is the group of 2×2 unitary matrices with determinant 1, i.e. unitary transformation in a 2-dimensional complex linear space. The dimension of the group is 3 as the number of the parameters and the generators of the group [180]. We define the set of parameters as $\vec{\alpha} = (\alpha_1, \alpha_2, \alpha_3)^\intercal$, where $^\intercal$ identifies the transpose, and the set of generator as $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)^\intercal$. The σ_i, where $i = 1, 2, 3$, are all Hermitian and unitary matrices and are called Pauli matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{42}$$

They square to one $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. They are traceless and satisfy the relations:

$$\sigma_i \sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k, \tag{43}$$

where the structure constant of the group is ϵ_{ijk} the Levi-Civita symbol. The commutator relations are:

$$[\sigma_i, \sigma_j] = \sigma_i\sigma_j - \sigma_j\sigma_i = 2i\epsilon_{ijk}\sigma_k. \tag{44}$$

Useful identities are: $\mathrm{Tr}(\sigma_i) = 0$, $\mathrm{Tr}(\sigma_i\sigma_j) = 2\delta_{ij}$, $\mathrm{Tr}([\sigma_i, \sigma_j]\sigma_k) = 4i\epsilon_{ijk}$.

For every vector $\vec{x} \in \mathbb{R}^3$ we identify an element x of the Lie algebra $\boldsymbol{su}(2)$ as follows:

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \leftrightarrow x = \sum_{i=1}^{3} x_i\sigma_i. \tag{45}$$

[§] The generators of the group $SO(3)$ has been multiplied by $-i$. It has been done to have the same notation for rotational operators along all the thesis.

Identified the element on the Lie algebra $\boldsymbol{su}(2)$ with the notation just specified above, the following relations are satisfied:

$$
\begin{aligned}
\vec{x} \times \vec{y} &= \frac{1}{2i}[x, y]\,, \\
\vec{x} \cdot \vec{y} &= \frac{1}{2}\mathrm{Tr}(xy)\,, \\
\vec{x} \times \vec{y} \cdot \vec{z} &= \frac{1}{4i}\mathrm{Tr}([x, y]z)\,.
\end{aligned}
\tag{46}
$$

Any element U of $SU(2)$ group can be expressed as:

$$U(\vec{\alpha}) = \mathrm{e}^{\left(\frac{i}{2}\sum_{j=1}^{3}\alpha_j\sigma_j\right)}\,, \tag{47}$$

where it is simple verify that $UU^{\dagger} = \mathbf{I}$ by the unitary, where † identifies the Hermitian conjugate. Indeed, using the property that the logarithm of the determinant of a matrix $\boldsymbol{A}$ it is equal to the trace of the logarithm of the matrix $\boldsymbol{A}$, we obtain:

$$\log\left(\det(U(\vec{\alpha}))\right) = \mathrm{Tr}(\log(U(\vec{\alpha}))) = \mathrm{Tr}(\frac{i}{2}\sum_{j=1}^{3}\alpha_j\sigma_j) = 0. \tag{48}$$

We can infer then that $\det(U(\vec{\alpha})) = 1$.

Using Taylor series we get that equation (47) can be rewritten as $U(\vec{\alpha}) = \cos(\frac{\alpha}{2})\mathbf{I} + i\sin(\frac{\alpha}{2})\vec{n}\cdot\boldsymbol{\sigma}$, where $\vec{n}$ is a unitary vector, which allows writing the array of parameter as $\vec{\alpha} = \alpha\vec{n}$, with α the angle of rotation. The existence of the factor $\frac{1}{2}$ in equation shows that for a rotation of 2π, we get $U = -\mathbf{I}$, whereas the rotation of a vector through an angle of 2π obviously leaves the vector unchanged. The mapping of $SU(2)$ in $SO(3)$ is not one to one. There are two elements of $SU(2)$ corresponding to each element of $SO(3)$. This mapping is called homomorphism.

Local coordinates on manifolds

In this section we introduce the concept of local coordinates on manifolds and how they are related to Lie groups.

A manifold with dimension m is:

"*a Hausdorff topological space such that every point has a neighbourhood homeomorphic to* $\mathbb{R}^m$*, i.e. an Euclidean space*" [93].

Formally, we can define a bijective map $\psi_i : U_i \to V_i$ such that for each open sets U_i on the manifold is associated only one open set $V_i \in \mathbb{R}^m$. The result of this simple operation is a definition of a local coordinate patch V_i of the open set U_i. In other words we are mapping a small piece of the curved space, the manifold, into a small piece of a flat space, the Euclidean space. The coordinates on this small piece of flat space are called local coordinates and are defined by the continuous map ψ_i.

It is also useful to define for a given point p on the manifold "*a change of coordinate*" between different local coordinate patches. To do that we imagine that the point p on the manifold is in the intersection of two open sets U_i an U_j, i.e. $U' = U_i \cap U_j$. Applying the bijective map defined above, we obtain two different open sets $V_i' = \psi_i(U')$ and $V_j' = \psi_j(U')$. To go from V_i' to V_j' we can therefore define two different bijective maps such that $\psi_{ij} = \psi_j \circ \psi_i^{-1} : V_i' \to V_j'$, $\psi_{ji} = \psi_i \circ \psi_j^{-1} = V_j' \to V_i'$ and $\psi_{ij} = \psi_{ji}^{-1}$. These maps have been built with the meaning of "*a change of coordinate on the manifold in local coordinate patches*" [48].

On a Lie group we can translate the local coordinates on manifold in the following way. Assume that we are close to the identity of the group, i.e. to the element e. We define the open set U_e on the manifold as the "*patch*" on the manifold close to the identity element of the group. By the left group action, it is possible building up patches $U_g = g(U_e)$ close to each point of the manifold. Using the bijective map defined before, local coordinates for patch U_g looks like $V_g = \psi_e[g^{-1}(U_g)]$.

From now on we will refer to the Lie group $SO(3)$, but the discussion can be immediately generalised to the Lie group $SU(2)$.

In local coordinate, it is possible write down the equation (40) as:

$$\boldsymbol{R}(\vec{\alpha}) = \boldsymbol{R}^* \boldsymbol{r}(\vec{\alpha}) = \boldsymbol{R}^* \mathrm{e}^{i \sum_i \alpha_i \boldsymbol{J}_i}, \tag{49}$$

where the $\boldsymbol{R}^*$ is a general rotation, and the $\boldsymbol{r}(\vec{\alpha})$ is a small rotation around $\boldsymbol{R}^*$. Equation (49) provides a one to one mapping between vectors in the open ball $\sqrt{\vec{\alpha}^\intercal \vec{\alpha}} < \pi$ and a local patch of the manifold $SO(3)$ centred around the point $\boldsymbol{R}^*$ [181]. The simple idea, here, is to use the fact that rotations are transformations which evolve continuously from the identity. This allows treating the dependence of the orientation on a local patch for a special orthogonal transformation as a continuously evolving rotation transformation from the identity.

The term $\boldsymbol{R}^*$, in equation (49), identifies a point on the manifold where a small rotation $\boldsymbol{r}(\vec{\alpha})$ on a local patch occurs. By conditions enumerated at the beginning of this appendix, the term $\boldsymbol{R}^*$ at time $t = 0$ identifies the identity transformation, i.e. the body frame is set to be coincident with the initial reference frame S. When the motion of the rigid body starts, we assume that the $\boldsymbol{R}^*$ is always updated to the last rotation [181], and we perform a small rotations $\boldsymbol{r}(\vec{\alpha})$, as a succession of infinitesimal rotation in the local patch. In other words we map the patch around the general $\boldsymbol{R}^*$ with a patch close to the identity. We have constructed at every point $\boldsymbol{R}^*$ on the manifold $SO(3)$, therefore, a continuous, invertible mapping between a neighbourhood of $\boldsymbol{R}^*$ on the manifold and an open set in $\mathbb{R}^3$ [181].

By the fact that the sum in the exponential in equation (49), $i \sum_k \alpha_k \boldsymbol{J}_k = d\boldsymbol{R}$, gives the same infinitesimal operator in equation (28), and by the observation that the elements of a skew-symmetric operator transform as a vector under proper rotation, we deduce that the open set in $\mathbb{R}^3$ defined by the patch around $\boldsymbol{R}^*$ is the Lie algebra $\boldsymbol{so}(3)$. The Lie algebra $\boldsymbol{so}(3)$ allows us to define the derivative with respect to parameters of the group manifold, which specify a small finite rotation on the local patch.

Numerically this operator reads:

$$\left[\boldsymbol{R}^{*}\frac{\partial}{\partial\vec{\alpha}}\mathrm{e}^{i\sum_{k}\alpha_{k}\boldsymbol{J}_{k}}\right]_{\vec{\alpha}=0} \tag{50}$$

where the derivatives are computed at $\vec{\alpha}=0$ because of infinitesimal rotation around the general rotation $\boldsymbol{R}^{*}$. In other words, we set the evolution of the rotation on the Lie algebra $\boldsymbol{so}(3)$, which is isomorphic to the Lie algebra $\mathbb{R}^{3}$ with cross product [93]. This Lie algebra is used to represent in a general way the orientational degrees of freedom by infinitesimal generators.

Appendix B

In this appendix we describe when the Itô or a Stratonovich stochastic integration are the same.

Let's start to write a Kramers-Langevin equation for a one dimensional translational Brownian motion with mass $m = 1$:

$$\begin{aligned} \frac{\partial x(t)}{\partial t} &= v(t), \\ \frac{\partial v(t)}{\partial t} &= a(x(t),t) + b(x(t),t)\xi(t), \end{aligned} \tag{51}$$

where $x(t)$ is the position of the Brownian particle, $v(t)$ the velocity, $a(x(t),t)$ the drift term and $b(x(t),t)$ the diffusion coefficient. $\xi(t)$ identifies the white noise, i.e. zero mean and correlation function equal to $\langle \xi(t)\xi(t')\rangle = \delta(t-t')$. For a sake of simplicity, we set $a(x(t),t) = 0$ and the equation (51) reads:

$$\begin{aligned} \mathrm{d}x(t) &= v(t)\mathrm{d}t, \\ \mathrm{d}v(t) &= b(x(t),t)\mathrm{d}W(t), \end{aligned} \tag{52}$$

where $W(t)$ is a Wiener process.

We define $\alpha \in [0,1]$ as parameter that shows if we are doing an integration by Itô (i.e. $\alpha = 0$) or by Stratonovich (i.e. $\alpha = 0.5$). The second equation in (52) can be written as:

$$\mathrm{d}v(t) = b(x(t) + \alpha \mathrm{d}x(t), t)\mathrm{d}W(t), \tag{53}$$

and expanding in Taylor's series one obtains:

$$\mathrm{d}v(t) = b(x(t),t)\mathrm{d}W(t) + \alpha b^{'}(x(t),t)\mathrm{d}x(t)\mathrm{d}W(t). \tag{54}$$

Replacing $\mathrm{d}x(t)$ of the equation (52) into the equation written above, one obtains that the term proportional to α is order $O(t^{\frac{3}{2}})$ and then negligible.

The first term in the r.h.s. of equation (54) is the term given by the Itô's integration, therefore in this case one does not have to specify what kind of integration is considered in the Langevin equation.

Whether the Langevin equation has the following form:

$$\frac{\partial s(t)}{\partial t} = a(s(t),t) + b(s(t),t)\xi(t), \tag{55}$$

where $s(t)$ is a degree of freedom of the system [28], one has to specify what kind of integration is considered.

My Contribution to the Papers

I In ***Paper I*** I performed the analytical investigation under the supervision of Erik Aurell and Ralf Eichhorn. I wrote the draft of the paper and played an active role in the discussion of the content.

II In collaboration with the co-authors I developed the analytical calculations for ***Paper II***. I performed the numerical analysis, implemented the code for simulations and analyzed the data. I also participated in the discussion for working out some analytical calculations and for drawing up the manuscript.

III ***Paper III*** is a follow-up of ***Paper II*** and was conducted principally by me. I formulated the research question of the project and performed the analytical calculations in collaboration with my supervisor Erik Aurell. I implemented the code for simulations and did the numerical analysis for the project. I made all the plots used in the paper. I wrote the manuscript in collaboration with the co-author and I followed the whole review process.

IV ***Paper IV*** is in collaboration with Ralf Eichhorn. I performed under his supervision the analytical investigation. I performed also the numerical analysis, implemented the code for the simulation and analyzed the results. I also played an active role in the discussion for drawing up the manuscript.

Bibliography

[1] Jorge Kurchan. "In and out of equilibrium". *Nature*, 433(7023):222–225, 2005.

[2] Ken Sekimoto. *"Stochastic Energetics"*. Springer, Berlin, Heidelberg, 2010.

[3] ResearchItaly. "Computational sciences: new algorithm to solve hard problems". `https://www.researchitaly.it/en/success-stories/computational-sciences-new-algorithm-to-solve-hard-problems/`, 2016.

[4] William G. Gray and Cass T. Miller. "Elements of thermodynamically constrained averaging theory". In *Introduction to the Thermodynamically Constrained Averaging Theory for Porous Medium Systems*, pages 1–36. Springer, 2014.

[5] N. G. Van Kampen. *"Stochastic Processes in Physics and Chemistry"*. North Holland, Amsterdam, 2007.

[6] E. K. H. Salje. "Multi-scaling and mesoscopic structures". *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 368(1914):1163–1174, 2010.

[7] Yoseph Imry. *"Introduction to Mesoscopic Physics"*. Oxford University Press, Oxford New York, 1997.

[8] J. K. G. Dhont. *"An introduction to dynamics of colloids"*. Elsevier, Amsterdam, 1996.

[9] P.E. Austrell and L. Kari. *"Constitutive Models for Rubber IV: Proceedings of the Fourth European Conference on Constitutive Models for Rubber, ECCMR 2005, Stockholm, Sweden, 27-29 June 2005, in 'On the constitution of filled rubber components and utilization in non linear finite element analysis' of M. Achenbach, J. Duarte"*. Taylor & Francis, 2005.

[10] Robert Zwanzig. *"Non equilibrium Statistical Mechanics"*. Oxford Univ. Press., Oxford, 2001.

[11] Scott A. Socolofsky and Gerhard H. Jirka. "Engineering–lectures", 2002.

[12] S.R. de Groot and P. Mazur. *"Non-Equilibrium Thermodynamics"*. Dover, New York, 1984.

[13] Uriel Frisch. *"Turbulence"*. Cambridge University Press, Cambridge, 1996.

[14] Mohammad Faghri and Bengt Sundén. *"Transport phenomena in fires"*, volume 20. WIT press, 2008.

[15] R. Huang, I. Chavez, K. M. Taute, B. Lukic, S. Jeney, M. Raizen, and E. Florin. "Direct observation of the full transition from ballistic to diffusive Brownian motion in a liquid". *Nat. Phys.*, 7:576–580, 2011.

[16] SV Patankar. *"Numerical Heat Transfer and Fluid Flow"*. McGraw Hill, 1980.

[17] Frank M White and Isla Corfield. *"Viscous fluid flow"*, volume 3. McGraw-Hill New York, 2006.

[18] A. Einstein. "Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen". *Annalen der Physik*, 322(8):549–560, 1905.

[19] M. von Smoluchowski. "Zur kinetischen theorie der brownschen molekularbewegung und der suspensionen". *Annalen der Physik*, 326(14):756–780, 1906.

[20] P. Langevin. "Sur la théorie du mouvement brownien". *Comptes Rendus de l'Academie des Sciences (Paris)*, 146:530–533, 1908.

[21] Robert M. Mazo. *"Brownian Motion: Fluctuations, Dynamics and Applications"*. Oxford University Press, Oxford, 2002.

[22] S. Chandrasekhar. "Stochastic problems in physics and astronomy". *Rev. Mod. Phys.*, 15:1–89, Jan 1943.

[23] Peter Reimann. "Brownian motors: noisy transport far from equilibrium". *Physics reports*, 361(2):57–265, 2002.

[24] Giorgio Parisi. "Brownian motion". *Nature*, 433(7023):221–221, 2005.

[25] I. Snook. *"The Langevin and generalised Langevin approach to the dynamics of atomic, polymeric and colloidal systems"*. Elsevier, Amsterdam, 2007.

[26] E. M. Purcell. "Life at low Reynolds number". *American Journal of Physics*, 45(1):3–11, Jan 1977.

[27] Phil Attard. *"Non-Equilibrium Thermodynamics and Statistical Mechanics"*. Oxford University Press, Oxford, 2012.

[28] Crispin Gardiner. *"Stochastic Method: A Handbook for the Natural and Social Sciences"*. Springer-Verlag Berlin Heidelberg, Berlin, 2009.

[29] C. Di Castro and R. Raimondi. *"Statistical Mechanics and Applications in Condensed Matter"*. Cambridge University Press, 2015.

[30] Peter J. Mucha and M. P. Brenner. "Diffusivities and front propagation in sedimentation". *Physics of Fluids*, 15(5):1305–1313, 2003.

[31] H. Brenner. "Coupling between the translational and rotational Brownian motions of rigid particles of arbitrary shape". *J. Colloid Sci.*, 20, 1955.

[32] Michael P. Brenner. "Screening mechanisms in sedimentation". *Physics of fluids*, 11(4):754–772, 1999.

[33] A. Einstein. *"Investigations on the Theory of the Brownian Movement"*. Dover Books on Physics Series. Dover Publications, 1956.

[34] P.J.W. Debye. *"Polar molecules"*. Dover books explaining science. Dover Publications, 1929.

[35] W. H. Furry. "Isotropic rotational Brownian motion". *Phys. Rev.*, 107:7–13, Jul 1957.

[36] L. Dale Favro. "Theory of the rotational Brownian motion of a free rigid body". *Phys. Rev.*, 119:53–62, Jul 1960.

[37] D. E. Woessner. "Nuclear spin relaxation in ellipsoids undergoing rotational Brownian motion". *The Journal of Chemical Physics*, 37(3):647–654, 1962.

[38] Paul S. Hubbard. "Rotational Brownian motion". *Phys. Rev. A*, 6:2421–2433, Dec 1972.

[39] K. A. Valiev and E. N. Ivanov. "Rotational Brownian motion". *Soviet Physics Uspekhi*, 16(1):1–16, 1973.

[40] Duane W Condiff and John S Dahler. "Brownian motion of polyatomic molecules: the coupling of rotational and translational motions". *The Journal of Chemical Physics*, 44(10):3988–4004, 1966.

[41] Herbert Goldstein. *"Classical mechanics"*. World student series. Addison-Wesley, Reading, Menlo Park, Amsterdam, 1980.

[42] Y. Han, A. M. Alsayed, M. Nobili, J. Zhang, T. C. Lubensky, and A. G. Yodh. "Brownian motion of an ellipsoid". *Science*, 314(5799):626–630, 2006.

[43] AL Schwab and JP Meijaard. "How to draw Euler angles and utilize Euler parameters". In *ASME 2006 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, pages 259–265. American Society of Mechanical Engineers, 2006.

[44] Evangelos A. Coutsias and Louis Romero. "The quaternions with an application to rigid body dynamics". Technical report, Sandia National Laboraties, 2004.

[45] Howard Brenner. "Taylor dispersion in systems of sedimenting nonspherical Brownian particles: II Homogeneous ellipsoidal particles". *J. Colloid. Interface Sci.*, 80, 1981.

[46] Raffaele Marino, Ralf Eichhorn, and Erik Aurell. "Entropy production of a Brownian ellipsoid in the overdamped limit". *Phys. Rev. E*, 93:012132, 2016.

[47] Yueheng Lan and Erik Aurell. "The stochastic thermodynamics of a rotating Brownian particle in a gradient flow". *Scientific Reports*, 5, 2015.

[48] Raffaele Marino and Erik Aurell. "Advective-diffusive motion on large scales from small-scale dynamics with an internal symmetry". *Phys. Rev. E*, 93:062147, Jun 2016.

[49] James Diebel. "Representing attitude: Euler angles, unit quaternions, and rotation vectors". *Matrix*, 58(15-16):1–35, 2006.

[50] A. Messiah. *"Quantum Mechanics"*. Number v. 2 in Quantum Mechanics. North-Holland, 1981.

[51] Eugene Wigner. *"Group theory: and its application to the quantum mechanics of atomic spectra"*, volume 5. Elsevier, 2012.

[52] A. N. Kolmogorov. "Zur umkehrbarkeit der statistischen naturgesetze". *Mathematische Annalen*, 113(1):766–772, 1937.

[53] A. N. Shiryayev. "On the reversibility of the statistical laws of nature". In *Selected Works of A. N. Kolmogorov*, pages 209–215. Springer, 1992.

[54] Martin Koeber and Uwe Schäfer. "The unique square root of a positive semidefinite matrix". *International Journal of Mathematical Education in Science and Technology*, 37(8):990–992, 2006.

[55] M. Doi and S. F. Edwards. *"The Theory of Polymer Dynamics"*. Clarendon Press, Oxford, 1986.

[56] Joseph R Lakowicz. "Fluorescence anisotropy". In *Principles of fluorescence spectroscopy*, pages 291–319. Springer, 1999.

[57] Gregorio Weber. "Rotational Brownian motion and polarization of the fluorescence of solutions". *Advances in protein chemistry*, 8:415–459, 1953.

[58] Andreas C Albrecht. "Polarizations and assignments of transitions: the method of photoselection". *Journal of Molecular Spectroscopy*, 6:84–108, 1961.

[59] D. Edwardes. "Steady motion of a viscous liquid in which an. ellipsoid is constrained to rotate about a principal axis". *Q. J. Math.*, 26:70–78, 1892.

[60] Glenn H Fredrickson. "Concentration depolarization of fluorescence in the presence of molecular rotation". *The Journal of chemical physics*, 88(9):5291–5299, 1988.

[61] M Ehrenberg and R Rigler. "Rotational Brownian motion and fluorescence intensify fluctuations". *Chemical Physics*, 4(3):390–401, 1974.

[62] R Rigler, Ülo Mets, J Widengren, and P Kask. "Fluorescence correlation spectroscopy with high count rate and low background: analysis of translational diffusion". *European Biophysics Journal*, 22(3):169–175, 1993.

[63] Stuart B. Dubin, Noel A. Clark, and George B. Benedek. "Measurement of the rotational diffusion coefficient of lysozyme by depolarized light scattering: configuration of lysozyme in solution". *The Journal of Chemical Physics*, 54(12):5158–5164, 1971.

[64] Bruce J. Berne and Robert Pecora. *"Dynamic light scattering: with applications to chemistry, biology, and physics".* Courier Corporation, 2000.

[65] Z. Cheng and T. G. Mason. "Rotational diffusion microrheology". *Phys. Rev. Lett.*, 90:018304, Jan 2003.

[66] K. Solc and W. H. Stockmayer. "Kinetics of diffusion-controlled reaction between chemically asymmetric molecules. i. general theory". *The Journal of Chemical Physics*, 54(7):2981–2988, 1971.

[67] Richard A. Cone. "Rotational diffusion of rhodopsin in the visual receptor membrane". *Nature*, 236(63):39–43, 1972.

[68] R. J. Cherry, A. Bürkli, M. Busslinger, and G. Schneider. "Rotational diffusion of band 3 proteins in the human erythrocyte membrane". *Nature*, 263:389–393, 1976.

[69] A. N. Kolmogorov. "über die analytischen methoden in der wahrscheinlichkeitsrechnung". *Mathematische Annalen*, 104(1):415–458, 1931.

[70] A. N. Kolmogorov. "Zur theorie der stetigen zufälligen prozesse". *Mathematische Annalen*, 108(1):149–160, 1933.

[71] A. N. Kolmogorov. "Zufallige bewegungen (zur theorie der brownschen bewegung)". *Annals of Mathematics*, pages 116–117, 1934.

[72] K. Itô. "Brownian motion on a Lie group". *Proc. Japan Acad.*, 26:4–10, 1950.

[73] K. Yosida. "Brownian motion in a homogeneous riemannian space". *Pacific J.*, 2:263–96, 1952.

[74] H. P. Jr. McKean. "Brownian motions on the 3-dimensional rotation group". *Mem. College Sci. Univ. Kyoto Ser. A Math.*, 33(1):25–38, 1960.

[75] H. P. Jr. McKean. *"Stochastic Integrals"*. Academic Press, New York and London, 1969.

[76] Elton P. Hsu. *"Stochastic Analysis on Manifolds"*. American Mathematican Society, Providence Rhode Island, 2002.

[77] Gregory S. Chirikjian. *"Stochastic Models, Information Theory, and Lie Groups, Volume 1"*. Birkähuser, Boston Basel Berlin, 2009.

[78] Gregory S. Chirikjian. *"Stochastic Models, Information Theory, and Lie Groups, Volume 2"*. Birkhäuser, Boston Basel Berlin, 2009.

[79] Kôsaku Yosida. "Brownian motion on the surface of the 3-sphere". *The Annals of Mathematical Statistics*, 20(2):292–296, 1949.

[80] Fabio Manca, Pierre-Michel Déjardin, and Stefano Giordano. "Statistical mechanics of holonomic systems as a Brownian motion on smooth manifolds". *Annalen der Physik*, 2015.

[81] Yvonne Choquet-Bruhat and Cecile DeWitt-Morette. *"Analysis, Manifolds and Physics, Part 1: Basics; Revised Edition"*. North Holland, 1996.

[82] Steven Weinberg. *"Gravitation and cosmology: principles and applications of the general theory of relativity"*. Wiley, New York, 1972.

[83] Peter Häunggi and Peter Jung. *"Colored Noise in Dynamical Systems"*, pages 239–326. John Wiley & Sons, Inc., 2007.

[84] Peter Hänggi and Harry Thomas. "Stochastic processes: Time evolution, symmetries and linear response". *Physics Reports*, 88(4):"207 – 319, 1982.

[85] Yu.L. Klimontovich. "Ito, stratonovich and kinetic forms of stochastic equations". *"Physica A: Statistical Mechanics and its Applications"*, 163(2):515 – 532, 1990.

[86] Gregory S Chirikjian. "group theory and biomolecular conformation: I. mathematical and computational models". *Journal of Physics: Condensed Matter*, 22(32):323103, 2010.

[87] R. Alterovitz, A. Lim, K. Goldberg, G. S. Chirikjian, and A. M. Okamura. "Steering flexible needles under markov motion uncertainty". In *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1570–1575, Aug 2005.

[88] Nilanjan Sarkar, Xiaoping Yun, and Vijay Kumar. "Control of mechanical systems with rolling constraints: Application to dynamic control of mobile robots". *The International Journal of Robotics Research*, 13(1):55–69, 1994.

[89] Robert J. Webster III, Jin Seob Kim, Noah J. Cowan, Gregory S. Chirikjian, and Allison M. Okamura. "Nonholonomic modeling of needle steering". *The International Journal of Robotics Research*, 25(5-6):509–525, 2006.

[90] Giuseppe Oriolo. *"Wheeled Robots"*, pages 1548–1554. Springer London, London, 2015.

[91] Peter J. Olver. *"Applications of Lie Groups to Differential Equations"*. Spring-Verlag, New York Berlin Heidelberg Tokyo, 1986.

[92] Claude Chevalley. *"Theory of Lie Groups I"*. Princeton University Press, Princeton, 1946.

[93] L. R. Bishop and R. J. Crittenden. *"Geometry of Manifolds"*. Academic Press, New York and London, 1964.

[94] Keith J. Devlin. *"The millennium problems: the seven greatest unsolved mathematical puzzles of our time"*. Basic Books, 2002.

[95] Grisha Perelman. "The entropy formula for the ricci flow and its geometric applications". *arXiv preprint math/0211159*, 2002.

[96] Grisha Perelman. "Ricci flow with surgery on three-manifolds". *arXiv preprint math/0303109*, 2003.

[97] Grisha Perelman. "Finite extinction time for the solutions to the ricci flow on certain three-manifolds". *arXiv preprint math/0307245*, 2003.

[98] Willy Feller. "Zur theorie der stochastischen prozesse". *Mathematische Annalen*, 113(1):113–160, 1937.

[99] E Joergensen. "Construction of the Brownian motion and the ornstein-uhlenbeck process in a riemannian manifold on basis of the gangolli-mc. kean injection scheme". *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 44(1):71–87, 1978.

[100] Ramesh Gangolli. "On the construction of certain diffusions on a differentiable manifold". *Probability Theory and Related Fields*, 2(5):406–419, 1964.

[101] Massimo Vergassola and Marco Avellaneda. "Scalar transport in compressible flow". *Physica D*, 106(1-2):148–166, 1997.

[102] Andrea Mazzino, Stefano Musacchio, and Angelo Vulpiani. "Multiple-scale analysis and renormalization for preasymptotic scalar transport". *Phys. Rev. E*, 71:011113, 2005.

[103] G. A. Pavliotis and A. M. Stuart. "Parameter estimation for multiscale diffusions". *Journal of Statistical Physics*, 127(4):741–781, 2007.

[104] C. M. Bender and S. A. Orszag. *"Advanced Mathematical Methods for Scientists and Engineers: Asymptotic Methods and Perturbation Theory"*. Springer Verlag, 1999.

[105] Grigoris Pavliotis and Andrew Stuart. *"Multiscale methods: averaging and homogenization"*. Springer, 2008.

[106] DW McLaughlin, GC Papanicolaou, and OR Pironneau. "Convection of microstructure and related problems". *SIAM Journal on Applied Mathematics*, 45(5):780–797, 1985.

[107] Erik Aurell, Stefano Bo, Marcelo Dias, Ralf Eichhorn, and Raffaele Marino. "Diffusion of a Brownian ellipsoid in a force field". *EPL (Europhysics Letters)*, 114(3):30005, 2016.

[108] Christian Van den Broeck et al. "Stochastic thermodynamics: a brief introduction". *Physics of Complex Colloids*, 184:155–193, 2013.

[109] Kerson Huang. *"Statistical Mechanics"*. John Wiley and Sons, 2012.

[110] K. Sekimoto. "Langevin equation and thermodynamics". *Prog. Theor. Phys. Suppl.*, 180:17, 1998.

[111] Eugene Wong and Moshe Zakai. "On the convergence of ordinary integrals to stochastic integrals". *The Annals of Mathematical Statistics*, 36(5):1560–1564, 1965.

[112] George E. Uhlenbeck and Leonard S. Ornstein. "On the theory of the Brownian motion". *Physical review*, 36(5):823, 1930.

[113] W. Moon and J.S. Wettlaufer. "On the interpretation of stratonovich calculus". *New Journal of Physics*, 16(5):055017, 2014.

[114] Steven E. Shreve. *"Stochastic calculus for finance II: Continuous-time models"*, volume 11. Springer Science & Business Media, 2004.

[115] Valentin Blickle and Clemens Bechinger. "Realization of a micrometre-sized stochastic heat engine". *Nature Physics*, 8(2):143–146, 2012.

[116] D. M. Carberry, J. C. Reid, G. M. Wang, E. M. Sevick, D. J. Searles, and D J Evans. "Fluctuations and irreversibility: an experimental demonstration of a second-law-like theorem using a colloidal particle held in an optical trap". *Phys. Rev. Lett.*, 92:140601, 2004.

[117] Haye Hinrichsen. "Entropy and information". In *Lecture notes DPG Summer school, Bad Honnef*, pages 47–94. DPG Summer school, Bad Honnef, 2015.

[118] Dénes Petz. "Entropy, von neumann and the von neumann entropy". In *John von Neumann and the foundations of quantum physics*, pages 83–96. Springer, 2001.

[119] Myron Tribus and Edward C. McIrvine. "Energy and information". *Scientific American*, 225(3):179–188, 1971.

[120] Radu Badescu. *"Equilibrium and Nonequilibrium Statistical Mechanics"*. John Wiley and Sons, 1975.

[121] Charles H Bennett. Demons, engines and the second law. *Scientific American*, 257(5):108–116, 1987.

[122] Harvey S Leff and Andrew F Rex. "resource letter md-1: Maxwell's demon". *American Journal of Physics*, 58(3):201–209, 1990.

[123] Koji Maruyama, Franco Nori, and Vlatko Vedral. "Colloquium: The physics of maxwell's demon and information". *Reviews of Modern Physics*, 81(1):1, 2009.

[124] Andrei N Kolmogorov. "On the shannon theory of information transmission in the case of continuous signals". *IRE Trans. Information Theory*, 2(4):102–108, 1956.

[125] A. N. Kolmogorov. "Theory of transmission of information". *Amer. Math. Soc. Translation, Ser.*, 2(33):291–321, 1963.

[126] D. J. Evans, E. G. D. Cohen, and G. P. Morriss. "Probability of second law violations in shearing steady flows". *Phys. Rev. Lett.*, 71:2401, 1993.

[127] G. Gallavotti and E. G. D. Cohen. "Dynamical ensembles in nonequilibrium statistical mechanics". *Phys. Rev. Lett.*, 94:2694, 1995.

[128] G. E. Crooks. "Nonequilibrium measurements of free energy differences for microscopically reversible markovian systems". *J. Stat. Phys.*, 90:1572, 1998.

[129] G. E. Crooks. "Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences". *Phys. Rev. E*, 60(3):2721, 1999.

[130] J. Kurchan. "Fluctuation theorem for stochastic dynamics". *J. Phys. A*, 31:3719, 1998.

[131] Jorge Kurchan. "Fluctuation theorem for stochastic dynamics". *Journal of Physics A: Mathematical and General*, 31(16):3719, 1998.

[132] J. L. Lebowitz and H. Spohn. "A Gallavotti-Cohen-type symmetry in the large deviation functional for stochastic dynamics". *J. Stat. Phys.*, 95:333, 1999.

[133] U. Seifert. "Entropy production along a stochastic trajectory and an integral fluctuation theorem". *Phys. Rev. Lett.*, 95:040602, 2005.

[134] Udo Seifert. "Stochastic thermodynamics, fluctuation theorems and molecular machines". *Reports on Progress in Physics*, 75(12):126001, 2012.

[135] Massimiliano Esposito, Upendra Harbola, and Shaul Mukamel. "Nonequilibrium fluctuations, fluctuation theorems, and counting statistics in quantum systems". *Rev. Mod. Phys.*, 81:1665–1702, 2009.

[136] E.M. Sevick, R. Prabhakar, Stephen R. Williams, , and Debra J. Searles. "Fluctuation theorems". *Annual Review of Physical Chemistry*, 59:603–633, 2008.

[137] Saar Rahav and Christopher Jarzynski. "Nonequilibrium fluctuation theorems from equilibrium fluctuations". *New Journal of Physics*, 15(12):125029, 2013.

[138] V. Y. Chernyak, M. Chertkov, and C. Jarzynski. "Path-integral analysis of fluctuation theorems for general Langevin processes". *J. Stat. Mech.*, 2006(08):P08001, 2006.

[139] Raphael Chetrite. "Fluctuation relations for diffusion that is thermally driven by a nonstationnary bath". *Phys. Rev. E*, 80:051107, 2009.

[140] R. Chetrite and K. Gawędzki. "Fluctuation relations for diffusion processes". *Commun. Math. Phys.*, 282:469–518, 2007.

[141] B. Cleuren, C. Van den Broeck, and R. Kawai. "Fluctuation and dissipation". *C. R. Physique*, 8:567, 2007.

[142] B. Cleuren, K. Willaert, A. Engel, and C. Van den Broeck. "Fluctuation theorem for entropy production during effusion of a relativistic ideal gas". *Phys. Rev. E*, 77:022103, 2008.

[143] G. E. Crooks and C. Jarzynski. "Work distribution for the adiabatic compression of a dilute and interacting classical gas". *Phys. Rev. E*, 75(2):021116, 2007.

[144] A. Imparato and L. Peliti. "Work and heat probability distributions in out-of-equilibrium systems". *C. R. Physique*, 8:556, 2007.

[145] A. Imparato, L. Peliti, G. Pesce, G. Rusciano, and A. Sasso. "Work and heat probability distribution of an optically driven Brownian particle: Theory and experiments". *Phys. Rev. E*, 76:050101, 2007.

[146] C. Jarzynski. "Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach". *Phys. Rev. E*, 56(5):5018–5035, 1997.

[147] Denis J. Evans and Debra J. Searles. "The fluctuation theorem". *Advances in Physics*, 51(7):1529–1585, 2002.

[148] Sergio Ciliberto, Nicolas Garnier, S Hernandez, Cédrick Lacpatia, J-F Pinton, and G Ruiz Chavarria. "Experimental test of the Gallavotti–Cohen fluctuation theorem in turbulent flows". *Physica A: Statistical Mechanics and its Applications*, 340(1):240–250, 2004.

[149] Delphine Collin, Felix Ritort, Christopher Jarzynski, Steven B. Smith, Ignacio Tinoco, and Carlos Bustamante. "Verification of the Crooks fluctuation theorem and recovery of rna folding free energies". *Nature*, 437(7056):231–234, 2005.

[150] G. M. Wang, Edith M. Sevick, Emil Mittag, Debra J. Searles, and Denis J. Evans. "Experimental demonstration of violations of the second law of thermodynamics for small systems and short time scales". *Physical Review Letters*, 89(5):050601, 2002.

[151] David Lacoste and Michael A Lomholt. "Stochastic thermodynamics of a tagged particle within a harmonic chain". *Physical Review E*, 91(2):022114, 2015.

[152] A. Puglisi, S. Pigolotti, Lamberto Rondoni, and A. Vulpiani. "Entropy production and coarse graining in markov processes". *Journal of Statistical Mechanics: Theory and Experiment*, 2010(05):P05015, 2010.

[153] G. N. Bochkov and Yu E. Kuzovlev. "Nonlinear fluctuation-dissipation relations and stochastic models in nonequilibrium thermodynamics: I. generalized fluctuation-dissipation theorem". *Physica A: Statistical Mechanics and its Applications*, 106(3):443–479, 1981.

[154] K. H. Andersen, P. Castiglione, A. Mazzino, and A. Vulpiani. "Simple stochastic models showing strong anomalous diffusion". *The European Physical Journal B-Condensed Matter and Complex Systems*, 18(3):447–452, 2000.

[155] G. Gradenigo, A. Puglisi, A. Sarracino, Angelo Vulpiani, and D. Villamaina. "The out of equilibrium response function in sub-diffusive systems". *Physica Scripta*, 86(5):058516, 2012.

[156] Shinichi Sasa. "Collective dynamics from stochastic thermodynamics". *New Journal of Physics*, 17(4):045024, 2015.

[157] Christian Van den Broeck, Shinichi Sasa, and Udo Seifert. "Focus on stochastic thermodynamics". *New Journal of Physics*, 18(2):020401, 2016.

[158] Sebastian Goldt and Udo Seifert. "Stochastic thermodynamics of learning". *Physical Review Letters*, 118(1):010601, 2017.

[159] Udo Seifert. "First and second law of thermodynamics at strong coupling". *Physical review letters*, 116(2):020601, 2016.

[160] Christopher Jarzynski. "Stochastic and macroscopic thermodynamics of strongly coupled systems". *Physical Review X*, 7(1):011008, 2017.

[161] Anna Alemany and Felix Ritort. "Fluctuation theorems in small systems: extending thermodynamics to the nanoscale". *Europhysics News*, 41(2):27–30, 2010.

[162] C. Jarzynski. "Comparison of far-from-equilibrium work relations". *C. R. Physique*, 8:495, 2007.

[163] C. Jarzynski. "Nonequilibrium equality for free energy differences". *Phys. Rev. Lett.*, 78:2690, 1997.

[164] Christopher Jarzynski. "Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale". *Annu. Rev. Condens. Matter Phys.*, 2, 2011.

[165] Udo Seifert. "Stochastic thermodynamics: principles and perspectives". *The European Physical Journal B-Condensed Matter and Complex Systems*, 64(3):423–431, 2008.

[166] Thomas Speck, Jakob Mehl, and Udo Seifert. "Role of external flow and frame invariance in stochastic thermodynamics". *Phys. Rev. Lett.*, 100:178302, 2008.

[167] S. Kullback and R. A. Leibler. "On information and sufficiency". *Ann. Math. Statist.*, 22(1):79–86, 03 1951.

[168] NG Van Kampen. "The definition of entropy in non-equilibrium states". *Physica*, 25(7-12):1294–1302, 1959.

[169] Edwin T. Jaynes. "Information theory and statistical mechanics". *Physical review*, 106(4):620, 1957.

[170] Edwin T. Jaynes. "Information theory and statistical mechanics. ii". *Physical review*, 108(2):171, 1957.

[171] Antonio Celani, Stefano Bo, Ralf Eichhorn, and Erik Aurell. "Anomalous thermodynamics at the microscale". *Phys. Rev. Lett.*, 109:260603, 2012.

[172] Stefano Bo and Antonio Celani. "Multiple-scale stochastic processes: decimation, averaging and beyond". *Physics Reports*, 2016.

[173] Howard Brenner. "Coupling between the translational and rotational Brownian motions of rigid particles of arbitrary shape: II General theory". *J. Colloid. Interface Sci.*, 23, 1967.

[174] JM Rallison. "The effects of Brownian rotations in a dilute suspension of rigid particles of arbitrary shape". *Journal of Fluid Mechanics*, 84(02):237–263, 1978.

[175] Lev Davidovič Landau and Evgenii Mikhailovich Lifshitz. *"Mechanics and electrodynamics"*. Elsevier, 2013.

[176] Vladimir I. Arnold. *"Mathematical Methods of Classical Mechanics, 2nd Ed."*, volume 60. Springer Verlag, 1989.

[177] W. L. Bade and Herbert Jehle. "An introduction to spinors". *Rev. Mod. Phys.*, 25:714–728, 1953.

[178] John Stuelpnagel. "On the parameterization of the three-dimensional rotation group". *SIAM Review*, 6(4), 1964.

[179] Morton Curtis. *"Matrix Group"*. Spinger-Verlag, New York, 1979.

[180] Ian D. Lawrie. *"A Unified Grand Tour of Theoretical Physics Third Edition"*. CRC Press, Boca Raton London New York, 2012.

[181] Camillo J. Taylor and David J. Kriegman. "Minimization on the lie group SO(3) and related manifolds". Technical Report 9405, Center for Systems Science, Dept. of Electrical Engineering, Yale University, New Haven, CT, 1994.

Papers