Stochastic Dynamics of Reacting Biomolecules


World Scientific Singapore 2002
Chapter 2

Tools of Stochastic Dynamics

*L. Schimansky-Geier and P. Talkner*

2.1 Introduction

For most topics treated in this book it is of fundamental importance to understand the motion of classical particles in complex potential landscapes and, in particular, the escape of particles from locally stable states. An escape process only is possible if the particle can exchange energy with its environment in order to trespass intermediate states with high potential energy. These events can be treated by the help of various mathematical and computational tools some of which will be introduced in this section.

Molecular dynamics provides a most powerful approach to gain information about the dynamical properties of large assemblies of particles with known interactions. The primary objects of this methodology are the trajectories of the considered system, say a macromolecule swimming in water. The mutual interactions of the macromolecule and the water molecules require to simulate the time evolution of the full system. From the resulting huge amount of detailed information about the state of each single molecule the interesting properties of the macromolecule, as for example its diffusional properties, then have still to be extracted.

A less detailed description that also captures the relevant aspects of the problem under study can often be achieved by means of the theory of stochastic processes. There, the uninteresting part of the problem, the dynamics of the water molecules in the above example, is modeled in some crude manner that correctly reflects the statistical behavior of many collisions between the macromolecule and water molecules but not the detailed process of a single molecular encounter. In the mentioned example the force of the water molecules on the macromolecule is split into an average friction force and a fluctuating force. For small velocities of the macromolecule relative to the average motion of the water molecules the Stokes’ law describes the friction force. The fluctuating force then is often modeled by Gaussian white noise the strength of which is related to the temperature of the water and to the friction constant according to Einstein’s formula.

The simplifications of the stochastic method compared to molecular dynamics
pay off in a smaller computational effort needed to investigate the relevant part of
the system. In a simulation only a few stochastic differential equations describing
the dynamics of the relevant degrees of freedom have to be solved in contrast to a
huge number of Newton’s equations of motion for all microscopic degrees of freedom.

Both methods rely on statistical mechanics: In molecular dynamics the initial
conditions usually are drawn from a distribution of points in phase space describ-
ing the system in a particular thermodynamic state. In the stochastic framework
the statistical properties of the random forces and their relations to transport co-
efficients are dictated by relations from statistical mechanics, as for example the
Einstein relation in the case of a diffusing macromolecule. In general, the final re-
results of both methods are averaged quantities such as mean values of macroscopic
quantities and correlations of fluctuations about the mean values and their Fourier
transforms giving the according power spectra.

In the stochastic framework, the most probable fluctuations about the mean
values of macroscopic quantities turn out to be very small. However, there are
important exceptions from this rule where large deviations play a crucial role. The
escape from a metastable state is a typical example. In this process the system has
to visit remote regions in phase space where it has a much higher energy than in
its initial and final states and where therefore it is found only rarely.

This chapter gives an introduction to stochastic methods and their underpin-
nings by statistical mechanics (excellent enlarged presentations can be found, for
We will always deal with a few relevant degrees of freedom which we will refer to
as velocities and coordinates of a single or a few interacting particles which are in
contact with a thermal bath. We will study their dynamics in equilibrium and their
relaxation to equilibrium. For a particle with metastable states the waiting times
in these states and the transition rates between these states will be of particular
interest.

This chapter is a tutorial into stochastic dynamics. There are a large num-
ber of exciting monographs, reviews and collections of articles which show the
vitality of applications of stochastic methods in physics. We have listed only a
smaller number in the reference list (Stratonovich, 1963,1967; Kuznetsov et al.,
1965; Tikhonov, Mironov, 1979; Gardiner, 1982; Horsthemke, Lefever, 1983; Risken
1984; Malchow, Schimansky-Geier, 1985; Montroll, Lebowitz, 1987; Moss, McClint-
tock, 1990; van Kampen, 1992; Hänggi Jung, 1995; Hänggi, Talkner, 1995; Kogan,
1996; Schimansky-Geier, Pöschel, 1997; Anishchenko et al., 1999; Garcia-Ojalvo,
Sancho,1999; Freund, Pöschel, 2000; San Miguel, Toral, 2000; Schimansky-Geier
et al., 2000; Reimann, 2002; Anishchenko et al, 2002). Readers interested in
detailed mathematical foundations and background are refereed to mathematical
2.2 Fluctuations in statistical physics

2.2.1 The canonical distribution

We will deal with a system being in thermal equilibrium with its bath and assume that the combined system plus bath is isolated. The total system then possesses a constant energy $E_0 = H_s + H_b$ and the volume $V_0 = V_s + V_b$. The subscripts $b$ and $s$ stand for the bath and the system, respectively. Here we have neglected the contribution of the interaction between the system and the bath to the total energy. It indeed is small if both the system and the bath are macroscopic and the interaction is short ranged. Microstates of the bath and the system are labelled as $\Gamma_b$ and $\Gamma_s$ referring to points in the respective phase spaces.

For an isolated system the equilibrium probability to find the system and the bath in an infinitesimal volume of size $d\Gamma_b d\Gamma_s$ at the phase space point $(\Gamma_s, \Gamma_b)$ has the important property that it maximizes the entropy. As a consequence, the resulting microcanonical distribution is given by the equipartition of the microstates at the given energy:

$$d\rho = \text{constant} \delta(H_s(\Gamma_s) + H_b(\Gamma_b) - E_0) d\Gamma_s d\Gamma_b.$$  \hfill (2.1)

If the phase space volume is measured in units of Planck’s constant taken to the power of the total number $N$ of degrees of freedom, the normalization constant coincides with the inverse number of quantum states per unit energy of the total system at the energy $E_0$ and volume $V_0$:

$$\text{constant} = \frac{1}{\Omega_0(E_0, V_0)}.$$  \hfill (2.2)

Multiplying the density of states by an energy $\delta E$ which is small on the macroscopic scale we obtain the number of states $W_0(E_0, V_0) = \Omega_0(E_0, V_0)\delta E$ in the phase space shell of width $\delta E$ at the energy $E_0$. With Boltzmann’s formula this number can be expressed in terms of the thermodynamic entropy of the bath and system

$$S_0(E_0, V_0) = k_B \ln W_0(E_0, V_0)$$  \hfill (2.3)

where $k_B$ denotes the Boltzmann-constant which is the fundamental constant of statistical mechanics of equilibrium systems. It turns out that the actual value of the width of the energy shell has practically no influence on the entropy of a macroscopic system and therefore presents no real arbitrariness in the fundamental relation between the density of states and the entropy.

Our next goal is to determine the probability with which one finds the system in the volume $V_s$ at the energy $E_s$ if the isolated total system plus bath has the energy $E_0$ and the volume $V_0$. This probability is given by the sum of all those phase points weighted by the microcanonical distribution that fulfill the required
condition that the system has the energy $E$

\[ P(E, V_s) = \int \delta(H_s - E) d\rho \]

\[ = \Omega_s(E, V_s) \frac{\Omega_b(E_0 - E, V_0 - V_s)}{\Omega_0(E_0, V_0)} \]  \hspace{1cm} (2.4)

Here we used the microcanonical distribution and introduced the density of states of the isolated bath and system at given respective energies and volumes:

\[ \Omega_s(E, V) = \int \delta(H_s - E) d\Gamma_s \]

\[ \Omega_b(E, V) = \int \delta(H_b - E) d\Gamma_b \]  \hspace{1cm} (2.5)

The first factor in the second line of eq. (2.4) gives the number of states of the isolated system at energy $E$. Therefore, the second term represents the probability $p(E, V_s)$ to find the system at energy $E$. Multiplying the denominator and numerator of this factor by the same energy $E$ we obtain the ratio of numbers of states in the bath and in the total system at their energies and volumes. According to eq. (2.3) these numbers can be expressed in terms of the respective entropies:

\[ P(E, V_s) = \exp\left(\frac{1}{k_B} \Delta S_0(E, V_s)\right) \]  \hspace{1cm} (2.6)

where $\Delta S_0(E, V_s)$ denotes the entropy difference of the bath and the total system:

\[ \Delta S_0(E, V_s) = S_b(E_0 - E, V_b) - S_0(E_0, V_0) \]  \hspace{1cm} (2.7)

For a macroscopic bath the entropy is again related to the number of states $W_b(E, V) = \Omega_b(E, V) \delta E_b$ in the form of eq. (2.3):

\[ S_b(E_0 - E, V_0 - V_s) = k_B \ln W_b(E_0 - E, V_0 - V_s)) \]  \hspace{1cm} (2.8)

Introducing the free energies of the total system and the bath

\[ S_0(E_0, V_0) = \frac{1}{T} (E_0 - F_0(T, V_0)) ; \]

\[ S_b(E_0 - E, V_b) = \frac{1}{T} (E_0 - E - F_b(T, V_b)) \]

we find for the entropy difference

\[ \Delta S_0 = \frac{1}{T} (F_s(T, V_s) - E) \]  \hspace{1cm} (2.10)

where $F_s = F_0 - F_b$ is the free energy of the system. This result combined with eq. (2.4) yields the canonical distribution of a system in thermal contact with a bath of temperature $T$

\[ P(E, V_s) = \Omega_s(E, V_s) \exp\left(\frac{1}{k_B T} (F_s(T, V_s) - E)\right) \]  \hspace{1cm} (2.11)
From normalization one obtains the fundamental relation

$$F_s(T, V_s) = -k_B T \ln \int \Omega_s(E, V_s) \exp\left(\frac{-E}{k_B T}\right) dE,$$  \hspace{1cm} (2.12)

that expresses the free energy of the systems in terms of the bath temperature and the density of states of the isolated system.

We note that the distribution is determined by the difference of the free energy and its value in the constrained case, $F_s(T, V_s) - E$. Below, we will recover this particular structure for the distributions of more general observables.

### 2.2.2 Einstein’s formula

In 1905 Albert Einstein derived the probability distribution for arbitrary observables of a thermodynamic system. In thermodynamic equilibrium an observable which we denote by $x$ in general will fluctuate. A simple example is given by the end to end distance of a linear polymer chain with $N$ bonds immersed in a fluid. Another example is the position of an adiabatic piston between two gases at equal temperature. Both quantities fluctuate and are characterized by probability distributions.

For the sake of simplicity we will assume that $x$ can be expressed as a phase space function of the microscopic variables $\Gamma_s$

$$x = x(\Gamma_s).$$ \hspace{1cm} (2.13)

Further we restrict ourselves to isolated thermodynamic systems, i.e. $E_s, N_s, V_s$ are constants. So we look for the probability distribution $P(x) dx$ of the fluctuating value $x$ in the interval $[x, x + dx]$ in this isoenergetic case. Because the condition $x = const$ represents a cross-section at the hypersurface $E_s = const$ the probability to find the system at $x$ will be proportional to the number of states that lie in this cross-section.

Einstein solved this problem by using the statistical definition of the thermodynamic entropy. Under the assumption of a conditioned equilibrium he introduced the conditional entropy $S(x|E_s, N_s, V_s)$ defined by the number of microscopic states $\Omega(x|E_s, N_s, V_s)$ that realize this certain value of $x$. One again uses Boltzmann’s formula

$$S(x|E_s, N_s, V_s) = k_B \ln \Omega(x|E_s, N_s, V_s).$$ \hspace{1cm} (2.14)

The probability distribution density of finding $x$ at a certain value follows from its geometric definition

$$P(x|E_s, V_s, N_s) = \frac{\Omega(x|E_s, N_s, V_s)}{\Omega(E, N, V)}$$ \hspace{1cm} (2.15)
or, respectively, according to (2.13)

\[ P(x) = \exp \left( \frac{1}{k_B} \left( S(x|E_s, N_s, V_s) - S(E_s, N_s, V_s) \right) \right) \]  

(2.16)

where \( S(E_s, N_s, V_s) \) is the unconstrained thermodynamic entropy. Again the probability of the conditioned state is expressed by the reduction of entropy caused by the fixation of \( x \).

By virtue of eq.(2.15) one easily proves that normalization is obeyed. As a result one need not calculate the full thermodynamic entropy \( S(E_s, N_s, V_s) \) which sometimes may be difficult. Obviously, it holds

\[ P(x) = \frac{\exp \left( \frac{1}{k_B} S(x) \right)}{\int dx \exp \left( \frac{1}{k_B} S(x) \right)} \]  

(2.17)

where we write \( S(x|E_s, N_s, V_s) = S(x) \) for short. Only that part of entropy enters which functionally depends on the considered degree of freedom whereas constant distributions drop due to normalization.

### 2.2.3 Fluctuations around equilibrium

Let us list some general properties of the probability distribution near the equilibrium value \( x_{eq} \). First of all, at its equilibrium value \( x_{eq} \) thermodynamic potentials have extremal properties with respect to \( x \). As a statement of the second law in the isoenergetic case the entropy has a maximum there, \( S(x) \rightarrow S_{\text{max}} = S(x_{eq}) \) if \( x_{eq} \) is reached.

Therefore, in the neighborhood of a stable equilibrium state the entropy can be expanded as

\[ S(x) = S(x_{eq}) - \frac{1}{2} k_B g (x - x_{eq})^2 \]  

(2.18)

with

\[ g = - \frac{1}{k_B} \left( \frac{\partial^2 S}{\partial x^2} \right)_{x=x_{eq}} > 0 . \]  

(2.19)

If higher terms in the expansion can be neglected it follows that fluctuations around the equilibrium state are well approximated by Gaussian distributions. If the majority of probability is in the neighborhood of \( x_{eq} \) one may extend the range of the Gaussian till infinite values. The normalized probability density to find the state \( x \) is given by

\[ P(x) = \sqrt{\frac{g}{2\pi}} \exp \left[ -\frac{1}{2} g (x - x_{eq})^2 \right] . \]  

(2.20)
The standard deviation is determined by the second derivative of the entropy
\[ \langle (x - x_{eq})^2 \rangle = \frac{1}{g} . \]  

(2.21)

In case of thermodynamic variables, i.e. for \( x \) being temperature, pressure, volume or particle numbers, the second derivative of entropy is a susceptibility like the heat capacity, compressibility or derivations of the chemical potentials. These susceptibilities describe the elasticity of thermodynamic systems in response to applied changes of the system and define the stability of thermodynamic states. According to (2.21) these functions determine the width of the probability distributions around equilibrium.

For a set of \( n \) fluctuating variables \( \vec{x} = x_1 \ldots x_n \) \((n \ll N_L)\) the resulting Gaussian distribution near the equilibrium state \( \vec{x}_{eq} \) reads
\[ P(\vec{x}) = \sqrt{\frac{\det(g)}{(2\pi)^n}} \exp \left[ -\frac{1}{2} \sum_{i,j} g_{i,j} (x_i - x_{i,eq}) (x_j - x_{j,eq}) \right] . \]  

(2.22)

Therein
\[ g_{i,j} = -\frac{1}{k_B} \left( \frac{\partial^2 S}{\partial x_i \partial x_j} \right)_{\vec{x}=\vec{x}_{eq}} \]  

(2.23)

and the second moments are given by the inverse matrix
\[ \langle (x_i - x_{i,eq})(x_j - x_{j,eq}) \rangle = g_{i,j}^{-1} . \]  

(2.24)

Analogously, for other thermodynamic constraints (for example, isothermic or isobaric closed systems) one obtains Gaussian distributions where \( g \) is the positive second derivative of the corresponding thermodynamic potentials, \( F(x|T_s, V_s, N_s) \) or \( G(x|T_s, P_s, N_s) \).

### 2.2.4 Perrin’s pendulum

As an example we discuss the motion of a sensitive torsion pendulum in a gas at temperature \( T \). The gas together with the pendulum form an isolated thermal system with constant overall energy \( E_0 \). The single degree of freedom describing the angular motion of the pendulum is the angle \( \phi \). The surrounding gas with a finite number of particles permanently hits on the pendulum causing fluctuations of the pendulum around a rest position \( \phi_{eq} \) at which the entropy is maximal with respect to \( \phi \).

What then is the probability to find deviations from the equilibrated rest position \( \phi_{eq} \)? For simplicity we set \( \phi_{eq} = 0 \).

The advantage of Einstein’s approach consists in the calculation of the entropy reduction \( \Delta S \) based on thermodynamic balance equations. Hence, one may use the
The Gibbsian fundamental equation is given by:

$$dS(E, x_1, \ldots, x_n) = \frac{1}{T} \left( dE - \sum_i x_i dx_i \right)$$

with $x_1 = \phi$. In Fig. (2.2.4) the unconstrained entropy $S(E)$ is plotted with respect to energy $E$. Deviations from the equilibrium state of $\phi$ decreases the entropy along $E = \text{const}$ which is an irreversible process (see vertical line in Fig. (2.2.4)). $S(x, E)$ can be found from the concept of the reversible “Ersatzprozeß”. It means that we calculate the entropy shift performing reversible work with $S = \text{const}$ which is called the minimal work (horizontal line in Fig. (2.2.4)). Applied to the pendulum the internal energy changes as

$$dE(\phi) = -dR_{\text{min}} = -D \phi d\phi$$

with $D$ being the modulo of torsion. For small deviations (see Fig. 2.2.4) the needed entropy shift relates as

$$dS(x) = \left( \frac{\partial S}{\partial E} \right)_V dE$$

and one gets

$$S(\phi) = S(0) - \frac{1}{2T} D \phi^2.$$ 

Therefore we find

$$P(\phi) = \sqrt{\frac{D}{2\pi k_B T}} \exp \left[ -\frac{D}{2k_B T} \phi^2 \right].$$

Perrin measured in experiments the standard deviation

$$\langle \phi^2 \rangle = \frac{k_B T}{D}.$$
and therefrom he found Boltzmann’s constant \( k_B \). From this result he was able to find Avargado’s number by

\[ N_A = \frac{R}{k_B} \]  

(2.31)

with \( R \) being the universal gas constant. It was one of the first findings where the consideration of fluctuations gave a strong support to prove the molecular structure of matter.

### 2.2.5 General approach

Here, we give a more formal derivation of the probability density of a fluctuating variable \( x = x(\Gamma_s) \). Using Dirac’s \( \delta \)-function one obtains

\[ P(x) = \int \delta (x - x(\Gamma_s)) \rho(\Gamma_s) \, d\Gamma_s. \]  

(2.32)

Therein \( \rho(\Gamma_s) \) is the probability of a microstate in a certain thermodynamic contact. The integration of \( \rho \) multiplied with the \( \delta \)-function just collects the probability of those microstates which realize the given value \( x \) for the function \( x(\Gamma_s) \).

In case of an isolated system the conditional entropy (2.14) then reads

\[ S(x|E, N, V) = k_B \ln \int \delta (x - x(\Gamma_s)) \delta (H(\Gamma_s) - E) \, d\Gamma_s. \]  

(2.33)

The probability distribution follows accordingly

\[ P(x) = \exp \left( -\frac{S(E, V, N)}{k_B} \right) \int \delta (x - x(\Gamma_s)) \delta (H(\Gamma_s) - E) \, d\Gamma_s. \]  

(2.34)

Immediately one finds the distribution (2.16).

Formally this approach can also be extended to other distributions describing different thermodynamic constraints. For example if \( x \) is a fluctuating value inside an isothermic-isochoric many-particle system the projection procedure yields

\[ P(x|T, V) = \exp \left( \frac{F(T, V)}{k_B T} - \frac{F(x|T, V)}{k_B T} \right) \]  

(2.35)

where

\[ F(x|T, V) = -k_B T \ln \int \delta (x - x(\Gamma_s)) \exp \left( -\frac{H(\Gamma_s)}{k_B T} \right) \, d\Gamma_s \]  

(2.36)

is the conditional free energy of the isothermic-isochoric system for a fixed value of \( x \).


2.3 Linear relaxation processes

In statistical physics macroscopic values $x$ of a thermodynamic system permanently fluctuate around their equilibrium state with maximal entropy. The probability distribution for these fluctuations was given in the last chapter. The actual sample path of $x(t)$ most of the time visits nonequilibrium states.

In this section we are concerned with the time evolution of the mean value of sample paths brought out of equilibrium by an initial impact at $t_0$. That might be realized by a thermal fluctuation as well as by an external perturbation. We will look at the relaxation of the mean $\langle x(t) \rangle_{x(t_0)=x_0}$ conditioned that it exactly starts at $x_0$ at $t = t_0 = 0$.

From the 2nd law of thermodynamics one knows that the equilibrium state is a stable attractor of the dynamics of the mean value. Entropy is a Lyapunov function with respect to perturbations around $x_{eq}$. The mean value as function of time obeys a dissipative law.

Taking the entropy as function of the actual value and derivating with respect to time gives

$$\dot{S}(x) = \frac{\partial S}{\partial x} \dot{x} \geq 0 \ .$$

(2.37)

Two new important values of non-equilibrium thermodynamics appear in this equation,

$$X = \frac{\partial S}{\partial \langle x \rangle}$$

(2.38)

is called in analogy to mechanics the thermodynamic force, where the entropy plays the role of a mechanical potential. The second term on the right hand side (2.37)

$$J = \dot{x}$$

(2.39)

is the thermodynamic flux.

Lars Onsager gave a general connection between both values. Referring to the circumstance that in dissipative systems there is no flux without a force he concluded that near to equilibrium the linear connection

$$J = L X$$

(2.40)

must hold with $L$ being a constant. Obviously, the 2nd law (2.37) requires that $L$ is strictly positive, i.e.,

$$\dot{S}(x) = L X^2 \geq 0 \ .$$

(2.41)

This first of Onsager’s hypothesizes about the linear connection between thermodynamic forces and fluxes led to the formulation of the thermodynamics of linear dissipative system. The kinetic coefficients $L$ are called Onsager-coefficients.
Linear relaxation processes

In the neighborhood of the equilibrium state it holds with the approximated entropy from eq. (2.18)

\[ X = \frac{\partial S}{\partial x} = -k_B g x. \]  
(2.42)

With (2.39) and (2.40) we find the linear relaxational dynamics for non-equilibrium states

\[ \dot{x} = -L k_B g x = -\lambda x \]  
(2.43)

where

\[ \lambda = L k_B g \]  
(2.44)

denotes the relaxation coefficient of the variable \( x \). This value which occurs here as a result of our consideration describes the macroscopic behavior of a thermodynamic system brought out of equilibrium. Eq. (2.43) is readily integrated to yield for the relaxation of the variable \( x \) starting from the nonequilibrium value \( x_0 \)

\[ (x(t))_{x(0)=x_0} = x_0 \exp (-\lambda (t - t_0)) \].  
(2.45)

Therefore, \( \lambda^{-1} \) plays the role of the decay time of the initial state.

On the other hand, \( \lambda \) stands in close relation to the fluctuational properties. Indeed, inserting (2.21) we obtain the static fluctuation-dissipation relation usually called Einstein’s relation

\[ \lambda = k_B \frac{L}{\langle x^2 \rangle}. \]  
(2.46)

It connects the standard deviations of the fluctuations with the relaxation coefficient of the dissipative dynamics for the mean value (comp.also eq. (2.30)). This relation is plausible because if the standard deviation of the distribution is small the relaxation of possible deviations has to be fast, as well. Both, the restoration of equilibrium and the deviation from equilibrium are counteracting but strongly connected in equilibrium.

One can proceed in a quite similar way if dealing with several thermodynamic values. Indeed, taking the entropy as function of \( x_i \) (\( i = 1 \ldots n \)) the entropy production reads

\[ \dot{S}((<x_1>, \ldots, <x_n>)) = \frac{\partial S}{\partial <x_i>} (\dot{x}_i) \]  
(2.47)

We agree, here and furtheron, to sum over repeating indices.

According to Onsager, for the thermodynamic forces and fluxes

\[ X_i = \frac{\partial S}{\partial x_i}, \quad J_i = \dot{x}_i \]  
(2.48)
a linear relation exists

\[ J_i = L_{i,j} X_j \]  

(2.49)

with a symmetric and positive definite matrix of transport coefficients \( L_{i,j} \). We later will come back to the symmetry. The positivity is an immediate consequence of the second law of thermodynamics.

Performing the derivative in eq. (2.48) with the Taylor expanded entropy we obtain equations of motion for the variable \( x_i \)

\[ \dot{x}_i = -k_B L_{i,j} g_{j,k} x_k . \]  

(2.50)

The relaxational coefficients of the linear processes near equilibrium states are given by

\[ \lambda_{i,j} = k_B L_{i,k} g_{k,j} . \]  

(2.51)

Since \( \dot{\gamma} \) determines the stationary standard deviation we again meet the deep relation between fluctuations and and counteracting dissipation.

2.4 Correlations and spectra

Fluctuating processes can be quantified by their stationary or time dependent moments. Between them the first and second achieve central interest and are often sufficient to investigate the behavior of fluctuations.

In equilibrated systems the first moment of the probability distribution

\[ \langle x \rangle = \int x P(x) \, dx . \]  

(2.52)

gives a stationary value. It need not coincide with the equilibrium state of the system but with no loss of generality we furtheron suppose \( \langle x \rangle = 0 \).

In the following we consider time correlation functions and their relations to spectra. These functions contain important informations about the dynamics of the many particle systems.

A correlation function measures the dependence of the variable \( x \) taken at different times. Conventionally it is given by the average of the product of \( x(t) \) and \( x(t + \tau) \)

\[ C_{x,x}(t, \tau) = \langle x(t) x(t + \tau) \rangle = \int x_1 x_2 P(x_1, t; x_2, t + \tau) \, dx_1 \, dx_2 ; \quad \tau \geq 0 \]  

(2.53)

where \( P(x_1, t_1; x_2, t_2) \) \( dx_1 \, dx_2 \) with \( t_1 < t_2 \) is the joint probability that \( x(t) \) is inside \( [x_1, x_1 + dx_1] \) at time \( t_1 \) and, respectively in \( [x_2, x_2 + dx_2] \) at \( t_2 \). By definition the joint density can be expressed through the transition probability density as

\[ P(x_1, t_1; x_2, t_2) = P(x_2, t_2|x_1, t_1) P(x_1, t_1) . \]  

(2.54)
In equilibrium the fluctuations of the variables \( x \) form a stationary process which means that \( P(x, t) \) is independent of time and the transition density depends on time differences \( \tau = t_2 - t_1 \), only. Therefore the autocorrelation function is independent of the running time

\[
C_{x,x}(\tau) = \int \left( x_1 P(x_1) \int x_2 P(x_2, \tau|x_1) \, dx_2 \right) \, dx_1 .
\] (2.55)

The second integral defines the conditional mean value. Hence one might write compactly

\[
C_{x,x}(\tau) = \int x_1 \langle x(\tau) \rangle_{x(0)=x_1} P(x_1) \, dx_1 .
\] (2.56)

Thus the time correlation function is defined as the average over the stationary probability distribution \( P(x) \) which has been established at time \( t = 0 \).

For a multivariate stationary process \( x_i(t) \) \( (i = 1, \ldots n) \) the correlations form a matrix with elements

\[
C_{i,j}(\tau) = \int x_i \langle x_j(\tau) \rangle_{x_j(0)=x_j} P(x_1 \ldots x_n) \, dx_1 \ldots dx_n .
\] (2.57)

From stationarity it follows immediately:

(i) Since \( C_{i,j}(\tau) \) does not depend on the actual time we find by shifting \( t \rightarrow t'-\tau \)

\[
C_{i,j}(\tau) = \langle x_i(0) x_j(\tau) \rangle = \langle x_i(-\tau) x_j(0) \rangle = C_{j,i}(-\tau) \] (2.58)

At \( \tau = 0 \) it coincides with the matrix of second moments \( C_{i,j} = \langle x_i x_j \rangle \). The behaviour for large \( \tau \) is determined by the ergodic properties of the considered stationary process. Under mixing conditions the correlation function vanishes if \( \tau \) goes to infinity.

(ii) The Fourier-transform of the time correlation function

\[
S_{i,j}(\omega) = \int_{-\infty}^{+\infty} C_{i,j}(\tau) \, e^{i \omega \tau} \, d\tau
\] (2.59)

is called the spectrum of the fluctuating values. It can be calculated either by finding the time correlation function via Fourier transform (2.59) or by the analysis of the dynamics of the fluctuating values. Indeed, if the Fourier-components of \( x_i(t) \)

\[
x_{i\omega} = \int_{-\infty}^{+\infty} dt \, x_i(t) \, e^{i \omega t}
\] (2.60)

are multiplied by \( x_{j\omega} \) and averaged over the stationary distribution

\[
\langle x_{i\omega} x_{j\omega'} \rangle = \int_{-\infty}^{+\infty} dt \int_{-\infty}^{+\infty} dt' \, \langle x_i(t)x_j(t') \rangle \, e^{i(\omega+\omega')t+i\omega'(t'-t)}
\] (2.61)
it yields due to the stationarity
\[
\langle x_i \omega x_j \omega' \rangle = 2\pi \delta (\omega + \omega') S_{i,j}(\omega)
\] (2.62)
For real processes \(x_i \omega = x_i^* - \omega\) this relation expresses the fact that only the complex conjugated values are correlated and contribute to the Fourier transform of the correlation function. For each \(\omega\) it forms a nonnegative definite matrix.

(iii) It is easy to show by virtue of (2.62) that
\[
\int |x(t)|^2 dt = \frac{1}{2\pi} \int |x_\omega|^2 d\omega
\] (2.63)
(Parseval theorem) where \(x_\omega\) is defined in (2.60).

(iv) A physical meaning for the Fourier transform of the correlation function was found by Wiener and Khinchin and formulated in a theorem named after them. Let us look on finite sequences of the real stochastic trajectory over the length \(T\) with \(x_T(t) = x(t)\) for \(|t| \leq T/2\) and \(x_T(t) = 0\) otherwise and Fourier transform
\[
x^*_\omega = \int_{-T/2}^{T/2} dt x(t) \exp(i\omega t).
\] (2.64)
As \(T \to \infty\) the original full stochastic process is recovered.

Following Parseval’s theorem the expression
\[
\frac{1}{T} \int_{-\infty}^{+\infty} (x^T(t))^2 dt = \frac{1}{T} \int_{-T/2}^{T/2} x^2(t)dt = \frac{1}{2\pi T} \int_{-\infty}^{+\infty} |x^*_\omega|^2 d\omega
\] (2.65)
alows a spectral decomposition. For real processes
\[
x^*_\omega \ (x^*_\omega)^* = x^*_\omega \ x^*_\omega = |x^*_\omega|^2,
\] (2.66)
is symmetrical for positive and negative frequencies and, hence, the integration is restricted to positive frequencies.

The left and right hand sides of Parseval’s relation stand for the summed energies of the deviations from the mean \(\langle x \rangle = 0\), either as integrals over the trajectory or as integrals over the amplitudes of the Fourier components. Then the power contained in the frequency band \([\omega, \omega + d\omega]\) with \(\omega \geq 0\) can be assigned as
\[
P^T(\omega) \ d\omega = \frac{2}{T} |x^*_\omega|^2 d\omega
\] (2.67)
with \(P^T(\omega)\) being the spectral power density of the considered time series \(x(t)^T\).

Averaging over an ensemble of different realizations yields
\[
\langle P^T(\omega) \rangle = \frac{2}{T} \langle |x^*_\omega|^2 \rangle = \frac{2}{T} \int_{-T/2}^{T/2} \langle x(t_1) x(t_2) \rangle \exp(i\omega(t_1 - t_2)) dt_1 dt_2.
\] (2.68)
Correlations and spectra

With the stationary correlation function inserted one obtains with $\tau = t_1 - t_2$

$$\langle P^T(\omega) \rangle = 2 \int_{-T}^{T} \left(1 - \frac{|\tau|}{T}\right) C_{x,x}(\tau) \exp(i\omega \tau) \, d\tau . \quad (2.69)$$

If the integral

$$\int_{0}^{\infty} \tau \, C_{x,x}(\tau) \, d\tau < \infty \quad (2.70)$$

converges the long time limit relates the spectral power density of the process $x(t)$

$$P(\omega) = \lim_{T \to \infty} \langle P^T(\omega) \rangle$$

(2.71)

to the Fourier transform of the correlation function as follows:

$$P(\omega) = 2 \int_{-\infty}^{\infty} C_{x,x}(\tau) \exp(i\omega \tau) \, d\tau = 2S_{x,x}(\omega). \quad (2.72)$$

Later on we will omit the difference between both values and refer the Fourier transform of the correlation function as to the power spectral density or, respectively, the power spectrum.

(v) As an example we consider a diffusing particle with a second moment spreading in time according to

$$\langle x^2(t) - x^2(0) \rangle = 2D t . \quad (2.73)$$

The l.h.s. can be replaced by the velocity correlation function as

$$\langle x^2(t) - x^2(0) \rangle = \left< \int_{0}^{t} \dot{x}(s) \, ds \int_{0}^{t} \dot{x}(s') \, ds' \right> . \quad (2.74)$$

Therefore it holds

$$D = \frac{1}{2t} \int_{0}^{t} \int_{0}^{t} C_{v,v}(s - s') \, ds \, ds' \quad (2.75)$$

In the limit $t \to \infty$ one finds

$$D = \frac{1}{2} \int_{-\infty}^{+\infty} C_{v,v}(\tau) \, d\tau . \quad (2.76)$$

For the relaxation of a heavy particle with mass $m$ in a viscous medium one has (comp.eq. (2.108))

$$C_{v,v}(\tau) = \frac{k_B T}{m} \exp\left(-\frac{\gamma}{m} |\tau|\right) , \quad (2.77)$$
where $\gamma$ is Stoke’s friction coefficient. Insertion into eq. (2.76) yields the well known Einstein formula

$$D = \frac{k_B T}{\gamma}$$

(2.78)

which will be met and discussed several times later on.

### 2.5 Linear response

In the section (2.2) we derived the linear relaxational dynamics for the mean values taking into account the thermodynamic force but neglecting fluctuating forces. Arriving at this conclusion Onsager made a second long ranging Ansatz with reading consequences. In the derivation of dynamics we did not specify whether the initial non-equilibrium state was prepared as a result of an external force like usually by considering the dynamics of mean values or in result of a fluctuation permanently generated in thermodynamic systems.

Onsager postulated that there is no difference in the dynamics of the regression of a fluctuating value and the behavior of the mean value perturbed from its marginal equilibrium state by an external force. In consequence this regression hypothesis means that the behavior of the fluctuations are determined by the relaxational coefficients. Also the time correlation function and the spectrums are obtained from the kinetic laws for the mean values.

We will see later that a more profound description of the fluctuation will lead us to the Langevin-equations where fluctuating forces have to be taken into the description. It turns out that this is equivalent to Onsager’s regression hypothesis, saying that the dynamics of the conditioned averages that enter the expression (2.57) for the correlation function is given by the relaxational dynamics of the mean values.

To prove the regression hypothesis of Onsager we consider the relaxation of the mean value $\langle x(t) \rangle$ for $t > 0$ if the thermodynamic system was driven out of equilibrium by an external force $F(t)$ with properties

$$F(t) = \begin{cases} F_0 & t \leq 0 \\ 0 & t > 0 \end{cases}$$

(2.79)

We assume that the force has acted for a sufficiently long time thus that the stationary probability distribution is reached at $t = 0$ when the force is switched off. For $t = 0$ we find for the average value of $x$

$$\langle x_0 \rangle_{F_0} = \int x_0 P(x_0, F_0) \, dx$$

(2.80)

with

$$P(x_0, F_0) \sim e^{\frac{S(x_0)}{k_B} - \frac{H(x_0, F_0)}{k_B t}}$$

(2.81)
where $H(x_0)$ is the Hamilton-function $H = -x_0 F_0$ associated to the force $F(t)$.

In the absence of a force at $t > 0$, the mean value of $x$ will relax to its unconstrained equilibrium value $\langle x \rangle_{eq} = 0$. Obviously, the evolution of mean $\langle x(t) \rangle_{x(0)=x_0}$ at $t > 0$ is conditioned to its fluctuating initial value $x_0$ which is distributed according to (2.81). Then, the relaxation process averaged over initial values reads

$$\langle x(t) \rangle_{F_0} = \frac{\int dx_0 \langle x(t) \rangle_{x(0)=x_0} \exp \left( \frac{S(x_0) - H(x_0)}{k_B T} \right)}{\int dx_0 \exp \left( \frac{S(x_0) - H(x_0)}{k_B T} \right)} \; ; \; t > 0 .$$  \tag{2.82}$$

For small forces the averaging in (2.82) can be performed by the unperturbed distribution without force. Indeed, expanding in the numerator and denominator as

$$\exp \left( \frac{H(x_0)}{k_B T} \right) \approx 1 + \frac{H(x_0)}{k_B T} ,$$

we find eventually

$$\langle x(t) \rangle_{F_0} = \frac{F_0}{k_B T} \int \langle x(t) \rangle_{x(0)=x_0} x_0 P(x_0) \, dx_0 .$$  \tag{2.84}$$

The integral on the r.h.s. defines just the time correlation function in (2.53). For small forces it holds

$$\langle x(t) \rangle_{F_0} = \frac{F_0}{k_B T} C_{x,x}(t) .$$  \tag{2.85}$$

Thus, in case of small perturbations the time correlation function is proportional to the mean value of the relaxational process.

In conclusion we are able to write down the kinetic equation for the time correlation function which are identically to the the linear relaxation dynamics. For the autocorrelation function of $x(t)$ we formulate

$$\frac{d}{d\tau} C_{x,x}(\tau) = -\lambda C_{x,x}(\tau)$$  \tag{2.86}$$

with the initial conditions

$$C_{x,x}(\tau = 0) = \langle x^2 \rangle = g^{-1} .$$  \tag{2.87}$$

Integrating (2.86) yields

$$C_{x,x}(\tau) = g^{-1} e^{-\lambda |\tau|}$$  \tag{2.88}$$

by virtue of (2.58). The generalization to several fluctuating variables is straightforward. The application of Onsager’s regression hypothesis leads to the kinetic equations

$$\frac{d}{d\tau} C_{i,j}(\tau) = -\lambda_{i,k} C_{k,j}(\tau)$$  \tag{2.89}$$
that has to be solved with the initial conditions
\[ C_{i;j}(0) = \langle x_i x_j \rangle = g_{i;j}^{-1}. \] (2.90)

This can be done by means of the one-side Fourier-transformation of the time correlation functions
\[ S_{i;j}^+(\omega) = \int_0^\infty e^{i\omega \tau} C_{i;j}(\tau) \, d\tau. \] (2.91)

It gives an algebraic equation for \( S_{i;j}^+(\omega) \)
\[ (-i\omega \delta_{i,k} + \lambda_{i,k}) S_{k;j}^+(\omega) = \langle x_i x_j \rangle \] (2.92)

The expression (2.91) together with the complex conjugated value
\[ [S_{j,i}^+(\omega)]^* = \int_0^\infty e^{-i\omega \tau} C_{j,i}(\tau) \, d\tau = \int_0^\infty d\tau e^{-i\omega \tau} C_{i;j}(-\tau) = \int_{-\infty}^\infty e^{i\omega \tau} C_{i;j}(\tau) \, d\tau \] (2.93)
determines the spectrum of the considered process
\[ S_{i;j}(\omega) = S_{i;j}^+(\omega) + [S_{j,i}^+(\omega)]^* \] (2.94)
where we have used (2.58). Thus we find
\[ S_{i;j}(\omega) = (-i\omega \delta_{i,k} + \lambda_{i,k})^{-1} g_{k;j}^{-1} + g_{i;k}^{-1} (-i\omega \delta_{k,j} + \lambda_{k,j})^{-1} \] (2.95)
or in a more compact form
\[ S_{i;j}(\omega) = [(-i\omega \delta_{i,k} + \lambda_{i,k}) (i\omega \delta_{j,l} + \lambda_{j,l})]^{-1} (L_{k,j} + L_{l,i}) \] (2.96)
with \( L_{i;j} \) being the Onsager coefficients. The correlation function \( C_{i;j}(\tau) \) follows from the inverse Fourier transform.

At the end of this chapter we will come back to Onsager’s coefficient \( L_{i;j} \) that establish the linear relation between thermodynamic fluxes and forces. Empirically it was found that the coefficients are symmetric
\[ L_{i;j} = \pm L_{j,i} \] (2.97)

This macroscopic property means that if a force \( X_j \) causes the thermodynamic flux \( J_i \) in turn the thermodynamic force \( X_i \) gives rise to a flux \( J_j \). In this general form this law was first pointed out by Onsager. Special cases had been known long before.

The reason for the symmetry of the transport coefficients lies in the stationarity of the process and the reversibility of the underlying microscopic Hamiltonian motion. As a consequence the motions forward and backward in time are related by
\[ x_i(t) = \epsilon_i x_i(-t) \] (2.98)
where $\varepsilon_i$ is the parity under time reversal which is $+1$ for even variables and $-1$ for odd ones. For the correlation function this yields

$$C_{i,j}(\tau) = \langle x_i(t)x_j(t+\tau) \rangle = \varepsilon_i\varepsilon_j \langle x_i(-t)x_j(-t-\tau) \rangle = \varepsilon_i\varepsilon_j C_{i,j}(-\tau)$$  \hspace{1cm} (2.99)

and with the stationarity (2.58) eventually

$$C_{i,j}(\tau) = \varepsilon_i\varepsilon_j C_{j,i}(\tau).$$  \hspace{1cm} (2.100)

Further we write down the dynamical equations for both correlation functions

$$\frac{d}{d\tau} C_{i,j} = -\lambda_{i,k} C_{k,j}(\tau)$$  \hspace{1cm} (2.101)

$$\frac{d}{d\tau} C_{j,i} = -\lambda_{j,k} C_{k,i}(\tau).$$  \hspace{1cm} (2.102)

Hence we obtain

$$\sum_k \lambda_{i,k} C_{k,j}(\tau) = \varepsilon_i\varepsilon_j \sum_k \lambda_{j,k} C_{k,i}(\tau)$$  \hspace{1cm} (2.103)

For $\tau = 0$ the correlation function yield the standard deviation with $C_{i,j}(0) = g_{i,j}^{-1}$ and with the definition of the relaxational coefficients (2.51) one finds the Onsager-Casimir relation

$$L_{i,j} = \varepsilon_i\varepsilon_j L_{j,i}$$  \hspace{1cm} (2.104)

### 2.5.1 Colored noise

As an example we consider the motion of a heavy particle with mass $m$ in a viscous liquid. For the mean velocity conditioned to its initial value $v(t)$ we have

$$m\frac{d}{d\tau} \langle v(t+\tau) \rangle = -\gamma \langle v(t+\tau) \rangle.$$  \hspace{1cm} (2.105)

and hence for the stationary correlation function

$$m\frac{d}{d\tau} \langle v(t+\tau)v(t) \rangle = -\gamma \langle v(t+\tau)v(t) \rangle$$  \hspace{1cm} (2.106)

with the initial condition

$$\langle v(t)^2 \rangle = \frac{k_B T}{m}.$$  \hspace{1cm} (2.107)

It gives the decaying correlation function

$$\langle v(t+\tau)v(t) \rangle = \frac{k_B T}{m} \exp \left( -\frac{\gamma}{m} |\tau| \right).$$  \hspace{1cm} (2.108)
The considered stochastic process \( v(t) \) is called the Ornstein-Uhlenbeck process (more detailed information about colored noise, its relation with other noise and its effect on nonlinear dynamics can be found in (van den Broeck, 1983; Hänggi, Jung, 1995). Since not all frequencies are excited with equal strength the velocity \( v(t) \) is referred to as a colored process. In particular, for the present prevailing low frequency excitations one finds it also is termed red noise.

### 2.5.2 Harmonic noise

A richer spectrum for a colored process is obtained if a harmonic force additionally acts on the particle. The system of equations for the velocity autocorrelation and the velocity position correlation function reads

\[
\begin{align*}
\dot{C}_{x,v}(\tau) &= C_{v,v}(\tau) \\
\dot{C}_{v,v}(\tau) &= -\frac{\gamma}{m} C_{v,v}(\tau) - \omega_0^2 C_{x,v}(\tau)
\end{align*}
\]

Following the approach (2.89-2.95) one simply finds the spectrum

\[
S_{v,v}(\omega) = \frac{k_B T}{m} \frac{2\gamma/m}{(\gamma/m)^2 + \omega^2 + \frac{\omega_0^2}{\omega^2}}
\]

which now possesses a peak at a finite frequency.

### 2.5.3 Fluctuation dissipation theorem

We come back to the reaction of the system on a small external force that has been adiabatically switched on at \( t \to -\infty \).

This problem usually is described within the linear response theory by introducing the response function \( \chi(t) \) of a many particle system. It describes the response of the mean value of \( x(t) \) on the external force \( F(t) \) related by

\[
\langle x(t) \rangle = \int_{-\infty}^{t} \chi(t - t') F(t') \, dt'.
\]

The response function coincides with the Green’s-function up to a multiplicative constants. Its Fourier transform \( \chi_\omega \) is known as the generalized susceptibility. Because of causality, the response function vanishes for negative times, \( \chi(\tau) = 0 \) if \( \tau \leq 0 \).

It is obvious that (2.113) holds only in for sufficiently small forces. As a consequence of the linear approximation, in the frequency regime there is only a response
at the frequency of the force. The Fourier transform convolute simply $\langle x_\omega \rangle$ with $F_\omega$ as

$$\langle x_\omega \rangle = \chi(\omega) F_\omega$$

(2.114)

On the other hand the mean $\langle x(t) \rangle$ and the correlation function are strongly related by (2.85). Therefore, it comes out that the response function is related to $C(\tau)$ as well. Indeed combining (2.84) and (2.113) and using the expression for the force $F(t)$ we get (now already independent of $F_0$ as result of the linear approximation)

$$C_{x,x}(\tau) = k_B T \int_\tau^\infty \chi(s) \, ds$$

(2.115)

with boundary condition $\int_0^\infty \chi(s) \, ds = \langle x^2 \rangle / k_B T$. Differentiation with respect to time gives

$$\frac{\partial C_{x,x}(\tau)}{\partial \tau} = -k_B T \chi(\tau) .$$

(2.116)

The expression (2.116) connects two fundamental physical processes and is known as the fluctuation-dissipation theorem. On the l.h.s. the correlation function describes the regression of fluctuations which is related to the response function upon external forces of a many particle system. The origin of this connection is in the common source of fluctuations and dissipation that is caused by the microscopic motion of the particles.

In the frequency domain it is sufficient to know the imaginary part of the susceptibility $\chi(\omega)$ in order to determine the power spectrum:

$$(\chi_\omega - \chi'_\omega) = -\frac{1}{k_B T} \int_{-\infty}^{+\infty} \frac{\partial C_{x,x}(\omega)}{\partial \omega} e^{i \omega t} \, dt = \frac{i \omega}{k_B T} S(\omega) .$$

(2.117)

Therefore, the spectrum of the considered value $x(t)$ is determined by the imaginary part of the $\chi_\omega$

$$S_{x,x}(\omega) = \frac{2 k_B T}{\omega} \Im \chi_\omega$$

(2.118)

It is this relation which gives the connection between fluctuations and dissipation in the frequency domain.

To show that $\Im \chi(\omega)$ is directly related to dissipation we consider the energy which is dissipated during one period if a system is driven by a periodic force (Landau, Lifshitz, 1971)

$$f(t) = \Re \left( f_0 e^{-i \omega t} \right) = \frac{1}{2} \left( F_0 e^{-i \omega t} + F_0^* e^{i \omega t} \right)$$

(2.119)
The energy change

$$\frac{dE}{dt} = - \langle x \rangle \frac{df}{dt} \quad (2.120)$$

with

$$\langle x \rangle = \frac{1}{2} \left( \chi(\omega) f_0 e^{-i\omega t} + \chi(-\omega) f_0^* e^{i\omega t} \right) \quad (2.121)$$

during one period $T$ stands for the dissipation rate $Q$ of the mechanical work in the many particle system. It is determined by the imaginary part of the susceptibility only

$$Q = \frac{1}{T} \int_0^T \frac{dE}{dt} dt = \frac{\omega}{2} |f_0|^2 \Im(\chi(\omega)) \quad (2.122)$$

It makes clear the meaning of eq. (2.118). Just the part responsible for the dissipation in many particle system determines the spectrum of the fluctuations.

As an example we again consider harmonic noise. The susceptibility is easily found

$$\chi_{\omega} = \frac{1}{m} \frac{1}{\omega^2 - \omega^2 + i\omega\gamma/m} \quad (2.123)$$

The spectrum becomes after insertion into (2.118)

$$S_{x,x}(\omega) = \frac{k_B T}{m} \frac{\gamma/m}{(\omega_0^2 - \omega^2)^2 + \omega^2(\gamma/m)^2} = \omega^{-2} S_{v,v}(\omega) \quad (2.124)$$

with $S_{v,v}(\omega)$ as defined in (2.112).

### 2.5.4 Nyquist theorem. White noise

Another question was posed in the work of Nyquist who considered current fluctuations in electronic circuits. He asked about the spectrum of the pondermotic forces which originate the permanent fluctuations from the rest states.

We consider an one dimensional conductor of length $l$ and area $A$ in which charged move under the influence of friction described by Stoke’s law. The conductivity of this conductor can be determines in linear response by applying a constant external electric field $E$.

For the dissipative motion of the averaged velocities of $N$ carriers we assume

$$m \langle \dot{v}_i \rangle = -\gamma \langle v_i \rangle + eE \quad (i = 1 \ldots N). \quad (2.125)$$

The stationary current density follows as

$$j = \frac{1}{V} e \sum_i \langle v_i \rangle = \frac{ne^2}{\gamma} E. \quad (2.126)$$
Therefore, the conductivity becomes

$$\sigma = \frac{n e^2}{\gamma}$$

and the resistance of the considered conductor is given by

$$R = \frac{l}{\sigma A} = \frac{l^2}{\gamma N e^2}.$$  

The conductor is embedded in a bath with temperature $T$. This leads to Gaussian distributed fluctuations (acc. to sect. (2.3)) of the velocities of the independent carriers. Its velocity spectrum is according to (2.109) given by

$$S_{v_i, v_j}(\omega) = \delta_{i,j} \frac{k_B T m}{m} \frac{2\gamma/m}{(\gamma/m)^2 + \omega^2}.$$  

The permanent motion of the particles will also induce a permanently fluctuating current

$$I(t) = \frac{e}{l} \sum_{i=1}^N v_i.$$  

Its spectrum is found from (2.129)

$$S_{I,I}(\omega) = N \frac{k_B T e^2}{m^2} \frac{2\gamma/m}{(\gamma/m)^2 + \omega^2}.$$  

Coming back to Nyquist’s problem we ask what the origin of the permanent motion of the carriers is that result in the current fluctuations. The linear dynamics (2.43,2.50) can not answer this question, because it describes the relaxation to the equilibrium states of vanishing mean values.

Earlier Langevin had answered a related question for Brownian motion which will be considered in the next section. In order to obtain fluctuations of the current noisy electric fields have to be incorporated into the description of the motion of the carriers.

Thus we will add to the equation of the fluctuating velocities time dependent electromotoric forces $eE_i(t)$

$$m \dot{v}_i = -\gamma v_i + eE_i(t)$$

where $E_i$ denotes the electric field at the position of the $i$-th charge. On average the electric fields vanish; Otherwise we would have non-vanishing mean velocities. For the current it results in an additional time dependent voltage over the resistance which induces current fluctuations

$$\xi(t) = \frac{l}{N} \sum_i E_i(t)$$
in Kirchhoff’s law

\[ L \dot{I} = -RI + \xi(t) \]  

(2.134)

where \( L \) denotes the inductance

\[ L = \frac{R m}{\gamma} = \frac{l^2 m}{Ne^2} \]  

(2.135)

and the After Fourier transform of eq. (2.134) we find

\[ I_\omega = \frac{\xi_\omega}{Z(\omega)} \]  

(2.136)

where \( Z(\omega) \) is the complex impedance of the conductor. It is the inverse of the susceptibility \( \chi(\omega) = 1/Z(\omega) \). The correlation of the Fourier transformed currents becomes

\[ \langle I_\omega I^*_\omega \rangle = \frac{1}{(1/\omega L + R) (1/\omega L + R)} \langle \xi_\omega \xi^{*}_\omega \rangle . \]  

(2.137)

Using eq. (2.62 one can express the current spectrum in terms of the voltage spectrum

\[ S_{I,I}(\omega) = |\chi(\omega)|^2 S_{\xi,\xi}(\omega) = \frac{1}{\omega^2 L^2 + R^2} S_{\xi,\xi}(\omega) \]  

(2.138)

where \( S_{\xi,\xi}(\omega) \) is the spectrum of the spontaneous voltage changes that drive the current fluctuations. Comparison with (2.131) leads to Nyquist’s result

\[ S_{\xi,\xi}(\omega) = 2Rk_B T . \]  

(2.139)

The spectrum of the spontaneously occurring voltage drops in the resistance does not depend on the frequency. It is called white noise. Because \( I(t) \) is Gaussian \( \xi(t) \) is Gaussian as well.

Obviously we have

\[ \langle \xi(t) \rangle = 0 . \]  

(2.140)

It is interesting to have a look at the correlation function of the voltage drops. Backward Fourier transformation results in

\[ \langle \xi(t) \xi(t') \rangle = 2\epsilon \delta(t - t') \]  

(2.141)

with

\[ \epsilon = Rk_B T \]  

(2.142)

where \( \delta(t) \) is Dirac’s-delta function. Hence white noise is uncorrelated in time. The factor \( \epsilon \) denotes the noise intensity. Its particular form as the product of thermal energy \( k_B T \) and the resistance \( R \), that characterizes the dissipation, is in accordance
with the Einstein relation. We note that specifying the first and second moment (2.140,2.141) is sufficient to characterize the Gaussian process $\xi(t)$.

The fact that $\xi(t)$ is uncorrelated at different times, and that its variance is infinite renders this process highly irregular jumping between plus and minus infinity. As a result it is nowhere continuous in time. Obviously, in physics white noise is an idealization. It is a good approximation for processes whose correlation time is much smaller than all other characteristic time scales.

One also observes that white noise forces acting in the linear dynamics (2.134) exhibit a fluctuating current with a red spectrum (2.138). The resistance acts as low pass filter for the applied white noise excitations. The current fluctuations are exponentially correlated

$$\langle I(t)I(t+\tau)\rangle = \frac{2k_B T}{L} \exp\left(-\frac{R}{L}|\tau|\right)$$

with correlation time $\tau_C = L/R$. Analogously, a RCL circuit would generate fluctuations of the current with a spectrum given by eq. (2.112) and $\omega_0$ is the eigenfrequency of the circuit $1/LC$.

In case of a strong resistance the correlation time in (2.143) becomes vanishingly small. The exponential function converges to a $\delta$-function and the current becomes white noise with intensity $2k_B T/R$.

### 2.5.5 White noise and the Wiener process

Though white noise is an idealization in physics it is widely used in the modern literature for modeling dynamical behavior under the influence of noise. We will list several properties of these systems in the next sections. The reason why it is often employed lies in the simpler mathematical treatment of processes originated by white noise. The important property of a process driven by white noise is that the change of the system within a small time step only depends on the actual state of the process but not on its history. Hence, the response on the white noise source is without memory which is not the case for correlated noise as will be seen in the section of the generalized Langevin equation. A dynamical variable of the system driven by white noise (as in the considered example, the current) is still a Markovian processes. In a differential shape $I(t+dt)$ is given in probability as a result of the value of the current at time $t$ and the random forces within the interval $[t,t+dt]$, independently from the former history.

In the mathematical literature white noise is defined in terms of its increment over a finite time step $dt$ after time $t$. Assigning the increment

$$dW(t) = dW_{dt} = \int_t^{t+dt} \xi(s) \, ds$$

(2.144)
the process is stationary and Gaussian with zero mean

$$\langle dW_{dt} \rangle = 0.$$  \hspace{1cm} (2.145)

Further, looking at the variance of the increments one finds

$$((dW_{dt})^2) = 2\epsilon dt$$  \hspace{1cm} (2.146)

which is Fick’s second law. For nonoverlapping time intervals $dt$ and $dt'$ the increments are uncorrelated.

The sum over the increments at subsequent time intervals is known as the Wiener process

$$W_t = \lim_{N \to \infty} \sum_{i=1}^{N} dW_{dt_i}$$  \hspace{1cm} (2.147)

with $dt_i = t_i - t_{i-1}$ and $t_N = t$ and $t_0 = 0$. It is continuous in time, nowhere differentiable and Markovian because its increments are independent one from each other.

2.6 Brownian Motion

2.6.1 Einstein’s relation

Stochastic methods became intrinsic part of the physical language in the context of the theoretical explanation of Brownian motion. This is the erratic motion of small particles immersed in a fluid. It is named after the botanist Robert Brown in memory of its discovery who observed highly random motion of pollen grains suspended in fluids in 1827. Montroll and West (1979) reported that the court physician to Maria Theresa whose name was Jan Ingenhousz observed erratic motion of finally powdered charcoal floating on alcohol surfaces earlier in 1785 which became not known to the non-Dutch speaking world.

Brown’s first intuition was that he had found the “Secret of Life”. Later he repeated the experiments with mineral particles and still found the permanent motion and his first hypothesis failed. His observation became a serious physical problem. First of all it was the first experimental proof showing deviations from the continuum theory of matter. The motion of the particles is the result of the molecular agitation resulting from their interaction with a finite number of the surrounding liquid molecules. In consequence Brownian motion supported the ideas of the molecular structure of matter yielding the possibility of the determining Boltzmann’s constant and Avargado’s number, see eq. (2.31).

The second deep conclusion concerns the diffusion coefficient $D$. It is responsible for two different physical processes. On the one hand it describes the relaxation of
an inhomogeneous distribution of a density inducing a matter flow by Fick’s law

\[ j_d = -D \frac{\partial n}{\partial x} \]  
(2.148)

The same quantity describes the random behavior of a single particle. The mean square displacement of a Brownian particle increases linear in time with a proportionality factor given by the diffusion coefficient

\[ \left\langle (x(t) - x(0))^2 \right\rangle = 2Dt \]  
(2.149)

Therefore the continuity equation

\[ \frac{\partial n}{\partial t} = D \frac{\partial^2 n}{\partial x^2} \]  
(2.150)

has an ambiguous character. On the one hand it models the homogenization of a macroscopic density \( n(x,t) \) and on the other hand it describes the evolution of probability density of single particles. In statistical mechanics the motion of the particle is expressed by mechanical laws. Connecting the deterministic diffusion with the fluctuating Brownian motion gives the possibility to determine the transport coefficient in terms of the mechanical properties of particles immersed in a fluid at a certain temperature.

We already have underlined this connection (eqs. (2.46), (2.78) and (2.139)). Here, we look at different fluxes which compensate each other in equilibrium. The distribution of gases in the gravitational field of the earth is due to the barometric formula

\[ \frac{dn}{dx} = -\frac{mg}{k_B T} n(x) \]  
(2.151)

The establishment of a stationary distribution is the compromise of two tendencies. Due to their weight the particles are attracted by gravity i.e. they fall to the earth-surface whereas the diffusion counteracts to homogenize the distribution. In equilibrium the diffusional flux (2.148) and the gravity flux

\[ j^{gr} = -n(x) \frac{mg}{\gamma} \]  
(2.152)

compensate each other

\[ j_d + j^{gr} = 0 \]  
(2.153)

Together with the Einstein relation (2.78)

\[ D = \frac{k_B T}{\gamma} \]  
(2.154)

the barometric formula results.
2.6.2 Brownian motion as Markovian dynamics

At the beginning of the theoretical treatment of Brownian motion stood the important insight that the erratic motion of the particles only could be interpreted within the frame of a probabilistic theory. It was Einstein and v. Smoluchowski (1906) and, later, Langevin who pursued this idea.

Einstein was unfamiliar with the experimental facts. He predicted in his work the agitation of suspended particles as a clever approach to determine experimentally Avargado’s number in which Perrin succeeded experimentally.

Einstein saw the reason for the ongoing irregular motion in the molecular nature of the liquid surrounding the Brownian particles. Evidently from the molecular chaos follows an out-of-balance of the impacts of the liquid molecules upon the Brownian particles which results in the random motion. Einstein supposed that the total momentum that the particle obtains from the liquid is statistically independent at any instant of time. It means that the position of the Brownian particle after a short time step only depends on the actual positions of the particle.

With this assumption the motion of the particle becomes a Markovian process for which Einstein was able to derive a kinetic law for the time evolution of the transition probability.

If we label the displacement of a Brownian particle during the time \( \tau \) by \( \Delta \) we find for the probability density \( P(x, t + \tau) \) of a particle to be in the position \( x \) at time \( t + \tau \)

\[
P(x, t + \tau) = \int_{-\infty}^{+\infty} P(x + \Delta, t) \Phi(\Delta, \tau) d\Delta
\]

(2.155)

where \( \Phi(\Delta, \tau) \) denotes the probability of performing a displacement of length \( \Delta \) during \( \tau \). It is symmetric

\[
\Phi(\Delta, \tau) = \Phi(-\Delta, \tau)
\]

(2.156)

and normalized

\[
\int_{-\infty}^{+\infty} \Phi(\Delta, \tau) d\Delta = 1.
\]

(2.157)

The expansion for small \( \tau \) and \( \Delta \)

\[
P(x, t + \tau) = P(x, t) + \tau \frac{\partial P}{\partial t}
\]

(2.158)

\[
P(x + \Delta, t) = P(x, t) + \Delta \frac{\partial P}{\partial x} + \frac{1}{2} \Delta^2 \frac{\partial^2 P}{\partial x^2}
\]

(2.159)
yields in the limit $\tau \to 0$
\[
\frac{\partial P}{\partial t} = \frac{\partial^2 P}{\partial x^2} \lim_{\tau \to \infty} \frac{1}{\tau} \int_{-\infty}^{+\infty} \frac{\Delta^2}{2} \Phi(\Delta, \tau) \, d\Delta.
\] (2.160)

Here the first derivative with respect to the position $x$ does not contribute because of the symmetry of $\Phi(\Delta, \tau)$, see eq. (2.156). If the limit
\[
D = \lim_{\tau \to \infty} \frac{1}{\tau} \int_{-\infty}^{+\infty} \frac{\Delta^2}{2} \Phi(\Delta, \tau) \, d\Delta
\] (2.161)
exists, the kinetic law for the time evolution of the density of a Brownian particle becomes
\[
\frac{\partial P}{\partial t} = D \frac{\partial^2 P}{\partial x^2}.
\] (2.162)

This law can be written in the form of a continuity equation for the total probability
\[
\frac{\partial P}{\partial t} + \frac{\partial}{\partial x} j(x, t) = 0.
\] (2.163)

For the current one recovers Fick’s law
\[
j = -D \frac{\partial P}{\partial x}.
\] (2.164)

If the total probability initially at $t = t_0$ is concentrated at $x = x_0$ the solution of eq. (2.162) becomes
\[
P(x, t|x_0, t_0) = \frac{1}{\sqrt{4\pi D(t-t_0)}} \exp \left[-\frac{(x-x_0)^2}{4D(t-t_0)}\right].
\] (2.165)

For the spreading of particle’s position this implies
\[
\langle (x-x_0)^2 \rangle = 2D(t-t_0).
\] (2.166)

This coincides with Fick’s second law.

We note that in this derivation of Fick’s second law the probability density of a single Brownian particle is assumed to be proportional to the mass density of a swarm of particles.

Thus Einstein and v. Smoluchowski outgoing from probabilistic point of view in the interpretation of Brownian motion found the main macroscopic results. The main result of their approach consists in the derivation of a kinetic equation for the probability density. This equation was later generalized by v. Smoluchowski, Fokker and Planck, Kolmogorov, Feller and others.

The diffusion coefficient in Einstein’s theory remains undetermined. It is expressed as the second moment of the jump width during unit time. Hence further information about the distribution $\Phi(\Delta, \tau)$ is required.
We mention that the Wiener process as defined previously has identical properties as the position of the Brownian particles. The increment \( W(t) - W(t_0) \) between times \( t \) and \( t_0 \) is distributed accordingly to eq. (2.165).

### 2.6.3 Langevin’s approach

The second general approach is due to Paul Langevin and was developed three years after Einstein’s paper. Langevin added a random force in Newton’s second law to compensate for the energy loss resulting from Stoke’s friction law (comp. sect.(2.5.4))

\[
m\ddot{v} = -\gamma v + \xi(t) .
\]  

(2.167)

The reason of its occurrence can be seen in the unbalanced impacts of the liquid molecules, in the thermal agitation. Langevin stressed that his approach was “infinitely simpler” than Einstein’s one.

In mathematical terms equation (2.167) constitutes a stochastic differential equation. The Brownian particle’s velocity becomes a stochastic process. On average we expect to find the deterministic damped motion. This requires obviously

\[
\langle \xi(t) \rangle = 0 .
\]  

(2.168)

We now multiply eq. (2.167) by \( x(t) \). Simple algebra gives

\[
\frac{d}{dt} (x\dot{x}) = -\frac{\gamma}{m} x \ddot{x} + \dot{x}^2 + \frac{1}{m} x(t) \xi(t) .
\]  

(2.169)

Averaging over an ensemble we find

\[
\frac{d}{dt} \langle x\dot{x} \rangle = -\frac{\gamma}{m} \langle \dot{x}^2 \rangle + \langle \dot{x}^2 \rangle + \frac{1}{m} \langle x(t) \xi(t) \rangle .
\]  

(2.170)

The second moment of the velocity can be determined from the kinetic energy in thermal equilibrium, yielding for the second term on the right hand side of eq. (2.170)

\[
\langle \dot{x}^2 \rangle = \frac{k_B T}{m} .
\]  

(2.171)

A further assumption concerns the stochastic force. Langevin postulated that the stochastic force changes much faster than the position of the Brownian particle does. The time scales of the random part \( \tau_c \) and the friction \( \tau_{brake} = m/\gamma \) should be separated (\( \tau_c \ll \tau_{brake} \)) that there can not be an effect of the stochastic force \( \xi(t) \) on the position of the particle \( x(t) \) at the same time. It makes the \( \xi(t) \) equivalently to Einstein’s independence of the jumps at different times. Therefore, the last term vanishes

\[
\langle x(t) \xi(t) \rangle = \langle x(t) \rangle \langle \xi(t) \rangle = 0 .
\]  

(2.172)
This means for the autocorrelation function of the random forces that it possesses a very short correlation time. For the sake of definiteness we assume an exponential decay of the autocorrelations:

\[
\langle \xi(t + \tau) \xi(t) \rangle = \frac{e}{\tau_c} \exp \left( -\frac{\tau}{\tau_c} \right).
\] (2.173)

The simplifying assumption here is that \( \tau_c \) is much smaller than any other time scale of the Brownian particles so that the limit \( \tau_c \to 0 \) yielding a \( \delta \)-function for the correlations of the random force provides a good approximation

\[
\lim_{\tau_c \to 0} \frac{e}{\tau_c} \exp \left( -\frac{\tau}{\tau_c} \right) = 2e \delta(\tau).
\] (2.174)

If the particle at \( t = 0 \) is located at the origin the integration of (2.170) results in

\[
\langle x^2(t) \rangle = 2 \frac{k_B T}{\gamma} \left[ t - \frac{m}{\gamma} \left( 1 - \exp \left( -\frac{\gamma}{m} t \right) \right) \right].
\] (2.175)

For times much smaller that the brake time \( t \ll \tau_{\text{brake}} = m \gamma^{-1} \) the particle moves ballistically like a free particle

\[
\langle x^2 \rangle = \frac{k_B T}{m} t^2 = \langle v_x^2 \rangle t^2.
\] (2.176)

For large-times \( t \gg \tau_{\text{brake}} \) the variance of the position spreads linear in time

\[
\langle x^2 \rangle = 2 \frac{k_B T}{\gamma} t - \frac{2k_B Tm}{\gamma^2}.
\] (2.177)

The additive constant is the squared length of the brake path of a Brownian particle

\[
l_{\text{brake}} = \sqrt{\frac{k_B T}{m} \frac{1}{\gamma}} = \sqrt{\langle v_x^2 \rangle} \frac{\mu}{\gamma}.
\] (2.178)

Diffusion with linear growth in time of the mean square displacement takes place at length scales larger the brake path where the additive constant can be neglected.

### 2.6.4 The overdamped limit

For strong damping and times larger than the braking time \( \tau_{\text{brake}} \) the description can be considerably simplified. The velocities \( v(t) \) at subsequent time steps that are sufficiently larger than the brake time are independent of each other. Therefore the velocity itself becomes a white process.

This can also be readily seen from the equation of motion (2.167) in which the acceleration can be neglected. For large \( \gamma \) it acts effectively on short time scales \( \propto \tau_{\text{brake}} \). This results in the simple equation

\[
v(t) = \dot{x}(t) = \frac{1}{\gamma} \xi(t)
\] (2.179)
where $\xi(t)$ should be due to (2.168) and (2.174). Equation (2.179) is readily integrated

$$x(t) = x_0 + \frac{1}{\gamma} \int_0^t \xi(t') dt'. \quad (2.180)$$

For the mean square displacement one obtains:

$$\left\langle (x(t) - x_0)^2 \right\rangle = \frac{1}{\gamma^2} \int_0^t \int_0^t \xi(t') \xi(t'') dt' dt''. \quad (2.181)$$

Using the properties of the random force one finds:

$$\left\langle (x(t) - x_0)^2 \right\rangle = 2 \frac{\epsilon}{\gamma^2} t. \quad (2.182)$$

This result confirms the statistical assumption made about the stochastic force by (2.168) and (2.174) and we find convergence with (2.177) at the considered coarse grained length scale. We also obtain that the intensity of the noise $\epsilon$ which was introduced in (2.174) relates as

$$\epsilon = D \gamma^2 = k_B T \gamma. \quad (2.183)$$

In summary, Langevin was able to find a dynamical approach basing on a mechanical description for the properties of the Brownian particle. The additionally acting random forces have statistical properties as a white noise process. Obviously the specified first and second moments do not completely characterize the random force. Only for Gaussian statistics of the random force the Maxwell distribution for the velocity is recovered.

### 2.6.5 Generalized Langevin equations

In this section we discuss a Hamiltonian model of a single degree of freedom, referred to as the particle, interacting with other degrees of freedom representing the environment of the particle. Initially, the environment is prepared in a thermal equilibrium state at temperature $T$. Because we will consider the thermodynamic limit for the environment it can be considered as a heat bath at the initial temperature $T$.

We will show how for the particle a random and dissipative process results that can be characterized by a (generalized) Langevin equation (Zwanzig, 1973; Reimann, 2001).
Brownian Motion

The particle with coordinate $Q$ and momentum $P$ moves in an external potential $V_{\text{ext}}(Q)$. The bath is assumed to consist of linear oscillators that are be linearly coupled to the particle:

$$H(Q, P, q_n, p_n) = H_S(Q, P) + H_{S,B}(q_n, p_n, Q),$$

where

$$H_S(Q, P) = \frac{P^2}{2M} + V_{\text{ext}}(Q),$$

$$H_{S,B}(q_n, p_n, Q) = \sum_n \left( \frac{p_n^2}{2m_n} + \frac{m_n\omega_n^2}{2} (q_n - Q)^2 \right).$$

In the absence of the external potential, the system is translation invariant, i.e. $H(Q, P, q_n, p_n) = H(Q + a, P, q_n + a, p_n)$. The dynamics results from the canonical equations of motion

$$\dot{Q} = \frac{P}{M},$$

$$\dot{P} = -\frac{\partial V_{\text{ext}}}{\partial Q} - Q \sum_n m_n\omega_n^2 + \sum_n m_n\omega_n^2 q_n,$$

$$\dot{q}_n = \frac{p_n}{m_n},$$

$$\dot{p}_n = -m_n\omega_n^2 q_n + m_n\omega_n^2 Q.$$

Because the equations for the bath oscillator are linear, their solution can be expressed in terms of the trajectory of the particle:

$$q_n(t) = q_n(0) \cos \omega_n t + \frac{p_n(0)}{m_n\omega_n} \sin \omega_n t + \omega_n \int_0^t \sin \omega_n (t - t') Q(t') \, dt'$$

where $q_n(0)$ and $p_n(0)$ denote the initial states at $t = 0$. Inserting these expressions into the dynamics of the particle we obtain

$$M \ddot{Q}(t) = -\frac{\partial V_{\text{ext}}}{\partial Q} - \left( \sum_n m_n\omega_n^2 \right) Q(t)$$

$$+ \int_0^t dt' \sum_n m_n\omega_n^2 \sin \omega_n (t - t') Q(t')$$

$$+ \sum_n m_n\omega_n^2 \left[ q_n(0) \cos \omega_n(t) + \frac{p_n(0)}{m_n\omega_n} \sin \omega_n t \right] Q(t') .$$
A partial integration of the third term on the right hand side yields

\[
M \ddot{Q} = -\frac{\partial V_{ext}}{\partial Q} + \int_0^t dt' \dot{Q}(t') \sum_n m_n \omega_n^2 \cos \omega_n (t - t')
\]

\[
+ \sum_n m_n \omega_n^2 \left\{ \frac{p_n(0)}{m_n \omega_n} \sin \omega_n t + [q_n(0) - Q(0)] \cos \omega_n t \right\}.
\]

There are two types of forces that the bath exerts on the particle: One force explicitly depends on the initial conditions of the bath whereas the other one describes the retarded back reaction of the particles motion mediated by the bath.

If one were able to control the initial conditions of the individual bath oscillators the dynamics still would be deterministic and reversible. However, in practice it is only to prepare the initial conditions of the bath according to some statistical law.

This introduces a randomness in the initial conditions and, via the third term of the right hand side of eq. (2.192), a random force:

\[
\xi(t) = \sum_n m_n \omega_n^2 \left\{ \frac{p_n(0)}{m_n \omega_n} \sin \omega_n t + [q_n(0) - Q(0)] \cos \omega_n t \right\}. 
\]

Because it is practically impossible to precisely convert the initial bath conditions to their time reversed values, the reversibility of the motion also is lost.

As already indicated, we assume the bath to be initially prepared in a thermal equilibrium state constrained to the initial position of the particle. Hence the probability of the bath initial momenta and positions are given by:

\[
\rho(q_n, p_n|Q) = Z^{-1} \exp \left( -\frac{H_{S, B}(q_n, p_n, Q)}{k_B T} \right)
\]

where \( T \) denotes the temperature and \( Z \) the partition function

\[
Z = \int dq_n dq_n Z dN \exp \left( -\frac{H_{S, B}(q_n, p_n, Q)}{k_B T} \right).
\]

This is a Gaussian probability distribution. It completely determines the statistical properties of the random force \( \xi(t) \). As a linear combination of Gaussian random variables it is a Gaussian process which is completely determined by its mean value and its autocorrelation function:

\[
\langle \xi(t) \rangle = 0
\]

\[
\langle \xi(t) \xi(t') \rangle = \frac{1}{M^2} \gamma (t - t')
\]

where

\[
\gamma (t - t') = k_B T \sum_n m_n \omega_n^2 \cos \omega_n (t - t').
\]
This function coincides with the kernel of the second term on the right hand side of eq. (2.192) and generally is referred to as the memory friction. Eq. (2.192) may now be written in the form of a generalized Langevin equation

\[ M \ddot{Q}(t) + \int_0^t dt' \gamma(t - t') \dot{Q}(t') + \frac{\partial V_{ext}}{\partial Q} = \xi(t) \]  

(2.199)

with a Gaussian random force specified by eqs. (2.196).

The fact that the correlation function of the random force is closely related to the memory friction is known as the fluctuation dissipation theorem. In the particular case of a \( \delta \)-correlated random force

\[ \langle \xi(t)\xi(t') \rangle = 2k_B T \delta(t - t') \]  

(2.200)

we both recover the Langevin equation for a Markovian process of coordinate and velocity of the particle (2.167) and the Einstein relation (2.78).

We note, that for countably many oscillators the memory friction is a quasiperiodic function of the time difference. In order to approach an instantaneous friction, the bath must consist of a continuum of oscillators.

In the case that \( \xi(t) \) is colored noise the generalized Langevin equation possesses a friction term with memory correspondingly to the friction memory \( \gamma(t - t') \). Depending on the model under consideration the exponentially correlated Ornstein-Uhlenbeck process with memory friction (2.108) or harmonic noise with spectrum (2.124) are used to account for a temporal structure of the bath.

2.7 The Fokker-Planck equation

2.7.1 Kolmogorov’s forward and backward equations

An alternative approach to describe a Markovian process is based on the fact, that the change of the process in time only depends on the present state but not on the history of the process. As a consequence for a Markovian process the change of the process over a finite time can be built up from changes in infinitesimal time steps. In particular, from the knowledge of the transition probability density \( P(x, t|x_0, t_0) \) for infinitesimally closed times \( t > t_0 \), the transition probability density at finitely separated times \( t > t_0 \) can be constructed. With a starting probability \( P(x_0, t_0) \) all multitime probabilities then can be constructed.

The time evolution of the transition probability density can be considered in two ways. One can either fix the condition \( x_0 \) and propagate the time \( t \) in the forward direction, or, alternatively, one fixes \( x \) at \( t \) and propagates from there to backward in time to the past. Both equations of motion, the forward and the backward equation are based on the Chapman-Kolmogorov equation which represents a necessary condition for a process to be Markovian. It states that the transitions probability
density to reach $x$ at time $t$ from $x_0$ at time $t_0$ can be split in two steps from $x_0, t_0$ to $y, s$ and from there to the final state $x, t$. The time $s$ can be taken any value $t_0 \leq s \leq t$ and the intermediate points $y$ have to be summed over all available states:

$$P(x, t|x_0, t_0) = \int dy P(x, t|y, s) P(y, s|x_0, t_0).$$  \hfill (2.201)

Fixing now $x_0$ and $t_0$ we calculate the change of transition probability density between $t$ and $t + dt$, $dt > 0$

$$P(x, t+dt|x_0, t_0) - P(x, t|x_0, t_0) = \int dy \left[ P(x, t+dt|y, t) - \delta(x-y) \right] P(y, t|x_0, t_0).$$  \hfill (2.202)

One can show under very general conditions that the right hand side is proportional to $dt$ for small time steps. Hence, we obtain the forward equation

$$\frac{\partial}{\partial t} P(x, t|x_0, t_0) = \mathbf{L}_x(t) P(x, t|x_0, t_0)$$

$$P(x, t_0|x_0, t_0) = \delta(x-x_0)$$  \hfill (2.203)

where $\mathbf{L}(t)$ denotes the forward operator

$$\mathbf{L}(t) \rho(x) = \lim_{dt \to 0} \frac{1}{dt} \left( \int dy P(x, t+dt|y, t) \rho(y) - \rho(x) \right)$$  \hfill (2.204)

We have affixed the index $x$ to the forward operator in the eq. (2.203) to indicate that $\mathbf{L}$ here acts on the forward variable.

Before we further evaluate possible forms of the forward operator we proceed in an analogous way and derive the backward equation. Considering now the increment $P(x, t|x_0, t_0) - P(x, t|x_0, t_0 - dt_0)$ we find

$$\frac{\partial}{\partial t_0} P(x, t|x_0, t_0) = \mathbf{L}^+_x(t_0) P(x, t|x_0, t_0)$$

$$P(x, t_0|x_0, t_0) = \delta(x-x_0)$$  \hfill (2.205)

where $\mathbf{L}^+(t)$ is the backward operator:

$$\mathbf{L}^+(t) f(x) = \lim_{dt \to 0} \frac{1}{dt} \left( \int dy P(y, t|x, t - dt) f(y) - f(x) \right).$$  \hfill (2.206)

Note that in the integral kernel of the forward and backward equation the arguments $x$ and $y$ are interchanged. Hence, the backward equation is the adjoint operator of the forward operator relative to the scalar product

$$(f, \rho) = \int dx f(x) \rho(x)$$  \hfill (2.207)

i.e.

$$(f, \mathbf{L}(t)\rho) = (\mathbf{L}^+(t)f, \rho).$$  \hfill (2.208)
A general property of the forward and backward equation is the conservation of total probability. In terms of the forward operator one thus has

$$\int dx L(t) \rho(x) = 0$$  \hspace{1cm} (2.209)

for all integrable functions $\rho(x)$, and in terms of the backward operator

$$L^+(t) 1 = 0.$$  \hspace{1cm} (2.210)

Both equations follow from the definitions of the forward and the backward operators (2.204) and (2.206), respectively.

Now we discuss the possible explicit forms of the forward and the backward operator. First we consider pure jump processes which sit for a random time in some state and then hop over some finite distance to another state. This process by definition has no continuous component.

We denote the probability for a jump from $x$ to $y$ taking place in the interval $[t, t + dt]$ by

$$r(x, y, t) dt = P(x, t + dt | y, t).$$  \hspace{1cm} (2.211)

In the case of a pure jump process the jump rates $r(x, y, t)$ completely determine the forward and the backward operator:

$$L(t) \rho(x) = \int dy r(x, y, t) \rho(y) - \int dy r(y, x, t) \rho(x)$$  \hspace{1cm} (2.212)

$$L^+(t) f(x) = \int dy r(y, x, t) f(y) - \int dy r(y, x, t) f(x).$$  \hspace{1cm} (2.213)

The first terms in the equations for $L$ and $L^+$ result from the first terms the eqs. (2.204,2.206) for $x \neq y$. The second, negative terms take care of the conservation of the total probability as given by the eqs.(2.209,2.210). The form of the forward equation for a pure jump process is analogous to a Master equation of a Markovian process with discrete phase space: the time rate of change of the probability density $\rho(x, t)$ at time $t$ consists of an increase resulting from jumps into the state $x$, $\int dy r(y, x, t) \rho(x, t)$, and a decrease because of jumps out of the state $x$, $\int dy r(y, x, t) \rho(x, t)$.

In order to observe the decay law of the state $x$ one has to prevent the process to again visit the state $x$ after it has left it. I. e., the backward equation simplifies to

$$- \frac{\partial}{\partial t_0} P(x, t | x_0, t_0) = - \kappa(x_0, t_0) P(x, t | x_0, t_0)$$  \hspace{1cm} (2.214)

where

$$\kappa(x_0, t_0) = \int dy r(y, x_0, t_0)$$  \hspace{1cm} (2.215)
denotes the total rate out of the state $x_0$. The integral of $P(x, t|x_0, t_0)$ over an
infinitesimally small region $G_{E}(x_0)$ containing the point $x_0$ gives the probability
$\rho_{x_0}(t, t_0) = \int dx P(x, t|x_0, t_0)$ that the process has not left the state $x_0$ up to time $t$
once it has been there at time $t_0$. From the backward equation one finds an equation
of motion for $\rho_{x_0}(t, t_0)$

$$\frac{\partial}{\partial t_0} \rho_{x_0}(t, t_0) = \kappa(x_0, t_0)\rho_{x_0}(t, t_0)$$

(2.216)

with the final condition

$$\rho_{x_0}(t, t) = 1.$$ \hfill (2.217)

This equation is readily integrated to yield

$$\rho_{x_0}(t, t_0) = \exp \left( - \int_{t_0}^{t} ds \kappa(x_0, s) \right).$$ \hfill (2.218)

In the case of a time homogeneous process the transition rates $r(x, y, t) = r(x, y)$
are independent of time and an exponential law for the lifetime of a state of a
Markovian jump process results:

$$\rho_{x_0}(t, t_0) = \exp \left( -\kappa(x_0)(t - t_0) \right).$$ \hfill (2.219)

where $\kappa(x_0) = \int dy r(y, x_0)$.

We note, that the process once it has reached the state $x$ will jump to another
state $y$ is given by

$$q(y, x, t) = \frac{r(y, x, t)}{\kappa(x, t)}.$$ \hfill (2.220)

The other extreme case of a Markovian process is one that moves continuously.
For such a process the probability for a change of the process $x$ in $dt$ vanishes faster
than $dt$ and consequently the jump rates $r(x, y, t)$ for $x \neq y$ are zero. In other words,
the transition probability density $P(x + u, t|x, t - dt)$ is a very narrow function in $u$
for small $dt$ and therefore we can represent the backward operator in the following
way:

$$L^+ (t) f(x) = \lim_{dt \to 0} \frac{1}{dt} \int du P(x + u, t|x, t - dt) f(x + u) - f(x)$$

$$= \sum_{n=1}^{\infty} \frac{1}{n!} K_n(x, t) \frac{\partial^n}{\partial x^n} f(x)$$

(2.221)

where we have expanded the function $f(x + u)$about $x$ in a Taylor series. The
coefficients $K_n(x, t)$ are the $n$th conditioned moments per time of the distance that
The Fokker-Planck equation

75

the process covers in the infinitesimal time $dt$:

$$K_n(x,t) = \lim_{dt \to 0} \frac{1}{dt} \int du u^n P(x + u, t|x, t - dt)$$

$$= \lim_{dt \to 0} \frac{(dx^n)_{x(t) = x}}{dt}$$

(2.222)

where the second line is a shorthand.

Because of the backward operator is the adjoint of the forward operator one immediately obtains for $L$

$$L(t)\rho(x) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial x^n} K_n(x,t) \rho(x)$$

(2.223)

Eqs. (2.221) and (2.223) are known as the Kramers-Moyal expansion of the backward and the forward operator, respectively and the conditional moments $K_n(x,t)$ as the Kramers-Moyal moments.

For a continuous process one may expect that only a finite number of the Kramers Moyal moments are different from zero because the existence of a high finite moment would indicate a large spreading of the process within a short time. Indeed, a theorem by Pawula states that if a Kramers Moyal moment of even order larger than 3 vanishes all Kramers Moyal moments except the first and the second ones vanish as well. Only those Markovian processes have continuous trajectories for which the forward and the backward operators are differential operators of at most second order. They are known as diffusion processes and the corresponding forward equation is a Fokker-Planck equation characterized by the Fokker-Planck operator

$$L(t) = -\frac{\partial}{\partial x} K_1(x,t) + \frac{1}{2} \frac{\partial^2}{\partial x^2} K_2(x,t).$$

(2.224)

By definition $K_2(x,t)$ is a nonnegative function and is now as the diffusion coefficient, whereas $K_1(x,t)$ is the drift. Before we discuss the connection of the Fokker-Planck equation and Markovian processes described by a Langevin equation we note, that a general forward operators of a Markovian process consists of a sum of a Fokker-Planck operator describing the continuous part of the motion and an integral operator as given in eq. (2.212) describing the contribution of the jumps.

Finally we note that the generalization to processes in more than one dimension is straightforward: the integrations in forward and backward equations of jump-type extend over the whole available $n$-dimensional state space. In diffusion equations the spatial derivatives became partial derivatives with respect to the different coordinates in phase space with the drift vector $\vec{K}(\vec{x},t) = \{K_1(x_1, \ldots, x_n, t)\}$ and the diffusion matrix $\vec{D}(\vec{x},t) = \{D_{i,j}(x_1, \ldots, x_n, t)\}$. For example, the Fokker-Planck
operator in \( n \) dimensions becomes
\[
L(t) = -\sum_{i=1}^{n} \frac{\partial}{\partial x_i} K_i(x_1, \ldots, x_n) + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j} D_{i,j}(x_1, \ldots, x_n). \tag{2.225}
\]
the diffusion matrix is by definition nonnegative definite. The operator acts on the forward variables of the transition probability density of the \( n \)-dimensional stochastic process.

### 2.7.2 Moments of the transition probabilities

In this section we derive the Kramers-Moyal moments for a Langevin equation with Gaussian white noise as stochastic source term. In particular we show that all higher than the second order Kramers-Moyal moments vanish for this particular class of processes.

Our starting point is the stochastic differential equation
\[
\dot{x} = f(x,t) + g(x,t) \xi(t) \tag{2.226}
\]
for a single variable \( x(t) \) and \( \xi(t) \) is Gaussian white noise with intensity \( \epsilon \) and the increments \( dW(t) \) during \( dt \) (see eqs. (2.144-2.146)). For a small time interval \( dt \) the increment of the process is given by
\[
dx = x(t + dt) - x(t) = \int_t^{t+dt} f(x(s), s) \, ds + \int_t^{t+dt} g(x(s), s) \, dW(t) \tag{2.227}
\]
and the integral has to be understood as described after eq. (2.147).

One only needs to know \( dx \) for short times \( dt \) in linear order in \( dt \) because higher terms \( O(dt^2) \) do not contribute to the Kramers-Moyal moments. Because of the discontinuous behaviour of \( \xi(t) \) and the resulting non-differentiable increments of the Wiener process the second integral depends on the particular position of the points \( s \) in the interval \([t, t + dt]\) even if \( dt \) approaches zero. To show this rather strange behaviour we write \( s \) in terms of \( t \) and \( dt \) and a parameter \( q \in [0, 1] \):
\[
s = t + q \, dt = q \, (t + dt) + (1 - q) \, t. \tag{2.228}
\]
and express \( x(s) = x(t + q \, dt) \) in terms of \( x(t) \) and the increment \( dx \):
\[
x(s) = x(t) + q \, dx(t). \tag{2.229}
\]
Now we are ready to evaluate the integrals up to order \( dt \)
\[
dx(t) = \left[ f(x(t), t) + q \frac{\partial f(x(t), t)}{\partial x} \, dx(t) \right] dt + \left[ g(x(t), t) + q \frac{\partial g(x(t), t)}{\partial x} \, dx(t) \right] dW(t). \tag{2.230}
\]
The bistable oscillator

We reinsert $dx(t)$ on the right hand side and obtain in order $dt$:

$$dx(t) = \left[ f(x(t), t) + 2\epsilon q g(x(t), t) \frac{\partial g(x(t), t)}{\partial x} \right] dt + g(x(t), t) dW(t). \quad (2.231)$$

The additional term of order $dt$ that is proportional to $q$ results from the squared increment of the Wiener process:

$$(dW(t))^2 = 2dt + o(dt). \quad (2.232)$$

In mathematics the choice $q = 0$ is common. It is known as the Ito-interpretation of a stochastic differential equation, whereas physicists often use the Stratonovich interpretation corresponding to $q = 1/2$ (Stratonovich, 1990).

The conditioned moments of $dx(t)$ follow as:

$$\langle dx(t) \rangle = \left[ f(x(t), t) + 2\epsilon q g(x(t), t) \frac{\partial g(x(t), t)}{\partial x} \right] dt + o(dt) \quad (2.233)$$

$$\langle (dx(t))^2 \rangle = 2\epsilon g^2(x(t), t) dt + o(dt) \quad (2.234)$$

$$\langle (dx(t))^n \rangle = o(dt), \quad n \geq 3. \quad (2.235)$$

This gives for the first Kramers-Moyal moment:

$$K_1(x, t) = \frac{\langle dx \rangle}{dt} = f(x, t) + 2Dq \frac{\partial g(x, t)}{\partial x} g(x, t) \quad (2.236)$$

still depending on $q$. For the second moment one finds

$$K_2(x, t) = \frac{\langle dx^2 \rangle}{dt} = 2Dg^2(x, t) \quad (2.237)$$

and the higher moments starting from $n \geq 3$ vanish, i.e. $K_n(x, t) = 0$. Hence to a stochastic differential equation with a prescribed interpretation, i.e. a fixed value of $q$, there corresponds a Fokker-Planck equation and vice versa to each Fokker-Planck equation a Langevin equation with a specified interpretation may be assigned.

2.8 The bistable oscillator

In the stochastic theory, the bistable oscillator coupled to a thermal bath is one of the best investigated nonlinear models. A mechanical realization is a Brownian particle of mass $m$ moving in a potential $U(x)$ with two local minima. The bath exerts a linear damping force and a Gaussian white random force. The dynamics of the bistable oscillator then is described by a Langevin equation which can be written as

$$\frac{dx}{dt} = v, \quad \frac{dv}{dt} = -\gamma v - \frac{\partial U(x)}{\partial x} + \sqrt{2\gamma k_B T} \zeta(t) \quad (2.238)$$
where $\xi(t)$ is a normalized Gaussian white random force with $\langle \xi(t) \rangle = 0$ and $\langle \xi(t) \xi(s) \rangle = \delta(t-s)$. The strength of the random force is related to the damping constant $\gamma$ and the temperature $T$ of the bath by the Einstein relation.

As a particular example we take a symmetric quartic potential

$$U(x) = -\frac{1}{2}x^2 + \frac{b}{4}x^4, \quad a, b > 0$$  \hspace{1cm} (2.239)

that gives rise to the force

$$f(x) = -\frac{\partial U(x)}{\partial x} = ax - bx^3.$$  \hspace{1cm} (2.240)

If we neglect the random force but keep the damping, the resulting deterministic dynamics has two stable stationary points in phase space with coordinates

$$v_0 = 0, \quad x_{0,1} = \pm \sqrt{\frac{a}{b}}, \quad a, b > 0,$$  \hspace{1cm} (2.241)

and a hyperbolic point at the origin:

$$v_0 = 0, \quad x_b = 0.$$  \hspace{1cm} (2.242)

At weak damping, i.e. if $\gamma < 8a$, the stable points are foci and else nodes. These points coincide with the stationary points of the energy of the particle, $E = mv^2/2 + U(x)$. The stable points correspond to the two equally high minima of the energy and the hyperbolic point to a saddle point. There, the energy barrier separating the two wells is lowest. The energy difference between the saddle point and the minima is $\Delta U = a^2/(4b)$. In the deterministic approximation, the separatrix of the basins of attractions of the locally stable points is formed by the stable manifold of the hyperbolic point. The shape of the two basins of attraction is shown in fig.(2.3) for different values of $\gamma$. It resembles two tadpoles with infinitely long tails wrapped around each other.
The bistable oscillator

Fig. 2.3 The separatix between the domains of attraction of the locally stable points \( v = 0, x = \pm 1 \) for the motion of a particle of mass \( m = 1 \) in the potential \( U(x) = x^4/4 - x^2/2 \) for different values of the friction constant, \( \gamma = 0.1, 1, 3 \) in panels (a), (b) and (c), respectively. Note that the ratio of the velocity and friction scales is the same in the different panels but that the absolute scales differ.

Without the random force, the damping leads to a permanent loss of energy that is given by

\[
\frac{d}{dt} E = \frac{d}{dt} \left( \frac{m}{2} v^2 + U(x) \right) = -\gamma v^2 \leq 0.
\]  

It vanishes at the fixed points where \( v = 0 \).

When the damping constant is large, \( \gamma \to \infty \), or the particle is light, \( m \to 0 \), the inertial force becomes negligibly small compared to the damping and the potential force. Consequently, the velocity can be adiabatically eliminated and the particle moves according to

\[
\frac{dx}{dt} = -\frac{1}{\gamma} \frac{\partial U}{\partial x} + \sqrt{\frac{2k_B T}{\gamma}} \xi(t),
\]

where we again took into account the fluctuating force and a general (bistable) potential.

So far we have neglected the influence of the random force that is caused by the bath. It typically drives the particle out of the stable states. For small noise, however, the average typical deviation of the particle from the stable points is small compared to the distance between the stable points and the hyperbolic fixed point. On the other hand, there are realizations of the random force that drive the particle from one stable state over the saddle point to the other stable state. We will see that these events are rare if the thermal energy is small compared to the energy barrier separating the two stable states, but, also that they occur with certainty. Strictly speaking, an arbitrarily small white Gaussian random force destabilizes
the formerly stable states of a dynamical system and renders them metastable.
The bistable oscillator is just one of the simplest examples of this widespread and
important effect. It has many different applications in physics, chemistry, biology
and technical sciences.

To name an example we consider a molecule that may exist in two different
configurations A and B. In a solvent, transitions between these two forms may
occur:

\[ A = B. \] (2.245)

In the energy landscape of the molecule, the two configurations A and B correspond
to two minima which are separated by a barrier. The interaction with solvent
molecules provides the necessary energy to overcome this hindrance. The reaction
coordinate that leads over the barrier where it is lowest can be interpreted as a
particle’s coordinate. Assuming that the solvent acts with many weak collisions on
a fast time scale, the reaction can be described by a Langevin equation as given
by eq. (2.238) where \( U(x) \) is a conveniently defined potential of mean force. Other
examples of bistable behavior are more complicated chemical reactions, optical flip-
flop-devices, and optically bistable systems. Characteristic for all these cases and
many others is the passage of a saddle point that acts as a bottleneck for the
dynamics.

Before we discuss the escape from a metastable state in more detail we give the
Fokker-Planck equation for the probability density \( P(x,v,t) \) in phase space as it
follows from the Langevin equations (2.238):

\[
\frac{\partial}{\partial t} + v \frac{\partial}{\partial x} - \frac{1}{m} \frac{\partial U(x)}{\partial x} \frac{\partial}{\partial v} P(x,v,t) = \left\{ \gamma \frac{\partial}{\partial v} v + \frac{\gamma k_B T}{m^2} \frac{\partial^2}{\partial v^2} \right\} P(x,v,t).
\] (2.246)

This equation is also named after Klein who derived it first and Kramers. The left
hand side of the Klein-Kramers equation represents the reversible flow of probability
in phase space while the right hand side describes the irreversible effects caused by
the interaction with the heat bath. It has the form of a Stossintegral in the particular
limit of infinitely frequent and vanishingly small collisions. For a potential that
increases sufficiently fast for \( x \to \pm \infty \) the Klein Kramers equation has a uniquely
defined stationary solution that can be normalized to unity on the total phase space.
It coincides with the Maxwell Boltzmann distribution:

\[
P^{eq}(x,v) = N^{-1} \exp \left( -\frac{m}{2k_B T} v^2 - \frac{U(x)}{k_B T} \right). \] (2.247)

The fact that this probability density is approached from any initial distribution
clearly demonstrates that transitions between the metastable states must take place.
Otherwise the statistical weights that one can attribute to each metastable state
(1/2 for a symmetric double well potential) could not reach their equilibrium values
in accordance with the Maxwell-Boltzmann distribution (2.247). This also shows
The escape problem

that only the presence of noise brings into play the full nonlinearity of which otherwise the system would not pay much attention.

Because the Langevin equation defines an ergodic process we can infer ratios of the times that the system dwells in different states from the corresponding ratios of the ensemble density. Therefore, the ratio of times $t_{\text{stable}}$ and $t_{\text{saddle}}$ that the system dwells in equally large phase space regions of the size $dx dv$ at the stable point $(v_0, x_1)$ and the saddle point $(v_0, x_2)$, respectively, equals the ratio of the Maxwell Boltzmann distribution at the corresponding locations:

$$\frac{t_{\text{saddle}}}{t_{\text{stable}}} = \frac{P_{\text{eq}}(x = x_2, v = v_0)}{P_{\text{eq}}(x = x_1, v = v_0)} = \exp\left(-\frac{\Delta U}{k_B T}\right),$$

(2.248)

where $\Delta U = U(x_2) - U(x_1) = a^2/(4b)$ denotes the height of the energy barrier as seen from the bottom of either well. This corroborates our previous claim that for high barriers ($\Delta U \ll k_B T$), transitions between metastable states are rare events and, moreover, points to the importance of the Arrhenius factor $\exp\{-\Delta U/k_B T\}$ as a smallness parameter in this problem.

2.9 The escape problem

So far we have seen that Gaussian noise, how small it be, leads to the destabilization of deterministically stable states. It is quite natural to assume that the decay of a metastable state follows an exponential law which consequently is uniquely characterized by a rate constant $k$. This was verified theoretically, numerically and experimentally in a wide range of situations. Hereby it is assumed that once the system has left its initial metastable state this state will not be repopulated. This, of course, requires certain modifications of the process that will be discussed below.

If we again consider the bistable oscillator as an example, we can ask how the probabilities develop in time to find it in either of its metastable states. In the case of an exponential decay, this coarse grained dynamics is given by a simple master equation describing the transitions between the two metastable states:

$$\dot{p}_-(t) = -kp_-(t) + kp_+(t)$$
$$\dot{p}_+(t) = kp_-(t) - kp_+(t),$$

(2.249)

where $p_-(t)$ and $p_+(t)$ denote the probabilities to find the system in the domains of the attractors at $(v_0, x_0)$ and $(v_0, x_1)$, respectively. Moreover, we assume a symmetric potential. Therefore the rates out of the metastable states are equal. They are denoted by $k$.

Within this simplified picture, still different physical processes can be described which correspond to particular experiments from which the rate can be determined. The decay of a single metastable state, say the one at $(v_0, x_0)$, is an example that was already mentioned. It can be observed if the back flow from the other metastable
state is prevented, i.e. if the second state is changed into an absorbing state. The same effect is reached by setting \( p_+ (t) = 0 \). The master equation then simplifies to

\[
\dot{p}_{\text{abs}} (t) = - kp_{\text{abs}} (t), \tag{2.250}
\]

yielding \( p_{\text{abs}} (t) = \exp \{- kt\} \) for the decay of the population of the initial state. This population gives the probability that the system has not left the initial state up to the time \( t \) and therefore \( p_{\text{abs}} (t) \) is also called the waiting time distribution.

The probability density of exit times, \( \rho(t) \), follows as the negative derivative of the waiting time distribution, \( \rho(t) = k \exp \{- kt\} \), from which the moments of the exit time can be calculated:

\[
\langle t^n \rangle = \int_0^\infty dt t^n \rho(t) = n! k^{-n}. \tag{2.251}
\]

Of particular importance of course is the first moment, i.e. the mean exit time, which is given by the inverse rate.

\[ \langle t \rangle = 1/k. \tag{2.252} \]

In another experiment one allows for jumps in both directions and observes how a distribution initially localized in one state relaxes toward the equilibrium distribution. From the master equation the time dependence of this process readily follows:

\[
p_\pm (t) = \frac{1}{2} (1 \pm e^{-2kt}) \tag{2.253}
\]

Note that in this case the relaxation is characterized by the sum of the rates of both states resulting in \( 2k \) which is just the negative non-vanishing eigenvalue of the coefficient matrix on the right hand side of the master equation. According to the regression theorem one finds the same relaxation law for equilibrium correlations of the population. This fact is used in the reactive flux method which is an effective way to numerically determine rates.

One can also think of a non-equilibrium situation in which a steady current is maintained by a source that emits particles at a constant rate in one state, and a sink that instantly removes particles arriving at the other state. The probability current \( j \) that flows, say, from the left replenished to the right absorbing state, then follows as \( j = kp_- \) resulting in the flux-over-population formula for the rate:

\[
k = \frac{j}{n}, \tag{2.254}
\]

where \( n = p_- \) denotes the population of the replenished state.

These different situations can also be described in terms of the more detailed model of the Langevin equation, or the equivalent Fokker-Planck equation. We will shortly describe these methods before we discuss some of them in more detail in the following sections.
The escape problem

In order to determine the rate from the decay of the metastable state one again has to modify the process in such a way that back reactions are excluded. This conveniently is done by introducing an absorbing boundary enclosing the final state. The mean first passage time of the boundary then gives the inverse of the rate provided the absorbing boundary is close enough to the final state.

The decay of a non-equilibrium initial state according to the Fokker-Planck dynamics seemingly is much more complicated than the same process described by the master equation because the Fokker-Planck operator has infinitely many eigenvalues each of which gives rise to an exponentially decaying component of the probability density. The very fact, however, that the transitions between the metastable states occur only rarely, also shows itself in the spectrum of the Fokker-Planck operator: It has a pronounced gap separating two eigenvalues from the rest of the spectrum. One of these two eigenvalues is zero, \( \lambda_0 = 0 \) corresponding to the equilibrium solution of the Fokker-Planck equation; the other one is negative and has a small absolute value. The corresponding eigenfunction has a node line leading through the saddle point and in the neighborhood of the stable states its absolute value coincides with the Maxwell-Boltzmann distribution. All eigenvalues belonging to the rest of the spectrum possess a negative real part with an absolute value that is much larger than the one of \( \lambda_1 \): \( |\Re \lambda_k| \gg |\lambda_1|, \ k > 1 \). This mathematically reflects the time scale separation: The “large” eigenvalues describe the fast relaxations that take place within each of the domains of attraction and the “small” one, \( \lambda_1 \), is responsible for the transitions between the metastable states. As for the master equation the small eigenvalue \( \lambda_1 \) is the negative sum of the rates out of the two metastable states. There exist effective analytical and numerical methods to determine the small eigenvalue of a Fokker-Planck operator that describes a system with metastable states.

At the first sight it seems very complicated to use the flux-over-population method for the Fokker-Planck equation: In principle, one has to solve the stationary Fokker-Planck equation that is complemented by source and sink terms. This problem was circumvented by Kramers in a very elegant way: Rather than specifying particular sources and sinks he constructed a current carrying probability density that fulfills three requirements: First, that the sources maintain the Maxwell-Boltzmann distribution in the initial well, second, that the sinks lead to a vanishing distribution in the other well, and, third, that the region close to the saddle point is free from sources and sinks. In this region the potential force can be linearized and the resulting Fokker-Planck equation can be solved with asymptotic boundary conditions so that also the first two conditions are satisfied.

Before we further discuss this approach we present a rough estimate of the rate that has the virtue of providing an upper bound for the rate.
2.9.1 Transition state theory

The transition state theory makes use of the flux-over-population formula of the rate. Flux and population are calculated from the Maxwell-Boltzmann distribution by only taking into account positive velocities at the barrier which for the sake of simplicity is assumed to be located at $x_b = 0$. With these assumptions one finds for the flux:

$$j^{TST} = \int_{0}^{\infty} dv \, v \, P^{eq}(0, v) = \frac{k_B T}{m N} \frac{U(0)}{k_B T}$$

(2.255)

and for the population of the initial (left) well:

$$n = \int_{-\infty}^{\infty} dv \int_{0}^{\infty} dx \, P^{eq}(x, v) \approx \frac{2 \pi k_B T}{m \omega_0^2 N} \frac{\nu(0)}{\pi n^2}$$

(2.256)

where we have used a parabolic approximation of the potential near the initial stable state at $x = x_0$ and extended the integration to infinity. Both approximations are justified for low temperatures and for a symmetric potential. They are controlled by the smallness of the parameter $U^\prime(0)(x_0) k_B T / (8(U''(x_0))^2) = 3 k_B T / \Delta U$. Here $U''(x_0) = m \omega_0^2 = 2a$ and $U^\prime(0)(x_0)$ denote the second and fourth derivative, respectively.

With the flux-over-population formula one obtains the rate of transition state theory (TST):

$$k^{TST} = \frac{\omega_0}{2 \pi} e^{-\frac{\Delta U}{k_B T}}$$

(2.257)

This result has a simple interpretation: The frequency $\omega_0 / 2 \pi$ gives the number of attempts per time to overcome the barrier and the Arrhenius factor the fraction of successful approaches. The TST rate is based on the two assumptions that, first, the positive velocities are thermally distributed even at the barrier and, second, that there are no particles with negative velocities there. The first assumption requires a rather strong interaction of the particle with the heat bath such that the thermal distribution can be maintained even though particles escape, whereas the second assumption requires that the interaction with the bath is sufficiently weak in order that a particle that already has passed the barrier is not scattered back by the random force exerted by the bath. Both assumptions overestimate the probability flux at the barrier. Therefore the transition state theory represents an upper bound of the true rate:

$$k^{TST} \leq k.$$  

(2.258)

Transition state theory has widely been used in physics and chemistry. Often a considerable improvement of the rate can be achieved by taking into account other degrees of freedom than the reaction coordinate. The generalization of transition
state theory to multidimensional problems is also based on the flux over population expression for the rate. In the multidimensional phase space a dividing surface is introduced with the “reactants” on the one and the “products” on the other side of this surface, i.e. the reaction coordinate has to cross the dividing surface. The system is assumed to be in a thermodynamic equilibrium state fixing the probability distribution in phase space, as e.g. the canonical distribution. The population of reactants is determined by the integral over the respective part of phase space with the dividing surface as boundary, and the flux follows as the unidirectional probability current leading from reactants to products through the dividing surface. This again gives an upper bound for the rate that now may come very close to the true rate provided that those degrees of freedom are included that interact with the reaction coordinate and that the dividing surface is properly chosen. A variation of the dividing surface generally will lead to a change of the value of the TST rate. Together with the bounding property of the rate, eq. (2.258), one obtains a variational principle for the rate. This is the basis of the variational transition state theory.

A related question which is of relevance in signal communication and for many other technical and biological processes concerns the frequency with which a signal crosses a threshold \( x_s \). Under the assumption that the signal is stationary and Gaussian distributed with average value zero and that the velocity of the process can be defined and has a finite variance, Rice (1944) determined the crossing rate of the threshold at \( x = x_s \) as:

\[
   k_{Rice} = \frac{1}{2\pi} \frac{\psi_{x,v}(0)}{\psi_{x,x}(0)} \exp \left( -\frac{x_s^2}{\psi_{x,x}(0)} \right),
\]

(2.259)

where \( \psi_{x,x}(t) = \langle x(t) x \rangle \) and \( \psi_{x,v}(t) = \langle v(t) v(0) \rangle = -d^2 \psi_{x,x}(t)/dt^2 \) denote the correlation function of the signal and its velocity, respectively. The exponential factor has its counterpart in the Arrhenius factor of the transition state rate. Moreover, Rice showed that the prefactor just coincides with the expected number of zero crossings per second of the signal \( x \). Hence, also the prefactor has an analogous meaning as the one of transition state theory.

It is remarkable that no further details of the process enter than the variances of the signal itself and its velocity. These quantities also can be expressed in terms of the spectrum of the process.

2.9.2 Kramers’ rate formulae

In his pioneering work of 1940 Kramers derived expressions for the rate in the two cases of extremely weak damping and moderate to strong damping. It took almost half a century until Melnikov and Meshkov (1986) presented a crossover theory from the underdamped to a still weakly damped regime where transition state theory applies, and until Pollak, Grabert and Hänggi covered the full regime from weak
to strong damping. Here we will restrict ourselves to those cases originally treated by Kramers. For a discussion of the crossover theory we refer to the literature (Hänggi, Talkner, Borkovec, 1990; Melnikov, 1991). The model that was discussed by Kramers is that of the bistable oscillator, see eq. (2.238).

2.9.2.1 Moderate to strong damping

If the damping is sufficiently strong a typical escape path will cross the separatrix in the vicinity of the saddle point at \((v_0 = 0, x_b)\). Following Kramers we construct a probability density \(\rho(x,v)\) that in this region close to the saddle point is a solution of the stationary Fokker-Planck equation,

\[
L\rho(x,v) = 0 ,
\]

and that approaches the Maxwell Boltzmann distribution in the reactant well and vanishes in the product well. As explained above this amounts to the construction of a stationary flux carrying solution maintained by sources repopulating the reactant state and sinks taking out the products. Moreover, neither the sources nor the sinks must come too close to the saddle point.

In order to take into account the prescribed behavior in the two wells, we split off the Maxwell-Boltzmann distribution:

\[
\rho(x,v) = h(x,v)\rho_{eq}(x,v)
\]

where \(h(x,v)\) is a form function that approaches unity at the side of the reactants and zero at products:

\[
h(x,v) = \begin{cases} 
1 & \text{at reactants}, \\
0 & \text{at products}.
\end{cases}
\]

Finally, as a consequence of eq. (2.260), in the region of the saddle point, the form function is a solution of the modified Fokker-Planck equation:

\[
-v \frac{\partial h(x,v)}{\partial x} - \left[ \omega_b^2 x + \frac{\gamma}{m} v \right] \frac{\partial h(x,v)}{\partial v} + \frac{\gamma k_B T}{m^2} \frac{\partial^2 h(x,v)}{\partial v^2} = 0 .
\]

Because here the coordinate \(x\) is restricted to the close vicinity of the saddle point we have approximated the potential by a parabola:

\[
U(x) \approx U(0) - \frac{1}{2} m \omega_b^2 x^2 ,
\]

where \(m \omega_b^2 = -\partial^2 U(x_b)/\partial x^2\) is given by the second derivative of the potential at the barrier.

This partial differential equation reduces to an ordinary one for \(\zeta(u) = h(x,v)\) where \(u\) is a linear combination of position and velocity

\[
u + \left( -\frac{\omega_b^2}{\lambda} \right) v .
\]
The escape problem

Here the coefficient \( \lambda \) coincides with the positive Lyapunov exponent of the deterministic dynamics at the saddle point:

\[
\lambda = -\frac{\gamma}{2m} + \sqrt{\left(\frac{\gamma}{2m}\right)^2 + \omega_b^2}.
\]  
(2.266)

The differential equation for the auxiliary function \( \zeta(u) \) takes the form:

\[
ud\frac{d\zeta(u)}{du} + \frac{\gamma \lambda k_B T m^2 \omega_b^2}{d^2\zeta(u)} = 0.
\]  
(2.267)

It is readily solved and, together with the boundary conditions (2.9.2.1), yields for the form function:

\[
h(x, v) = \frac{1}{\sqrt{2\pi}} \int_{\frac{x}{\sqrt{2\omega_b}} - \gamma}^{\infty} e^{-z^2/2} dz.
\]  
(2.268)

We note that one also can reduce eq. (2.263) to an ordinary differential equation by using the second (negative) Lyapunov exponent in eq. (2.267). With the resulting solutions, however, one cannot satisfy the boundary conditions (2.9.2.1).

In contrast to transition state theory one now determines the net flux over the barrier rather than the unidirectional flux:

\[
j = \int_{-\infty}^{\infty} dv \psi(0, v) = \frac{\lambda k_B T m N e^{-\frac{v(0)}{\sqrt{2\omega_b}}}}{\omega_b m N}.
\]  
(2.269)

It differs from the flux obtained in the transition state theory by the “transmission” factor \( \lambda/\omega_b \). It is smaller than unity, approaches unity for \( \gamma/m \to 0 \), zero for \( \gamma/m \to \infty \), and takes into account that particles that have crossed the boundary are possibly scattered back and therefore must not be counted as successful escape events. The population of the reactant well is given by the same expression as in the transition state theory because by construction the flux carrying probability density there coincides with the Maxwell-Boltzmann distribution.

As a final result, one obtains for the rate:

\[
k_{Kramers} = \frac{\lambda}{\omega_b} k_{TST}.
\]  
(2.270)

In the limit of weak damping the result of transition state theory is reproduced. We will come back to this limiting behavior. The stronger the damping is, the more important becomes the influence of backscattering on the rate which leads to an asymptotic decay of the rate for large damping proportional to \( 1/\gamma \):

\[
k_{Kramers}^{\gamma \to \infty} = \frac{m \omega_b}{\gamma} k_{TST} \quad \text{for} \quad \frac{\gamma}{m \omega_b} \to \infty.
\]  
(2.271)

This result also directly follows from the Smoluchowski equation as we will see below.
The Kramers rate itself is an asymptotic result that becomes exact in the limit of high barriers, i.e. for $\Delta U/k_B T \to \infty$. There are two types of corrections to this asymptotic behavior: Analytic ones in $k_B T/\Delta U$ and non-analytic ones of the particular Arrhenius form $\exp \{-\Delta U/(k_B T)\}$. It turns out that, if the barrier height is so low that also the non-analytic corrections become relevant, the different approaches which we briefly discussed, will yield different results for the rate. This indicates that for too low barriers the rate picture starts failing. As a estimate when the rate description ceases to exist, one may consider a barrier height of $\Delta U/k_B T = 4$ that introduces a non-analytical error of roughly 1%. For barriers lower than this value a more detailed analysis of the considered system is necessary. For higher barriers the rate picture yields an adequate description but the analytical corrections may still be relevant. They can systematically be taken into account both in transition state theory (Pollak, Talkner, 1993) and on the level of the Fokker-Planck equation (Talkner, 1994).

2.9.2.2 Weak damping and energy diffusion

In the previous section we revised the assumptions of transition state theory related to backscattering. For a sufficiently large damping strength this indeed is the main mechanism that modifies the rate. Because the damping strength is not only a measure how effective any excess energy is taken from the system to the bath but also how effective the bath supplies the system with energy, the maintenance of local equilibrium on the time scale of the escape is guaranteed if the damping is sufficiently large. If it is small, the supply of energy becomes the relevant rate determining step. For that reason Kramers determined a reduced diffusion equation for the energy of the particle holding in the limit of vanishing damping:

$$\frac{\partial P(E,t)}{\partial t} = \frac{\gamma}{m} \frac{\partial}{\partial E} I(E) \left[ 1 + k_B T \frac{\partial}{\partial E} \right] \frac{\omega(E)}{2\pi} P(E,t), \quad (2.272)$$

where $I(E) = \int \sqrt{2m(E-U(x))} \, dx$ is the action and $\omega(E) = 2\pi \partial E/\partial I$ the frequency of the undamped system at the energy $E$. For this one dimensional diffusion equation one can find the exact stationary solution carrying the constant probability current $j$. Using the flux-over population expression one obtains for the rate

$$k_{\gamma=0} = \frac{I(E_b) \gamma}{mk_B T} T^{TST}, \quad (2.273)$$

where the population has been evaluated to leading order in $k_B T/\Delta U$. Here $I(E_b)$ is the action at the energy of the barrier. In contrast to the rate expression (2.9.2.1), the energy diffusion controlled rate vanishes with the damping constant. Obviously, when the factor in front of the transition state rate becomes of the order of unity, the energy diffusion rate becomes larger than the TST rate and can no longer be valid. We will not discuss the rather complicated theories that describe the crossover between the two regimes. For practical purposes it is often sufficient to use a simple
Pontryagin’s equation

Padé like interpolation formula:

\[ k_{\text{int}} = \left( k_{\text{Kramers}}^{-1} + k_{\gamma=0}^{-1} \right)^{-1} , \]

which may introduce an error of maximally 20%.

2.9.3 Transition rates in multidimensional landscapes

There are many situations which are described by a motion in a multidimensional energy landscape that cannot be reduced to a single coordinate. For systems that in the limit of long times approach a thermal equilibrium state and that obey the symmetry of detailed balance, transition rates between local minima of the energy can be determined in a way that is analogous to Kramers method described above. Landauer and Swanson (1961) determined the rates for a multidimensional overdamped system described by a Smoluchowski equation. Later, Langer (1969) also considered more general Fokker-Planck dynamics including the effect of inertia and also considered nucleation rates in field theories by performing a continuum limit. Here, we only will give general results for systems with a finite number of degrees of freedom.

The transition rate out of a metastable state over a barrier with an energy \( \Delta E \) above the initial well again is dominated by the Arrhenius factor \( \exp \left\{ -\frac{\Delta E}{k_B T} \right\} \).

The prefactor is given by the positive Lyapunov exponent \( \lambda \) of the deterministic motion at the saddle point multiplied by the square root of the ratio of the determinants of the second derivatives of the energy (the Hessians) at the initial energy minimum and at the barrier, \( H_0 = \left( \frac{\partial^2 E(x_0)}{\partial x_i \partial x_j} \right) \) and \( H_b = \left( \frac{\partial^2 E(x_b)}{\partial x_i \partial x_j} \right) \), respectively:

\[
  k_{\text{Langer}} = \frac{\lambda}{2\pi} \sqrt{\frac{\det H_0}{\det H_b}} e^{-\Delta E/k_B T} .
\]

Note that the absolute value of the Hessian at the barrier has to be taken because it is always negative as a consequence of the unstable direction. Here, we only have considered the most simple cases of a point like initial state and a single, point like barrier. For further details and more general cases we refer to the literature (Hänggi, Talkner, Borkovec, 1990).

2.10 Pontryagin’s equation

Several more specialized tools exist for Markov processes that allow one to determine certain relevant aspects of the process without the need of knowing the full conditional probability as the solution of either the forward or the backward equation. Often the introduction of boundaries and the modifications of the process at the boundaries are necessary for these methods. We first will introduce the most frequent types of boundary conditions and discuss how boundary conditions for the
forward equation can be related to ones for the backward equation and vice versa. Then we will discuss first passage times and so-called splitting probabilities.

### 2.10.1 Boundary conditions for the forward and the backward equation

For the sake of simplicity we restrict ourselves to one dimensional processes. Generalizations of the concepts discussed in this section to processes in higher dimensions are straightforward. Most exact solutions known in one dimension unfortunately do not simply translate to higher dimensions, but much of the qualitative behavior of particular solutions found in one dimension also applies in higher dimensions and often can be used as an inspiration and motivation for approximate solutions and particular ansätze.

We consider a time-homogeneous Markovian diffusion process \( x(t) \), i.e. the conditional probability \( P(x, t|y, s) \) of the process only depends on the time difference \( t - s \) and solves the forward and the backward equation:

\[
\begin{align*}
\frac{\partial}{\partial t} P(x, t|y) &= L_x P(x, t|y), \\
\frac{\partial}{\partial t} P(x, t|y) &= L_x^+ P(x, t|y),
\end{align*}
\]  

(2.276)

where \( L \) and \( L^+ \) denote the forward and backward Fokker-Planck operators, respectively. The indices of the respective operators indicate the variables on which they act: The state \( x \) at time \( t \) of observation in the forward and the condition \( y \) at the earlier time \( s = 0 \) in the backward equation. As we already have seen, the Fokker-Planck operator is a second order differential operator with coefficients \( K(x) \) and \( D(x) \) that characterize the drift and diffusion, respectively:

\[
\begin{align*}
L &= -\frac{\partial}{\partial x} K(x) + \frac{\partial^2}{\partial x^2} D(x), \\
L^+ &= K(x) \frac{\partial}{\partial x} + D(x) \frac{\partial^2}{\partial x^2}.
\end{align*}
\]  

(2.277)

In order to have a mathematically well defined problem the forward and backward equations have to be complemented by initial and boundary conditions. The initial condition is obvious:

\[
P(x, 0|y) = \delta(x - y).
\]  

(2.278)

The boundary conditions depend on the particular physical situation under consideration and require somewhat more thought. An important aspect here is that the forward and the backward operators are adjoint operators relative to each other:

\[
\int_G df(x) L \rho(x) = \int_G d\rho(x) L^+ \rho(x)
\]  

(2.279)
This has to hold for all admissible functions \( f(x) \) and \( \rho(x) \). It does not only determine the relation between the forward and the backward operator as it is evident from the eq. (2.277) but also the behavior of the admissible functions \( f(x) \) at the boundaries of the domain \( G \) once the properties of \( \rho(x) \) are fixed and vice versa.

As a first example for \( G \) we consider the interval \([y_1, y_2]\) with absorbing boundary conditions at \( y_1 \) and \( y_2 \). These are most naturally characterized in terms of the conditional variable of the transition probability by stating that no transitions may take place from the boundaries. This leads to the following conditions for the backward equation describing absorption:

\[
P(x, t|y_1) = P(x, t|y_2) = 0 \quad \text{for} \quad x \in (y_1, y_2).
\] (2.280)

In order to find the boundary conditions for the forward equation, one has to require that the condition (2.279) holds for all functions \( f(x) \) that vanish at the boundaries according to eq. (2.280). This is the case only if at the same time the function \( \rho(x) \) also vanishes at \( x = y_1 \) and \( x = y_2 \). Therefore, at the absorbing boundaries we find as boundary conditions for the forward equation:

\[
P(y_1, t|y) = P(y_2, t|y) = 0 \quad \text{for} \quad y \in (y_1, y_2).
\] (2.281)

As a second example we take the same interval \([y_1, y_2]\) as domain but now with reflecting boundaries at \( y_1 \) and \( y_2 \), i.e. every trajectory arriving from the interval at the boundary is send back to the interior of the interval. Hence, the probability fluxes through the boundaries vanish. This gives the reflecting boundary conditions for the forward operator:

\[
K(y_1) P(y_1, t|y) - \frac{\partial}{\partial x} [D(x) P(x, t|y)]_{x=y_1} = 0,
\]

\[
K(y_2) P(y_2, t|y) - \frac{\partial}{\partial x} [D(x) P(x, t|y)]_{x=y_2} = 0 \quad \text{for} \quad y \in (y_1, y_2).
\] (2.282)

Requiring such boundary conditions for the functions \( \rho(x) \) one finds the condition (2.279) satisfied only if the first derivative of the function \( f(x) \) vanishes at the boundaries. This leads to the reflecting boundary conditions for the backward equation:

\[
\frac{\partial P(x, t|y_1)}{\partial y} = \frac{\partial P(x, t|y_2)}{\partial y} = 0 \quad \text{for} \quad x \in (y_1, y_2).
\] (2.283)

So-called natural boundaries are never reached by a process in finite time. For the forward equation that amounts to the vanishing of the probability together with its first derivative. No condition follows for the backward equation in this case.

Finally we note that the left and the right boundary points may be of different character, e.g. \( y_1 \) may be reflecting and \( y_2 \) absorbing.
2.10.2 The first passage time distribution

The probability $\text{Prob}(x(s) \in G, s \leq t, x(0) = y) = W_G(t, y)$ that a process $x(t)$ has not left a certain region $G$ of its state space up to time $t$, in general depends on the starting point $x(0) = y \in G$. The probability $W_G(y, t)$ can be calculated from the transition probability $P_G(x, t|y)$ of a modified process that is stopped whenever the process has reached one of the boundaries of $G$. Before that happens it coincides with the original process. Consequently the conditional probability of the modified process fulfills the forward and backward equation of the original process on $G$ with absorbing boundary conditions at the boundaries of $G$. The waiting time distribution $W_G(y, t)$ coincides with the total amount of probability found in $G$ at time $t$:

$$W_G(y, t) = \int_G dx P_G(x, t|y) \quad (2.284)$$

We differentiate both sides with respect to time, use the backward equation for the time rate of change of the conditional probability, interchange the backward operator (acting on $y$) with the integral over $x$, and finally obtain the backward equation as the equation of motion of the waiting time probability:

$$\frac{\partial}{\partial t} W_G(y, t) = L^+ W_G(y, t) \quad \text{for } y \in G,$$

$$W_G(y, 0) = 1 \quad \text{for } y \in G,$$

$$W_G(y, t) = 0 \quad \text{for } y \in \partial G. \quad (2.285)$$

The initial and boundary conditions for $W_G(y, t)$ are direct consequences of the respective conditions for the conditional probability $P_G(x, t|y)$ in combination with the definition (2.284) of the waiting time distribution.

Once the waiting time distribution is known, the probability density $\rho_G(y, t)$ of the exit times follows as the negative derivative of $W_G(y, t)$ with respect to time:

$$\rho_G(y, t) = -\frac{\partial}{\partial t} W_G(y, t). \quad (2.286)$$

From the probability of exit times $\rho_G(y, t)$ moments of the first passage time follow:

$$\langle t^n(y) \rangle = \int_0^\infty dt t^n \rho_G(y, t). \quad (2.287)$$

Acting on both sides of this equation with the backward equation, using $L^+ \rho(y, t) = \partial \rho(y, t)/\partial t$ and integrating by parts one finds the following hierarchy of equations for the $n$th moments of the first passage time:

$$L^+ \langle t^n(y) \rangle = -n \langle t^{n-1}(y) \rangle$$

$$\langle t^n(y) \rangle = 0 \quad \text{for } y \in \partial G \quad (2.288)$$
where the absorbing boundary conditions follow from those of the waiting time distribution, see eq. (2.285). Moreover we have assumed that the moments exist. In particular, for the mean first passage time one obtains:

$$L^+(t(y)) = -1$$

$$\langle t(y) \rangle = 0 \quad \text{for} \quad y \in \partial G$$

(2.289)

This equation was derived by Pontryagin, Andronov, and Witt (1933) and is known as Pontryagin equation. An extension to higher dimensional cases can be find in (Weiss, 1967).

We mention that the same form of the Pontryagin equation also holds for Markovian processes in higher dimensions. The operator $L^+$ then denotes the backward operator of the considered process.

As a final comment we note that not all points of the boundary $\partial G$ need to be absorbing. For example, in the case of a one dimensional process with an interval as domain, one endpoint of the interval may be reflecting and the other one absorbing.

Before we discuss some examples for mean first passage times we introduce another quantity that characterizes a process in the presence of two absorbing boundaries.

### 2.10.3 Splitting probability

We again consider a one dimensional Fokker-Planck process $x(t)$ that starts at a point $y$ within an interval $[y_1, y_2]$. We pose the question with which probability the process will leave the interval at the boundary point $y_1$. This quantity is denoted by $\pi_1(y)$ and is called the splitting probability.

Obviously, it is sufficient to follow the process until it reaches either boundary and then to stop it. Hence, both boundaries are absorbing. The probability flowing out of the interval at $x = y_1$ per time is given by the probability current

$$j(y_1, t; y) = -\frac{\partial}{\partial x}[D(x)P(x, t|y)]_{x=y_1}.$$ 

(2.290)

Here the drift term goes not contribute to the probability current because it is proportional to the conditional probability $P(x, t|y)$ at the boundary $x = y_1$ which vanishes there.

The splitting probability results as the total probability flowing through $x = y_1$ and, hence is given by the integral of the respective probability current over all positive times:

$$\pi_1(y) = \int_0^\infty dt j(y_1, t; y).$$

(2.291)

Acting on both sides of this equation with the backward operator, interchanching it on the right hand side with both the time integration and the derivative with respect to $x$, and using the backward equation one obtains the difference of the
probability current \(j(y_1, t; y)\) at infinite and zero time. For \(y \in (y_1, y_2)\) it vanishes at both times and, hence, one finds as an equation for the splitting probability:

\[
L^+ \pi_1(y) = 0 \quad \text{for} \quad y \in (y_1, y_2),
\]

\[
\pi_6(y_1) = 1,
\]

\[
\pi_6(y_2) = 0. \tag{2.292}
\]

The boundary condition follow from the definition of the splitting probability in terms of the probability current (2.290) and its behavior at the boundaries. They have the simple interpretation that a particle that starts at the absorbing boundary \(y_1\) will never reach the other boundary \(y_2\) and vice versa.

For processes in higher dimensions that may leave a region \(G\) at different boundaries \(\partial G_i\), the splitting probability \(\pi_1(y)\) is defined as the relative frequency with which the particular boundary \(\partial G_1\) is reached before any other boundary \(\partial G_i, i \neq 1\) has been crossed. It satisfies the analogous equation (2.292) with the corresponding multidimensional backward operator. The boundary conditions are \(\pi_1(y) = 1\) for \(y \in \partial G_1\) and \(\pi_1(y) = 0\) for \(y \in \partial G_i, i \neq 1\).

### 2.10.4 Examples

In one dimensions the equations (2.289) and (2.292) for the mean first passage time and the splitting probability can be solved for arbitrary drift and diffusion.

#### 2.10.4.1 The splitting probability

The equation for the splitting probability out of the interval \([x_1, x_2]\) to the left boundary reads:

\[
\left[ K(y) \frac{d}{dy} + D(y) \frac{d^2}{dy^2} \right] \pi_1(y) = 0,
\]

\[
\pi_1(y_1) = 1, \quad \pi_1(y_2) = 0. \tag{2.293}
\]

It readily is solved to read:

\[
\pi_1(y) = \frac{\int_{y_1}^{y} dz \exp \{V(z)\}}{\int_{y_1}^{y_2} dz \exp \{V(z)\}}, \tag{2.294}
\]

where

\[
V(y) = -\int_{y_1}^{y} dz \frac{K(z)}{D(z)}. \tag{2.295}
\]
The other splitting probability to reach $y_2$ before $y_1$ is just the complement of $\pi_1(y)$ to one:

$$\pi_2(y) = \frac{\int_{y_1}^{y_2} dz \exp \{ V(z) \}}{\int_{y_1}^{y_2} dz \exp \{ V(z) \}} = 1 - \pi_1(y). \quad (2.296)$$

We specialize this general result to the two cases of pure diffusion and of the overdamped bistable oscillator of the previous section.

In the first case we find for vanishing drift, $K(y) = 0$, and constant diffusion $D(x) = D$ a linear dependence of the splitting probability on the starting point $y$:

$$\pi^{diff}_1(y) = \frac{y_2 - y}{y_2 - y_1}. \quad (2.297)$$

In the case of the overdamped oscillator we start the process in a point between the two metastable states and ask with which probability the right metastable state will be reached before the left one. This is given by the splitting probability

$$\pi_0(y) = \frac{\int_{x_0}^{x_1} dz \exp \left\{ \frac{U(z)}{k_B T} \right\}}{\int_{x_0}^{x_1} dz \exp \left\{ \frac{U(z)}{k_B T} \right\}}. \quad (2.298)$$

where, as defined above, $x_0$ and $x_1$ denote the metastable states to the left and the right of the barrier, respectively. The splitting probability connects the prescribed values at the boundaries in a monotonically decreasing way. For a symmetric potential it takes the value $1/2$ at the barrier. This point from where the system reaches both metastable states with equal probability is also called the stochastic separatrix. If the potential is not symmetric about the barrier the stochastic separatrix in general does not coincide with the deterministic one sitting on top of the barrier. Only in the limit of weak noise the stochastic separatrix approaches the deterministic one.

In this limit one can approximate the potential in eq. (2.298) by a parabola, $U(x) \approx U(x_b) - m\omega_b^2(x - x_b)^2/2$, extend the limits $x_0$ and $x_2$ to $-\infty$ and $\infty$, respectively, and perform the resulting Gaussian integrals. As result one obtains:

$$\pi_0(y) \approx \frac{1}{2} \text{erfc} \left( \frac{\sqrt{m\omega_b^2}}{2k_BT} y \right) \quad (2.299)$$

where $\text{erfc}(z)$ denotes the complementary error function. Hence, the splitting probability is almost unity for all points left of the barrier except for a thin boundary layer of the thickness of the order of $\sqrt{k_BT/(m\omega_b^2)}$ on which it rapidly falls to zero. That simply means that trajectories starting outside the thin layer at the barrier
almost behave as being not influenced by the noise and go to the next local equilibrium point. Only within the thin layer at the barrier the noise is effective and may redirect a trajectory to the “wrong” metastable state.

2.10.4.2 The mean first passage time

In one dimension the equation for the mean first passage time becomes:

\[
K(y) \frac{d}{dy} \left( \frac{d}{dy} \right) \langle t(y) \rangle = -1
\]

\[
\frac{d}{dy} \langle t(x_1) \rangle = 0, \quad \langle t(x_2) \rangle = 0
\]

(2.300)

where we only consider exits through the right boundary point \( y = y_2 \). Once the trajectory reaches the left boundary at \( y = y_1 \) it is reflected. Other situations can similarly be treated but will not be discussed here. The solution of this boundary value readily can be found:

\[
\langle t(y) \rangle = \int_{y_1}^{y_2} \int_{y}^{y_2} \frac{\exp \{ V(u) - V(v) \}}{D(v)} du dv
\]

(2.301)

where \( V(y) \) is defined by eq. (2.295)

For a freely diffusing particle one obtains:

\[
\langle t(y) \rangle_{\text{diff}} = \frac{1}{2D} \left( (y_1 - y)^2 - (y_1 - y_2)^2 \right)
\]

(2.302)

The parabolic profile reflects the typical behavior of diffusion.

For the bistable oscillator we choose the metastable state at the right of the barrier as absorbing state, \( y_2 = x_1 \) and take the limit \( x_1 \to -\infty \) for the left (reflecting) boundary point. The average time it takes to reach the metastable state \( x_1 \) from a starting point on its left hand side then becomes:

\[
\langle t(y) \rangle = \frac{\gamma}{k_B T} \int_{y}^{x_1} \int_{-\infty}^{u} \exp \left\{ \frac{U(u) - U(v)}{k_B T} \right\} dv du
\]

(2.303)

For sufficiently weak noise the double integral is dominated by a single pronounced maximum of the integrand at \( u = x_b = 0 \) and \( v = x_0 \). If the starting point \( y \) lies on the side of the barrier opposite to the final metastable state at \( x_1 \) in a distance of the barrier that is large compared to the thermal length \( l_{th} = \sqrt{k_B T/(m \omega_0^2)} \) the integral becomes independent of \( y \) and takes the value \( T \):

\[
\langle t(y) \rangle = T = \frac{\gamma}{k_B T} \int_{x_0}^{x_1} \exp \left\{ \frac{U(u)}{k_B T} \right\} \int_{-\infty}^{0} \exp \left\{ -\frac{U(v)}{k_B T} \right\} dv du \quad \text{for } y < x_b - l_{th}.
\]

(2.304)
Pontryagin’s equation

For low temperatures both integrals can be evaluated in Gaussian approximation to yield:

\[ T \approx \frac{m\omega_0^2 e^{-\frac{\Delta U}{k_B T}}}{2\pi\gamma} = \frac{1}{k_{\text{Kramers}}^2} \]  

(2.305)

If the starting point lies in the barrier region or on the side of the final metastable state, the two integrals in eq. (2.303) can be disentangled by taking zero as upper limit in the \( v \)-integral. This introduces an exponentially small error of the order \( \exp \left\{ -\frac{\Delta U}{k_B T} \right\} \). The \( u \)-integral can be expressed in terms of the splitting probability. For the mean first passage time this yields the appealing result:

\[ \langle t(y) \rangle = T \pi_0(y) \quad \text{for } x_0 \leq y \leq x_1 . \]  

(2.306)

It means that all the deterministic times it takes the system to move from the initial state to the next metastable state and from the barrier to the final metastable state are much shorter than the exponentially large waiting time in a metastable state. For \( y \)-values smaller than \( x_0 \) the mean first passage time consequently also remains constant provided the potential goes sufficiently fast to infinity for \( x \to \pm \infty \).

The fact that there is a single mean waiting time for the whole domain of attraction apart from a thin layer at the boundary also justifies the coarse grained rate picture discussed above. A further quantitative justification of the rate picture is the agreement of the mean waiting time with the inverse Kramers rate, see eq. (2.305) (Reimann et al, 1999).

References


J. Garcia-Ojalvo, J.M. Sancho, Noise in Spatially Extended Systems(Springer, New
York, Berlin, Heidelberg, 1999)


H.A. Kramers, Brownian Motion in a Field of Force and the Diffusion model of Chemical Reactions, Physica (Utrecht) 7, 284 (1940).

Pontryagin's equation


N. Wax (ed.), *Selected Papers on Noise* (Dover, New York 1954).

