Andreas Schadschneider

Nonequilibrium Physics with Interdisciplinary Applications

Version: October 19, 2013

Sommersemester 2013

Contents

Chapter 1

Introduction

1.1 Nonequilibrium

The concept of thermal equilibrium is the central aspect of the course in Statistical Physics. In principle one can distinguish thermal equilibrium of a system from the equilibrium of two systems in contact.
But what is the precise definition of "equilibrium" ? Usually it is postulated (and supported by experience) that any isolated system[1] tends to reach an "equilibrium state". In this state probability densities or density operators for the microstates become time-independent. This implies that expectation values of physical quantities approach "equilibrium values" which do not depend on time.
Systems in thermodynamic equilibrium can be characterized by a few macroscopic quantities, which are constant in time. Fluctuations around averages are possible (see statistical physics, Gibbs), but certain phenomena like Brownian motion are not correctly described by this.
For a large class of truly non-equilibrium systems no Hamiltonian (and thus no energy function) can be defined. Instead they are defined in terms of transition rates between different configurations. Such a system may attain a non-equilibrium stationary state or exhibit non-stationary behavior for ever. Therefore systems far from equilibrium can be divided further into two subclasses: (a) systems for which Hamiltonian can be defined and the stable thermodynamic equilibrium exists, so that the system evolves spontaneously towards this equilibrium; and (b) systems for which neither the Hamiltonian nor the Gibbsian equilibrium state exist. Systems of type (a) can attain thermodynamic equilibrium state after sufficiently long time. But, a system of type (b) may attain only a non-equilibrium steady-state (see Fig. 1.1.1).
Consider, for example, a window glass which is a supercooled liquid and is in a metastable state. It can, in principle, attain the corresponding crystalline equilibrium structure after sufficiently long time. But, in contrast, a conducting wire carrying an electrical current can, at best, attain a non-equilibrium steady-state after the transient currents die down. Systems in the latter category will be referred to as **driven non-equilibrium system** where, loosely speaking, an external

[1] In fact almost all systems are not perfectly closed. What is meant are systems which are isolated as good as possible from the surrounding world.

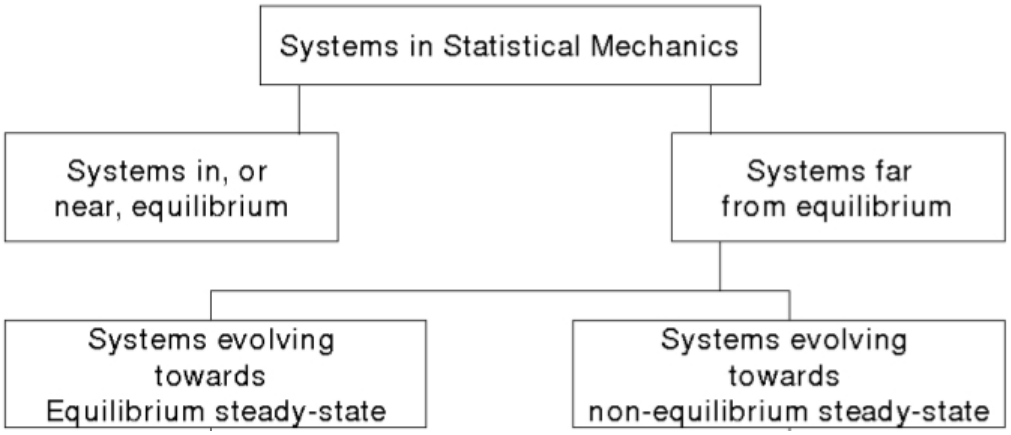

Figure 1.1.1: Equilibrium and non-equilibrium systems in statistical mechanics.

"drive" maintains the corresponding current. A difference of chemical potential between the two ends of a sample can drive a current of material particles. Similarly, an electric potential gradient can drive an electrical current.
Systems which are close to equilibrium can often be treated by **linear response theory**. As an example, consider a time-dependent field $\vec{F}(t)$. The average deviation of an observable $\vec{A}$ from its equilibrium value is then given by

$$\langle \delta \vec{A}(t) \rangle = \int dt' \underline{\underline{K}}(t-t') \vec{F}(t') \tag{1.1.1}$$

where $\underline{\underline{K}}$ is called **response matrix**. It is related to the **dynamical susceptibility**

$$\underline{\underline{\chi}}(\omega) = \int_0^\infty \underline{\underline{K}}(t) e^{i\omega t} dt. \tag{1.1.2}$$

According to the **fluctuation-dissipation theorem** the dynamical susceptibility is related to the correlation function $C_A(t) = \langle A(t)A \rangle$ of equilibrium (!!) fluctuations:

$$C_A(t) = \frac{k_B T}{i\pi} \mathcal{P} \int_{-\infty}^{\infty} d\omega \frac{\underline{\underline{\chi}}(\omega)}{\omega} \cos \omega t \,. \tag{1.1.3}$$

A special case are systems which are evolving very slowly on an experimental timescale, i.e. glasses.

Systems in equilibrium are described by the (canonical) Boltzmann distribution

$$P(C) = \frac{1}{Z} e^{-\beta E(C)} \tag{1.1.4}$$

with $\beta = \frac{1}{k_B T}$ where C specifies a certain configuration of the system. In mechanical systems, C is the set of the positions and momenta of the particles, whereas in the case of Ising spins, $C = \{\sigma_1, \ldots, \sigma_N\}$ with $\sigma_i = \pm 1$. $P(C)$ is the probability to find the system in state C and $E(C)$ is the corresponding energy. The partition function is

$$Z = \sum_C e^{-\beta E(C)} \,. \tag{1.1.5}$$

This is true for any system in thermal equilibrium which can exchange energy with a heat bath of temperature T. So if we know the energy function we can (at least in principle) determine all thermodynamical properties of the system.
An alternative point of view on any system would be not to specify its energies, but to look at the transitions $C \to C'$ between different configurations. This can be done using the **transition rates**[2] $W(C \to C')$. In equilibrium the so-called **detailed balance condition**

$$\boxed{W(C' \to C) e^{-\beta E(C')} = W(C \to C') e^{-\beta E(C)}} \tag{1.1.6}$$

must hold which we will derive later. It implies that the number of transitions, and thus the flux or probability current, from C to C' and $C' \to C$ are equal. This relation is also of some practical importance in computer simulations of equilibrium systems (Monte-Carlo algorithms).
Now we can give a first *Pragmatic definition*:

> **Nonequilibrium systems** are defined by their dynamics, not an energy function (Hamiltonian). The dynamics does not satisfy the detailed balance condition.

In this course we mainly consider systems which are far from equilibrium, especially systems which maintain a current (typically a particle current). Such sytems are generically driven by their environment instead of being in equilibrium with it.

1.2 Description of nonequilibrium systems

One can distinguish three classes of descriptions:

1. stochastic processes
2. kinetic theories
3. response theory

Approaches of type 1. assume that the process is stochastic. Either the process itself is stochastic or stochastic terms (*noise*) are added to the (deterministic) fundamental equations. Such stochastic theories are often used when we do not know the exact interactions. Therefore they appear in many interdisciplinary applications, especially when human behavior is involved.

[2] i.e. a transition probability per time

Kinetic theories try to derive nonequilibrium distribution functions (or density operators) from first principles. They satisfy the so-called kinetic equations.
Response theories describe nonequilibrium systems in terms of equilibrium quantities, e.g. the linear response theory described above. It usually starts from first principles, e.g. the Liouville equation.

Chapter 2

Diffusion

First we will look at **diffusion** which will become a kind of reference systems for us in the following. We try to describe the approach to equilibrium.

2.1 The diffusion equation

The system we are considering is assumed to be (chemically) homogeneous, i.e. it consists of a single type of particles which have a mass m. The total mass is then $M = Nm$ where N is the number of particles in the system. It is conserved and therefore

$$\frac{dN}{dt} = 0\,. \tag{2.1.1}$$

We introduce a particle density $n(\vec{r})$ with

$$N = \int n(\vec{r}) d^3r \tag{2.1.2}$$

and the current density $\vec{j}(\vec{r})$

$$\vec{J} = \int \vec{j}(\vec{r}) d^3r\,. \tag{2.1.3}$$

In the same way as in electrodynamics[1] we can then derive the **continuity equation** as a local expression for the conservation of particles:

$$\frac{\partial n}{\partial t} + \nabla \cdot \vec{j} = 0\,. \tag{2.1.4}$$

If we now start at time $t = 0$ with an inhomogeneous density distribution of the particles, experience tells us that it will become homogeneous through diffusion after a long time (see Fig. 2.1.1).

[1] By considering a finite volume V with surface F: the number of particles in V can only change if there is a current through F.

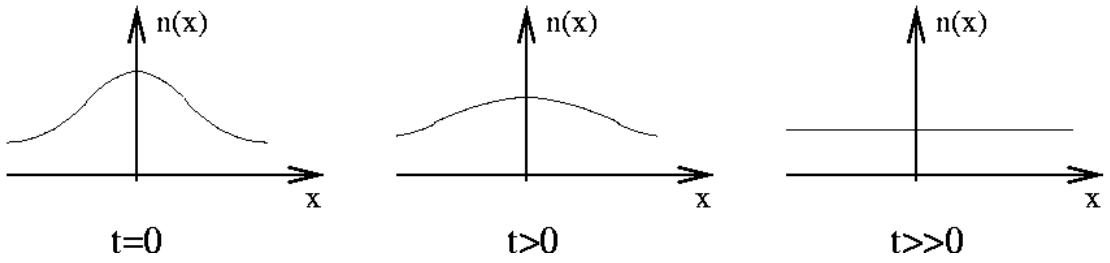

Figure 2.1.1: An inhomogeneous density distribution (left) will become homogeneous after a long time (right) by diffusion.

From the observation of colloidal particles in a suspension, A.E. Fick concluded in 1855 that origin of the current is the gradient of density distribution. This is summarized in **Fick's law**:

$$\vec{j}(\vec{r}) = -D\nabla n(\vec{r})\,. \tag{2.1.5}$$

The constant of proportionality D is called **diffusion coefficient**. It has the dimension $[D] = (\text{length})^2/\text{time}$.
Inserting Fick's law into the continuity equation yields the **diffusion equation** (or **heat equation**):

$$\boxed{\frac{\partial n(\vec{r},t)}{\partial t} = D\Delta n(\vec{r},t)\,,} \tag{2.1.6}$$

which relates the temporal change of the density distribution with its second spatial derivative.

2.2 Solution in one dimension

To obtain a better understanding of the dynamics described by the diffusion equation we solve it in the case of one spatial dimension. The initial density distribution is given by the function

$$n(x) := n(x, t = 0)\,. \tag{2.2.1}$$

For the solution of the 1d diffusion equation

$$\frac{\partial n}{\partial t} = D\frac{\partial^2 n}{\partial x^2} \tag{2.2.2}$$

we make the factorizing Ansatz

$$n(x,t) = f(x)g(t)\,. \tag{2.2.3}$$

After inserting into (2.2.2) we obtain

$$\frac{1}{g}\frac{\partial g}{\partial t} = \frac{D}{f}\frac{\partial^2 f}{\partial x^2} \tag{2.2.4}$$

where the left side is a function of t only and the right-hand side of x only. Therefore both sides have to constant in order to be true for all values of x and t. We denote this constant for convenience by $-k^2D$. Then (2.2.4) is reduced to two independent equations

$$\frac{dg}{dt} = -k^2Dg\,, \tag{2.2.5}$$

$$\frac{d^2f}{dx^2} = -k^2f\,. \tag{2.2.6}$$

Both equations are easily solved:

$$g(t) = Ae^{-k^2Dt}\,, \tag{2.2.7}$$

$$f(x) = Be^{ikx}\,. \tag{2.2.8}$$

We just remark here that we do not need to consider the full solution for $f(x)$ here which would also include the term e^{-ikx} (see below).
The most general solution is then obtained by superposition of the solutions for all k:

$$n(x,t) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} a(k) e^{ikx-k^2Dt} \tag{2.2.9}$$

where, for convenience, we have combined the product of the integration constants into a single function $a(k)/2\pi$. Here we see why the term e^{-ikx} has not been included in the solution of $f(x)$: since we integrate over all k-values, including $k < 0$. Including the e^{-ikx}-term would only lead to a redefinition of $a(k)$.
The solution has the form of a Fourier transformation. The Fourier transform of the initial condition $n(x)$ is denoted by $n(k)$, i.e.

$$n(x) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} n(k) e^{ikx}\,, \tag{2.2.10}$$

so that we find (by putting $t = 0$ in (2.2.9))

$$a(k) = \int_{-\infty}^{\infty} n(x) e^{-ikx} dx\,. \tag{2.2.11}$$

Inserting this into (2.2.9) we obtain

$$n(x,t) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \int_{-\infty}^{\infty} dx' n(x') e^{ik(x-x')-k^2Dt}\,. \tag{2.2.12}$$

This can be simplified by completing the square of the exponential function and using the Gauss integral

$$\int_{-\infty}^{\infty} e^{a^2x^2} dx = \frac{\sqrt{\pi}}{a}\,. \tag{2.2.13}$$

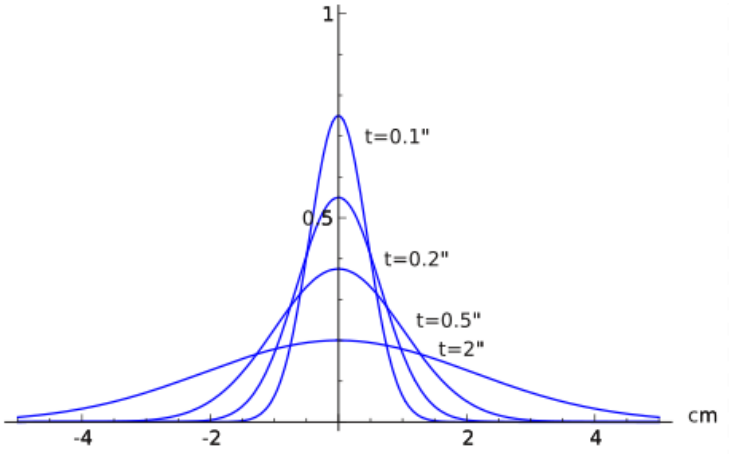

Figure 2.2.1: Time evolution of the density profile $n(x,t)$ due to diffusion for a delta-function initial condition [1].

Finally one obtains

$$n(x,t) = \int_{-\infty}^{\infty} dx' n(x') \frac{e^{-\frac{(x-x')^2}{4Dt}}}{\sqrt{4\pi Dt}} \,. \tag{2.2.14}$$

One can convince oneself that indeed the particle number is conserved, i.e.

$$\int_{-\infty}^{\infty} n(x,t)dx = \int_{-\infty}^{\infty} n(x)dx = N \,. \tag{2.2.15}$$

For illustration we consider the case of a delta function initial condition

$$n(x) = N\delta(x) \,, \tag{2.2.16}$$

i.e. at time $t = 0$ all particles are located at $x = 0$. Then the solution simplifies to

$$n(x,t) = N \frac{e^{-\frac{x^2}{4Dt}}}{\sqrt{4\pi Dt}} \tag{2.2.17}$$

which is also called the *standard solution*. It is a Gauss function of width $\sim \sqrt{Dt}$ which increases with $t^{1/2}$ (cf. Fig. 2.2.1).

2.3 Some properties of the solution

Finally we add some remarks:

1. Diffusion is infinitely fast, e.g. $n(x, t = \epsilon) \neq n(x, t = 0)$ for all x and $\epsilon > 0$.

2. The quantity
$$p(x,t) := \frac{n(x,t)}{N} \tag{2.3.1}$$
may be interpreted as the probability to find a particle in the interval $[x, x+dx]$. Obviously (compare with (2.2.15) it satisfies the necessary normalization condition
$$\int_{-\infty}^{\infty} p(x,t)dx = 1\,. \tag{2.3.2}$$

3. One can introduce a transition rate by
$$n(x,t) = \int dx_0\, W\big((x_0,t_0) \to (x,t)\big) n(x_0,t_0)\,, \tag{2.3.3}$$
where W also satisfies a diffusion equation and the definition of a Markov process (see later). This will be explored further in Exercise 1.

We can also calculate averages using the probability density (2.3.1). After a short calculation one finds
$$\langle x^2 \rangle = \int_{-\infty}^{\infty} x^2 p(x,t)dx = 2Dt\,, \tag{2.3.4}$$
where in the evaluation of the integral we have used
$$\int_{-\infty}^{\infty} x^2 e^{a^2x^2} dx = \frac{1}{2a}\frac{\sqrt{\pi}}{a}\,. \tag{2.3.5}$$

2.4 Diffusion under external forces

Next we consider diffusion under the influence of an external force $\vec{F}$. To study the consequences we briefly look at the system in a microscopic way and consider the dynamics of individual particles. Each particle obeys the equation of motion
$$m\dot{\vec{v}} = \vec{F} - \alpha\vec{v} \tag{2.4.1}$$
where $-\alpha\vec{v}$ is a force due to friction effects. The stationary solution of this equation of motion is obtain by setting $\dot{\vec{v}} = 0$:
$$\vec{v}_B = \frac{\vec{F}}{\alpha} =: B\vec{F} \tag{2.4.2}$$
where we have introduced the **mobility** $B = \frac{1}{\alpha}$.
For simplicity we again consider the one-dimensional case. The current has two contributions, one coming from diffusion and the other from the mobility:
$$j = j_D + j_B \tag{2.4.3}$$

where

$$j_D = -D\frac{\partial n}{\partial x}\,, \qquad j_B = nv_B = nBF\,. \tag{2.4.4}$$

The continuity equation reads

$$\frac{\partial n}{\partial t} = -\frac{\partial j}{\partial x} = -v_B\frac{\partial n}{\partial x} + D\frac{\partial^2 n}{\partial x^2}\,. \tag{2.4.5}$$

Its stationary solution (for $\frac{\partial n}{\partial t} = 0$) is given by

$$n(x) = n(x_0)e^{v_B(x-x_0)/D}\,, \tag{2.4.6}$$

which implies a vanishing current

$$j = \left(v_B - D\frac{v_B}{D}\right)n(x) = 0\,, \tag{2.4.7}$$

since $j_D = -v_B n$ and $j_B = v_B n$.
Now we assume that the stationary state is in thermal equilibrium. For constant F the density distribution becomes *barometric*, i.e. a Boltzmann distribution:

$$n(x) = n(x_0)e^{F(x-x_0)/k_BT}\,. \tag{2.4.8}$$

By comparison with (2.4.6) we see that $\frac{F}{k_BT} = \frac{v_B}{D}$. In combination with the definition (2.4.2) of the mobility this leads to the famous **Einstein relation** (1905)

$$D = k_BTB \tag{2.4.9}$$

which relates the diffusion coefficient to temperature. The temperature dependence indicates that diffusion is related to thermal collisions.

Chapter 3

Langevin equation

3.1 Derivation of the Langevin equation

Up to now we have described diffusion using a continuum description which does not refer to the motion of individual atoms or molecules. The only exception was the derivation of the mobility when external forces are included. Next we extend this line of reasoning to derive a phenomenological atomistic description of diffusion.

A good starting point, also historically, is the discovery of **Brownian motion**. In 1827, the botanist Robert Brown studied pollen grains in water under a microscope. Surprisingly he found that they moved and – even more surprisingly – the motion appeared to be rather irregular. An explanation was given in 1905 by Albert Einstein who argued that this irregular motion is caused by collisions of the pollen grain with the water molecules (which, of course, are invisible even under the microscope). These ideas were verified experimentally in 1908 by Jean Perrin. He received the Nobel price in 1926 because the experiments were an important step towards the establishment of atomic theory which – before that – was only hypothetical. Another conceptual progress was that Einstein's theory demonstrated how usefull a stochastic description of physical processes can be.

In the following we try to show that Brownian motion allows deep insights into the mechanism responsible for fluctuations and dissipation of energy and how background noise limits the accuracy.

Due to the random collisions with other particles the equation of motion of a specific particle can be written in the form

$$m\dot{v} = F_{\text{ext}}(t) + F(t)\,. \tag{3.1.1}$$

F_{ext} is an external force and $F(t)$ the force exerted by the other particles. The latter varies extremely fast on a typical timescale τ^* (depending on the changing positions of all particles!). τ^* can be estimated through the time between consecutive collisions:

$$\tau^* = \frac{\ell}{v_{\text{mol}}} \approx 10^{-13}\ s \tag{3.1.2}$$

where ℓ is the typical distance between particles (given by the density) and $v_{\text{mol}} \approx 1000$ m/s the typical molecular velocity.

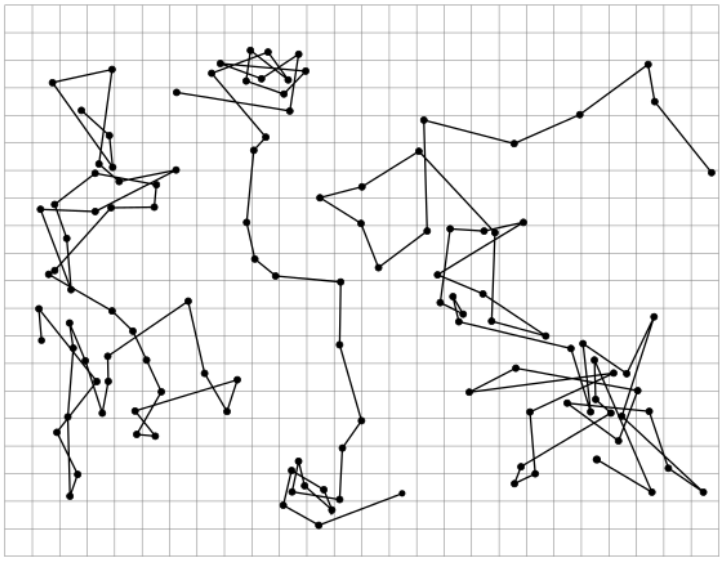

Figure 3.1.1: Brownian motion: Erratic trajectory of a pollen grain in water [1].

Since we are looking for a statistical description we consider an ensemble of identical systems[1]. We split the velocity into two contributions,

$$v = \langle v \rangle + v' \tag{3.1.3}$$

where $\langle v \rangle$ is the ensemble average of the velocity and v' a fast changing contribution. Similarly the force can be divided:

$$F = \langle F \rangle + F' \,. \tag{3.1.4}$$

Since $\langle F \rangle (v = 0) = 0$ the expansion of $\langle F \rangle$ with respect to v yields to lowest order

$$\langle F \rangle = -\alpha \langle v \rangle \,, \tag{3.1.5}$$

where α can be interpreted as friction constant. The minus sign appears because $\langle F \rangle$ has the tendency to reduce $\langle v \rangle$ (we must have $\lim_{t\to\infty} \langle v \rangle(t) = 0$).
Now we can give the equation of motion for the slowly varying part:

$$m\langle \dot{v} \rangle = F_{\text{ext}} + \langle F \rangle = F_{\text{ext}} - \alpha \langle v \rangle \,. \tag{3.1.6}$$

Note that it is not simply given by $m\langle \dot{v} \rangle = F_{\text{ext}}$ because then the system would not reach the equilibrium value $\langle v \rangle = 0$ for $F_{\text{ext}} = 0$.
If we include the fast-varying part we obtain[2] the **Langevin equation** (1908)

$$\boxed{m\dot{v} = F_{\text{ext}} - \alpha v + F'(t) \,,} \tag{3.1.7}$$

[1] Which differ in the realisation of the random force.
[2] using the approximation $\langle v \rangle \approx v$

which, for $F_{\text{ext}} = 0$ can be written in the form

$$\boxed{\dot{v} = -\gamma v + \xi(t)\,,} \tag{3.1.8}$$

The Langevin equation describes the force on a single molecule by splitting it into a slowly varying part $-\gamma v$ and a fast fluctuating random part $\xi(t)$ which has a vanishing average:

$$\langle \xi(t) \rangle = 0\,. \tag{3.1.9}$$

3.2 Averages

In order to understand the Langevin equation in more detail we take averages $\langle ... \rangle$ over many particles. Position x and the random force ξ are statistically independent so that we have

$$\langle x(t)\xi(t) \rangle = \langle x(t) \rangle \langle \xi(t) \rangle = 0\,, \tag{3.2.1}$$

due to (3.1.9).
If we assume that the particles are in a thermal equilibrium state we can apply the equipartition theorem

$$\frac{m}{2}\langle v^2 \rangle = \frac{1}{2}kT \tag{3.2.2}$$

where we have assumed for simplicity that we are dealing with a one-dimensional problem. In Excercise 2 we will show that then

$$\langle xv \rangle = \int_0^t e^{-\gamma(t-t')} kT\, dt' = \frac{kT}{\gamma}\left[1 - e^{-\gamma t}\right]\,. \tag{3.2.3}$$

and (by integration of this result)

$$\langle x^2 \rangle = 2\frac{kT}{\gamma}\left[t - \frac{1}{\gamma}\left(1 - e^{-\gamma t}\right)\right]\,. \tag{3.2.4}$$

Taking the long-time limit $t \to \infty$ the mean-square displacements simplifies to

$$\langle x^2 \rangle \approx \frac{2kT}{\gamma} t \tag{3.2.5}$$

so that $\langle x^2 \rangle \propto t$ as previously derived for diffusion. In the opposite (short-time limit) $t \ll 1$ the mean-square displacement is $\langle x^2 \rangle \propto t^2$ as for free thermal particles.

In general the fluctuating force is characterized by its vanishing average (3.1.9) and its temporal correlations

$$\langle \xi(t)\xi(t') \rangle = \phi(t - t')\,. \tag{3.2.6}$$

The function ϕ is symmetric, i.e. $\phi(t) = \phi(-t)$.
In a general **Ornstein-Uhlenbeck process** these correlations are only short-lived, i.e.

$$\phi(t) \approx 0 \qquad \text{for } t > t_c\,. \tag{3.2.7}$$

In this context t_c is called **correlation time**.

Chapter 4

Fokker-Planck equation

We now change slightly our point of view. Instead of considering a description of a single diffusing particle in terms of classical dynamics (plus a stochastic force) we look at an ensemble of many particles ($N \gg 1$). Based on the experience with many particle system a description in terms of distribution functions seems to be promising.

4.1 From Langevin to Fokker-Planck

The dynamics of each particle is assumed to be determined by a Langevin equation

$$\dot{v}_i + \gamma v_i = \xi(t)\,. \tag{4.1.1}$$

We now try to determine the velocity distribution function $P(v,t)$ which gives the average number of particles which have velocity v at time t. A formal definition of this function is

$$P(v,t) = \lim_{N\to\infty} \sum_{i=1}^{N} \delta(v - v_i(t)) =: \langle \delta(v - v(t)) \rangle\,. \tag{4.1.2}$$

Since we are interested in the time evolution of this quantity we start with a formal expansion in the time increment Δt:

$$\begin{aligned} P(v,t+\Delta t) &= \langle \delta(v - v(t) - \Delta v) \rangle \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial v^n} \langle (\Delta v)^n \delta(v - v(t)) \rangle\,. \end{aligned} \tag{4.1.3}$$

The velocity increment Δv can be related to Δt by integrating the Langevin equation:

$$\Delta v = \int_t^{t+\Delta t} \dot{v}(t')dt' = -\gamma v \Delta t + \int_t^{t+\Delta t} \xi(t')dt'\,, \tag{4.1.4}$$

where we have used that $-\gamma v$ is a slowly varying variable, in contrast to $\xi(t)$.
In order to derive a dynamical equation for $P(v,t)$ we will now consider the averages $A_n := \langle (\Delta v)^n \delta(v - v(t)) \rangle$ in (4.1.3) for all values of n.

- For $n = 0$ we have
$$A_0 = \langle \delta(v - v(t)) \rangle = P(v,t) \tag{4.1.5}$$
according to the definition (4.1.2).

- For $n = 1$ we use (4.1.4) and obtain
$$\begin{aligned} A_1 &= \langle \Delta v \delta(v - v(t)) \rangle \\ &= -\gamma \langle v(t) \delta(v - v(t)) \rangle + \int_t^{t+\Delta t} \langle \xi(t') \delta(v - v(t')) \rangle dt' \,. \end{aligned} \tag{4.1.6}$$
The first term can be simplified using the properties of δ-functions:
$$\langle v(t) \delta(v - v(t)) \rangle = \langle v \delta(v - v(t)) \rangle = v \langle \delta(v - v(t)) \rangle = v P(v,t) \tag{4.1.7}$$
where in the second step we have used that v is a (constant) parameter.
To treat the integral in the second term we assume that the fluctuating force $\xi(t)$ and the velocity $v(t)$ are not correlated:
$$\int_t^{t+\Delta t} \langle \xi(t') \delta(v - v(t')) \rangle dt' \int_t^{t+\Delta t} \langle \xi(t') \rangle \langle \delta(v - v(t')) \rangle dt' = 0 \,, \tag{4.1.8}$$
since $\langle \xi(t) \rangle = 0$.
Combining these results we find
$$A_1 = \langle \Delta v \delta(v - v(t)) \rangle = -\gamma v \Delta t P(v,t) \,. \tag{4.1.9}$$

- In the case $n = 2$ we use that the stochastic force is δ-correlated, i.e.
$$\langle \xi(t_1) \xi(t_2) \rangle = 2\gamma^2 D \delta(t_1 - t_2) \tag{4.1.10}$$
where the form of the coefficient[1] has been chosen for later convenience. Using a similar argument (and e.g. (4.1.2)) as in the case $n = 1$ we find
$$\begin{aligned} A_2 &= \int_t^{t+\Delta t} dt_1 \int_t^{t+\Delta t} dt_2 \langle \xi(t_1) \xi(t_2) \rangle P(v,t) - 2\gamma \Delta t \langle v(t) \delta(v - v(t)) \rangle \int_t^{t+\Delta t} \langle \xi(t') \rangle dt' \\ &+ \mathcal{O}\left((\Delta t)^2 \right) \end{aligned} \tag{4.1.11}$$
where the terms proportional to $(\Delta t)^2$, which are generated by calculating $(\Delta v)^2$, are neglected. Since the average of the fluctuating force vanishes we finally obtain
$$A_2 = 2\gamma^2 D \Delta t P(v,t) + \mathcal{O}\left((\Delta t)^2 \right) \,. \tag{4.1.12}$$

- The terms generated by $n \geq 3$ do not generate further contributions. Assuming that the stochastic process is Gaussian one can show that they of the order $\mathcal{O}\left((\Delta t)^n \right)$.

[1] corresponding at this point to the definition of D

Now we can collect the results for all the terms and determine the expansion (4.1.3). Since

$$\lim_{\Delta t\to 0}\frac{P(v,t+\Delta t)-P(v,t)}{\Delta t}=\frac{\partial P}{\partial t} \tag{4.1.13}$$

we then obtain the **Fokker-Planck equation**

$$\boxed{\frac{\partial P}{\partial t}=\gamma\frac{\partial}{\partial v}\left(v+\gamma D\frac{\partial}{\partial v}\right)P(v,t)} \tag{4.1.14}$$

derived by Fokker in 1913 and Planck in 1927.

4.1.1 Stationary solution

The stationary solution of the Fokker-Planck equation is easy to determine. Since $\frac{\partial P}{\partial t}=0$ in the stationary state, $vP(v,t)+\gamma D\frac{\partial}{\partial v}P(v,t)$ has to be constant. This can be integrated and finally yields the stationary solution (using the Einstein relation $\gamma D=kT/m$ see (2.4.9))

$$P(v)=Ce^{-\frac{1}{2}mv^2/kT} \tag{4.1.15}$$

which is the well-known **Maxwell distribution**.

4.2 Reduction to diffusion equation

In the following we will relate the Fokker-Planck equation to a diffusion problem. This is achieved by first making a transformation of variables by introducing

$$u:=ve^{\gamma t}\,. \tag{4.2.1}$$

Then we have

$$P(v,t)=P\left(ue^{-\gamma t},t\right)=:\tilde{P}(,u,t) \tag{4.2.2}$$

so that we can consider the probability distribution to be a function of t and the rescaled variable u (instead of v). The time evolution of this probability distribution is given by

$$\begin{aligned}\frac{\partial\tilde{P}}{\partial t} &= \frac{\partial P}{\partial v}\frac{\partial v}{\partial t}+\frac{\partial P}{\partial t}=-\gamma v\frac{\partial P}{\partial v}+\left(\gamma P+\gamma v\frac{\partial P}{\partial v}+\gamma^2 D\frac{\partial^2 P}{\partial v^2}\right)\\ &= \gamma\tilde{P}\gamma^2 D\frac{\partial^2\tilde{P}}{\partial u^2}\,.\end{aligned} \tag{4.2.3}$$

where we have used $\frac{\partial v}{\partial t}=-\gamma v$ and the Fokker-Planck equation for P.
In the next step we introduce a new function a related to the rescaled probability distribution and a rescaled time s by

$$\tilde{P}=ae^{\gamma t}\,,\qquad\text{and}\quad s=\frac{1}{2\gamma}\left(e^{2\gamma t}-1\right)\,. \tag{4.2.4}$$

Since $ds = e^{2\gamma t}dt$ after inserting into (4.2.3) we obtain

$$\frac{\partial a(u,s)}{\partial s} = \epsilon \frac{\partial^2 a}{\partial u^2} \qquad \text{with } \epsilon = \gamma^2 D\,. \tag{4.2.5}$$

Thus the function a satisfies a diffusion equation! Using the standard solution

$$a(u,s) = \frac{1}{\sqrt{4\pi\epsilon s}} e^{-\frac{(u-u_0)^2}{4\epsilon s}} \tag{4.2.6}$$

we finally obtain the probability distribution

$$P(v,t) = \sqrt{\frac{m}{2\pi kT\left(1-e^{-2\gamma t}\right)}} \exp\left(-\frac{m\left(v-v_0 e^{-\gamma t}\right)}{2kT\left(1-e^{-2\gamma t}\right)}\right) \tag{4.2.7}$$

for $t > 0$. This function describes the monotonous approach to the equilibrium distribution. It is a Gaussian distribution with average

$$\langle v(t)\rangle = v_0 e^{-\gamma t} \tag{4.2.8}$$

and variance

$$\sigma^2(t) = \gamma D\left[1 - e^{-2\gamma t}\right]\,. \tag{4.2.9}$$

4.3 Fokker-Planck equation with external field

The Fokker-Planck equation can be generalized to include an external field

$$F(x) = -\frac{dV}{dx}\,. \tag{4.3.1}$$

The probability function P then becomes a function not only of the velocity v and time t, but also the position x, i.e. $P = P(x,v,t)$. Starting from the Langevin equation

$$\dot{v} + \gamma v \frac{F(x)}{m} + f\,, \tag{4.3.2}$$

we can derive the corresponding Fokker-Planck equation along the lines of Sec. 4.1. It takes the form

$$\boxed{\frac{\partial P(x,v,t)}{\partial t} = \gamma \frac{\partial}{\partial v}\left[v + \gamma D \frac{\partial}{\partial v}\right] P(x,v,t) - v\frac{\partial P}{\partial x} - \frac{F(x)}{m}\frac{\partial P}{\partial v}\,.} \tag{4.3.3}$$

The corresponding stationary solution is

$$P(x,v) = C \exp\left(-\frac{\frac{1}{2}mv^2 + V(x)}{kT}\right) \tag{4.3.4}$$

which is the **Maxwell-Boltzmann distribution**.

Chapter 5

Random Walk

So far we have introduced a microscopic description of diffusion based on ideas from classical mechanics (forces!) with additional random fluctuations (due to random collisions). Next we look at the problem from a slightly different perspective without refering to forces. Instead the new description does not explicitly refer to the origin of the observed motion.

5.1 Symmetric random walk

We consider a single particle moving on a discrete lattice with lattice sites n. In each timestep it can move either to the left neighbour site $n-1$ with probability 1/2 or to the right neigbhour $n+1$ also with probability 1/2 (Fig. 5.1.1. This is the so-called **random walk** which can be considered as a simple model for a diffusing particle.

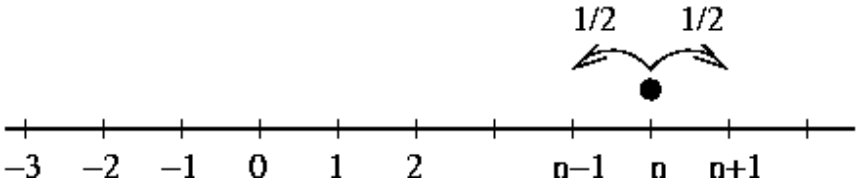

Figure 5.1.1: The symmetric random walk: in each step the particle moves with probability 1/2 either to the left or the right neighbour site.

Next we ask the following question: if the particle starts at the origin $n = 0$, what is the probability $P(m, N)$ that it ends at site m after performing N steps?
First, one easily convinces oneself that in order to end on site m in N steps the particle has to perform

$$N_+ = \frac{1}{2}(N + m) \tag{5.1.1}$$

steps to the right and

$$N_- = N - N_+ = \frac{1}{2}(N - m) \tag{5.1.2}$$

steps to the left. The order in which these steps are performed is arbitrary! Overall there are $\binom{N}{N_+}$ different ways of reaching m. Thus the probability to reach m in N steps is given by

$$P(m,N) = \binom{N}{N_+}\left(\frac{1}{2}\right)^{N_+}\left(\frac{1}{2}\right)^{N-N_+} = \frac{N!}{\left(\frac{1}{2}(N+m)\right)!\left(\frac{1}{2}(N-m)\right)!}\left(\frac{1}{2}\right)^{N} \tag{5.1.3}$$

which is a binomial distribution.
This exact expression can be simplified for large N (and finite m) using **Stirling's formula**

$$n! \simeq \sqrt{2\pi n}\left(\frac{n}{e}\right)^n , \tag{5.1.4}$$

or, in logarithmic form,

$$\ln n! \simeq \frac{1}{2}\ln 2\pi + \frac{1}{2}\ln n + n\ln n - n\,. \tag{5.1.5}$$

This gives

$$\begin{aligned}\ln P(m,N) &\simeq \left(N+\frac{1}{2}\right)\ln N - \frac{1}{2}\ln 2\pi - N\ln 2 - \frac{1}{2}(N+m+1)\ln\left(\frac{N}{2}(1+\frac{m}{N})\right)\\ &- \frac{1}{2}(N-m+1)\ln\left(\frac{N}{2}(1-\frac{m}{N})\right)\\ &\simeq -\frac{1}{2}\ln N + \ln 2 - \frac{1}{2}\ln 2\pi - \frac{m^2}{2N}\end{aligned} \tag{5.1.6}$$

where we have used $\ln(1\pm\frac{m}{N}) \approx \pm\frac{m}{N} - \frac{m^2}{2N^2}$ in the last step. Finally we obtain

$$P(m,N) = \sqrt{\frac{2}{N\pi}}e^{-\frac{m^2}{2N}}\,, \tag{5.1.7}$$

i.e. a Gaussian distribution as expected for a diffusion problem.

Next we perform the continuum limit by introducing a lattice constant a which is the distance between neighbouring lattice points. A lattice point m (located at ma) then represents a continuous space interval $[x, x+dx]$ with the continuous coordinate $x = ma$. The corresponding probability distribution is now[1]

$$P(x,N)dx = P(m,N)\frac{dx}{a} \tag{5.1.8}$$

and therefore

$$P(x,N) = \frac{1}{\sqrt{\pi N a^2}}e^{-\frac{x^2}{2Na}}\,. \tag{5.1.9}$$

We can introduce a time variable t by assuming that the particle performs N steps in time t, i.e. $\nu = \frac{N}{t}$ steps per unit time (sec). Then

$$P(x,N) = \frac{1}{\sqrt{4\pi D t}}e^{-\frac{x^2}{4Dt}} \qquad \text{where } D := \frac{\nu a^2}{2} = \frac{Na^2}{2t}\,. \tag{5.1.10}$$

[1] according to the rules of transformation under a change of random variables

Thus indeed we recover the diffusive behaviour in form of the standard solution (2.2.17). Note the typical scaling behaviour (similarity law) of this solution. It is invariant under the transformation

$$t \to \alpha t \qquad \text{and } x \to \sqrt{\alpha} x \tag{5.1.11}$$

for arbitrary α. This implies that for diffusion the relevant quantity is x^2/t, not x/t.

5.2 Asymmetric random walk

Next we generalize the random walk to asymmetric hopping, i.e. the particle moves to the right with rate p and to the left with rate q. This corresponds to a situation where an external field is applied. We are now considering the random walk in continuous time t instead of looking at the discrete steps. A **rate** is a probability per time. Thus "rate p" means that in a time interval Δt the hopping occurs with probability $p\Delta t$. Finally we introduce periodic boundary conditions (Fig. 5.2.1) where sites j and $L + j$ are equivalent. This has the big adavantage of making the system translational invariant so that we do not have to consider effects due its boundaries. In this way the theoretical analysis is considerably simplified.

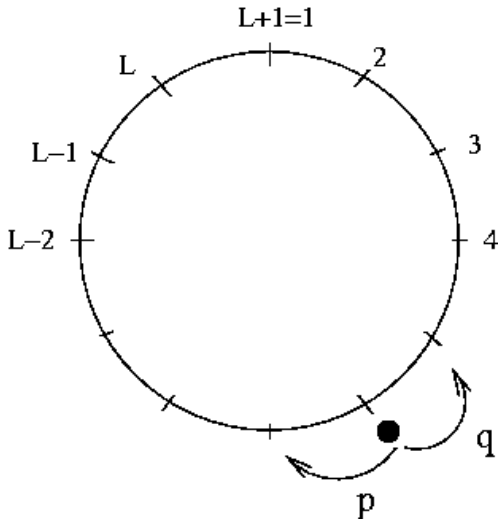

Figure 5.2.1: Asymmetric random walk with periodic boundary conditions. Here site $L + 1$ is identified with site 1.

Again we are interested in the probability $P(n, t)$ to find the particle at site n at time t. This time we will choose a slightly different approach to determine this quantity which will become our standard method in the following. This method starts from a dynamical equation for the probability $P(n, t)$ which has the form

$$\boxed{\frac{\partial P(n,t)}{\partial t} = pP(n-1,t) + qP(n+1,t) - (p+q)P(n,t)}\,. \tag{5.2.1}$$

This is called **master equation** and can be considered as a kind of continuity equation for the probabilites. Its origin is easy to understand. The left-hand-side determines the change of $P(n,t)$ in an infinitesimal time interval dt. Such a change can occur for three different reasons:

1. A particle moves from site $n-1$ to site n. This occurs with rate p and leads to an increase of $P(n,t)$.

2. A particle moves from site $n+1$ to site n. This again results in an increase of $P(n,t)$ and occurs with rate q.

3. A particle moves from site n to site $n-1$ or $n+1$. This occurs with rate $p+q$ and decreases $P(n,t)$.

So the right-hand-side of (5.2.1) consists of gain terms (first two terms) which increase $P(n,t)$ and loss terms (last term) which leads to a decrease. This is the general basis for any master equation and will be discussed later in more detail.
The master equation (5.2.1) for the asymmetric random walk with periodic boundary conditions can be solved using the **generating function technique**. In the present case the generating function takes the form of a Fourier series reflecting the spatial periodicity of the system. Details of the solution will be discussed in *Excercise 5*. It turns out that $P(n,t)$ is given by

$$P(n,t) = \sum_{k=1}^{L} e^{\lambda_k t} z_k^{-n} \tilde{P}(k,0) \tag{5.2.2}$$

with

$$z_k = e^{2\pi i k/L} \qquad \text{and} \quad \lambda_k = p(z_k - 1) + q\left(\frac{1}{z_k} - 1\right) \qquad (k = 1, \ldots, L)\,. \tag{5.2.3}$$

$\tilde{P}(k,0)$ is related to the initial condition $P(n, t=0)$ by

$$\tilde{P}(k,0) = \frac{1}{L}\sum_{n=1}^{L} P(n,0) z_k^n\,. \tag{5.2.4}$$

The stationary state is reached for $t \to \infty$. In this case only the term with $k = L$ in (5.2.2) survives. The mode $k = L$ corresponds to $z_L = 1$ and $\lambda_L = 0$ according to (5.2.3) and for $t \to \infty$ we have

$$P(n,t) \to P(n) = \tilde{P}(L,0) = \frac{1}{L}\,. \tag{5.2.5}$$

The probabilities in the stationary state become constant, i.e. independent of the position n. This comes not unexpected because it reflects the translational invariance of the problem.
The other eigenvalues λ_k determine the decay or **relaxation times** of mode k,

$$\tau_k := \frac{1}{|\mathrm{Re}(\lambda_k)|}\,, \tag{5.2.6}$$

i.e. how quick the stationary state is reached. For finite k and $L \gg 1$ one finds

$$\lambda_k \simeq (p-1)\frac{2\pi i k}{L} - (p+q)\frac{4\pi^2 k^2}{L^2}\,. \tag{5.2.7}$$

The **dynamical critical exponent** z describes how the relaxation time increases with the system size L. It is defined by

$$\tau_k \sim L^z\,. \tag{5.2.8}$$

This describes e.g. the phenomenon of **critical slowing down** in Monte Carlo simulations of systems near a critical point (which you might know from the course on Computional Physics). Near criticality the system needs a long time to reach the equilibrium state, especially for large system sizes.
In our case we find that the dynamical critical exponent is given by

$$\boxed{z = z_{\text{diffusion}} = 2} \tag{5.2.9}$$

which is typical for diffusion.

5.3 Random walk on infinite lattice

As final example we consider an asymmetric random walk on an infinite lattice as Fig. 5.1.1. The corresponding master equation has the same form as (5.2.1) but without the identification of sites k and $L + k$.
Introducing the generating function

$$F(z,t) := \sum_{n=-\infty}^{\infty} P(n,t)\left(\frac{q}{p}\right)^{n/2} z^n\,, \tag{5.3.1}$$

where the factor $\left(\frac{q}{p}\right)^{n/2}$ has been introduced for later convenience. By taking the derivative and using the master equation (5.2.1) one then finds

$$\frac{\partial F(z,t)}{\partial t} = \left[-(p+q) + \sqrt{pq}\left(z + \frac{1}{z}\right)\right] F(z,t) \tag{5.3.2}$$

and thus

$$F(z,t) = F(z,0)\exp\left(-(p+q)t\right)\exp\left(\sqrt{pq}\left(z+\frac{1}{z}\right)t\right)\,. \tag{5.3.3}$$

The explicit form of the probabilities $P(n,t)$ can be expressed by modified Bessel functions of the first kind $I_n(x)$ which are defined by their generating function

$$\exp\left(\frac{1}{2}\left(z+\frac{1}{z}\right)u\right) = \sum_{n=-\infty}^{\infty} I_n(u) z^n\,. \tag{5.3.4}$$

In terms of these Bessel functions the probability distribution takes the form

$$P(n,t) = e^{-(p+q)t} \sum_{l=-\infty}^{\infty} \left(\frac{p}{q}\right)^{(n-l)/2} P(l,0) I_{n-l}(2\sqrt{pq}t)\,. \tag{5.3.5}$$

For the special case of a symmetric random walker ($p = q = 1/2$) starting at the origin $n = 0$ (i.e. $P(n,0) = \delta_{n,0}$) one obtains the simple form

$$P(n,t) = I_n(t) e^{-t}\,. \tag{5.3.6}$$

5.4 Random walk and detailed balance

In the beginning we have characterized a nonequilibrium system by the absence of detailed balance (1.1.6) in the transition rates between the different states (see Sec. 1.1). Is the random walk indeed a nonequilibrium system in this sense, i.e. does it violate the detailed balance condition

$$w(C \to C')P(C) = w(C' \to C)P(C') \tag{5.4.1}$$

where $P(C)$ is the probability to find the system in state C (given by the Boltzmann distribution for equilibrium systems) and $w(C \to C')$ the transition rates between different states.
We consider the asymmetric random walk on a periodic lattice because of the simple result (5.2.5) for the probability distribution in the stationary state. It is independent of the position (which defines the state C of the system), i.e.

$$P(C) = P(n) = \frac{1}{L}\,. \tag{5.4.2}$$

The transition rates are given by

$$w(C \to C') = w(n \to n \pm 1) = \begin{cases} p & \text{for } "+" \\ q & \text{for } "-" \end{cases}\,. \tag{5.4.3}$$

Since $P(C)$ is constant (and non-zero) it drops out of (5.4.1) and the remaining detailed balance condition is

$$w(n \to n \pm 1) = w(n \pm 1 \to n)\,. \tag{5.4.4}$$

Due to (5.4.3) one site of this equation is equal to p, the other to q. Thus the random walk satisfies the detailed balance condition only in the symmetric case $p = q$. In the asymmetric case $p \neq q$ its stationary state is an nonequilibrium steady state.

Chapter 6

Ratchets

We have seen that diffusion is an isotropic process which can not generate directed motion. However, e.g. for biological systems directed motion is essential, e.g. in intracellular transport. Is it possible to build a rectifier for diffusion which leads to a directed transport? The answer to this question is the so-called **ratchet** or **Brownian ratchet** [3]. Before we discuss ratchets for diffusion in more detail we have a brief look at the famous **Feynman ratchet and pawl**.

6.1 Feynman's ratchet and pawl

In his lectures [4] Feynman used the ratchet and pawl to illustrate the laws of thermodynamics. In fact the system which was first introduced by M. Smoluchowski in 1912.
The **Feynman-Smoluchowski ratchet** is a simple machine which consists of a paddle wheel and a ratchet. The ratchet has asymmetric paddles so that the pawl only allows its motion in one direction[1] (Fig. 6.1.1).

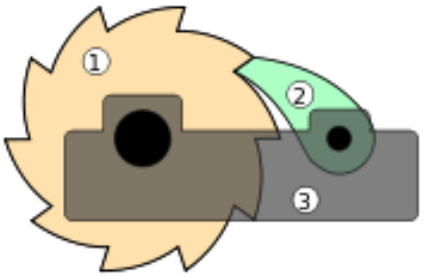

Figure 6.1.1: A ratchet consists of a wheel with asymmetric paddles. The pawl (2) allows its motion only in one direction. Here clockwise motion is not possible. (from [1])

The full Feynman-Smoluchowski ratchet is shown in Fig. 6.1.2. The ratchet is kept in a heat bath of temperature T_2 and coupled to a paddle wheel in a different heat bath with temperature

[1] You might find a ratchet also in your toolbox. It is a special kind of wrench.

T_2. It appears as it is possible to use this machine to extract useful work from heat at thermal equilibrium and lift the weight m. This would be a violation of Second Law of Thermodynamics. The idea behind this is the following. The molecules in the heat baths undergo random Brownian motion with a mean kinetic energy determined by its temperature. Assuming that the device is small enough so that the even single collisions with molecules can turn the paddle. These collisions tend to turn the paddle in both directions with the same probability. However, the ratchet prevents the motion in one direction. Effectively this appears to lead to a turing of the system in one direction, lifting the weight in the process.

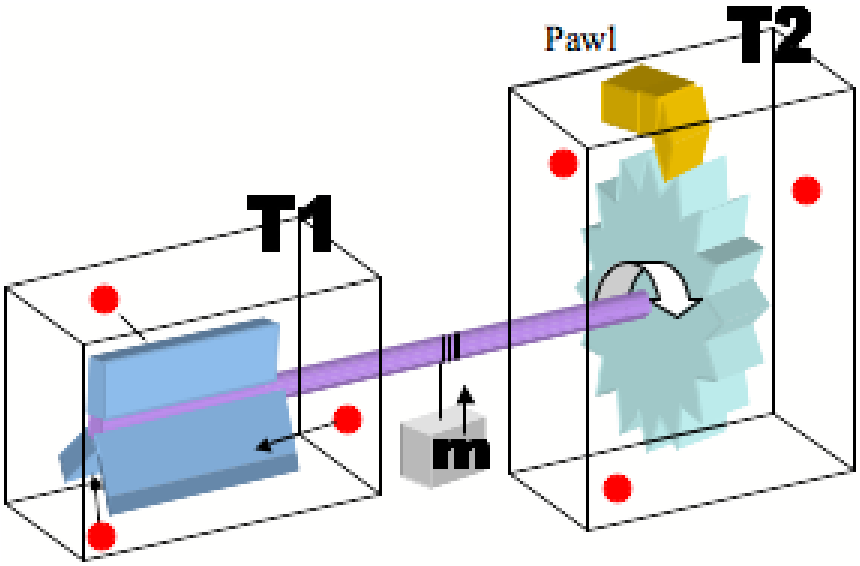

Figure 6.1.2: The Feynman-Smoluchowski ratchet consists of a paddle wheel (in heat bath of temperature T_1) and a ratchet (in heat bath of temperature T_2). At first sight it appears that this system can extract useful work (lifting the weight m) from heat (random fluctuations) in a system in thermal equilibrium. This would imply a violation of the Second Law of Thermodynamics. (from [1])

Feynman's analysis shows that this is not true and so the Second Law is not violated. Without going into the details, which can be found in [4] we just mention that one has also to consider collisions of the molecules with the pawl. These will lift the pawl from time to time allowing motion in the "forbidden" direction. Effectively no net rotation arise if the heat baths are at the same temperature $T_1 = T_2$.
Meanwhile there is even an experimental realization of the Feynman-Smoluchowski ratchet where the paddle and the ratchet are realized by special molecules [5]. The results are in agreement with the above analysis, i.e. no unidirectional motion was observed.

The Feynman-Smoluchowski ratchet is related to **Maxwell's demon**, a thought experiment developed by Maxwell in 1871. A box containing a case is divided into two compartments A and B, separated by a trap door (see Fig. 6.1.3). Initially both parts are in equilibrium, i.e. have the

same temperature. The trapdoor is guarded by an (imaginary) demon. He allows only molecules that are faster-than-average to move from A to B. So after some time the fast molecules are in B and the slow ones in A. This implies that the temperature in B is larger than that in A. Since (apparently) no work had to be spend this would be a violation of the Second Law.
The resolution of this problem is rather subtle. A proper analysis shows that it takes more thermodynamic work for the demon to determine the speed of the molecules (he needs some device for measuring molecular speeds!) than exergy[2] can be gained from the resulting temperature difference.

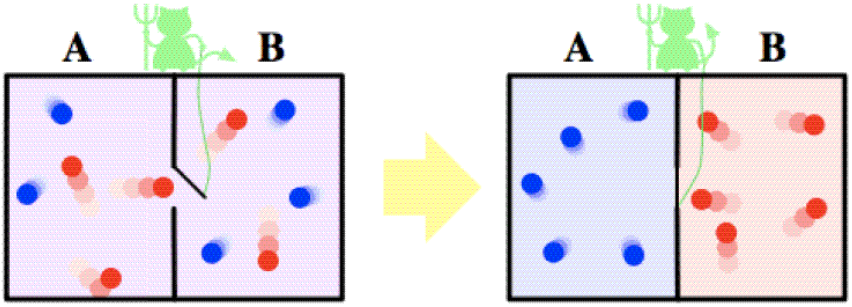

Figure 6.1.3: Maxwell's demon (green) opens the trapdoor between the compartments only for fast molecules (red) moving from the left part to the right. After some time, the temperature in B is larger than that in A. Without expenditure of work this would be a violation of the Second Law. (from [1])

6.2 Flashing ratchet

Now we are returning to the problem of rectification of diffusion. This can be achieved by using appropriate ratchet mechanisms. The simplest one is the so-called **flashing ratchet**. Here the diffusing particles are subjected to a ratchet potential which is switched on and off (explaining the name *flashing* ratchet). A **ratchet potential** $V(x)$ is in general periodic, time-dependent and not reflection symmetric. A typical form is shown in Fig. 6.2.1(a). Such potentials allow to generate directed motion by combining unoriented nonequilibrium fluctuations with spatial anisotropy (through the shape of the potential).
If the potential is switched on the particles are driven by the force $F = -V'(x)$ towards the minima of the potential (Fig. 6.2.1(a)). If now the potential is switched off, the particles start to diffuse isotropically (Fig. 6.2.1(b)). Next the potential is switched on again after some time. Again the particles are driven by the force to the minima of the potential. However due to the lack of reflection symmetry more particles will be captured in the minimum to the left than in the minimum to the right of the original position (Fig. 6.2.1(c)). Repeating this switching will

[2]Exergy is the maximum useful work possible during a process that brings a system in equilibrium with a heat reservoir.

then generate an effective current to the left (in general: in the direction of the maximum which is closer to the original position).

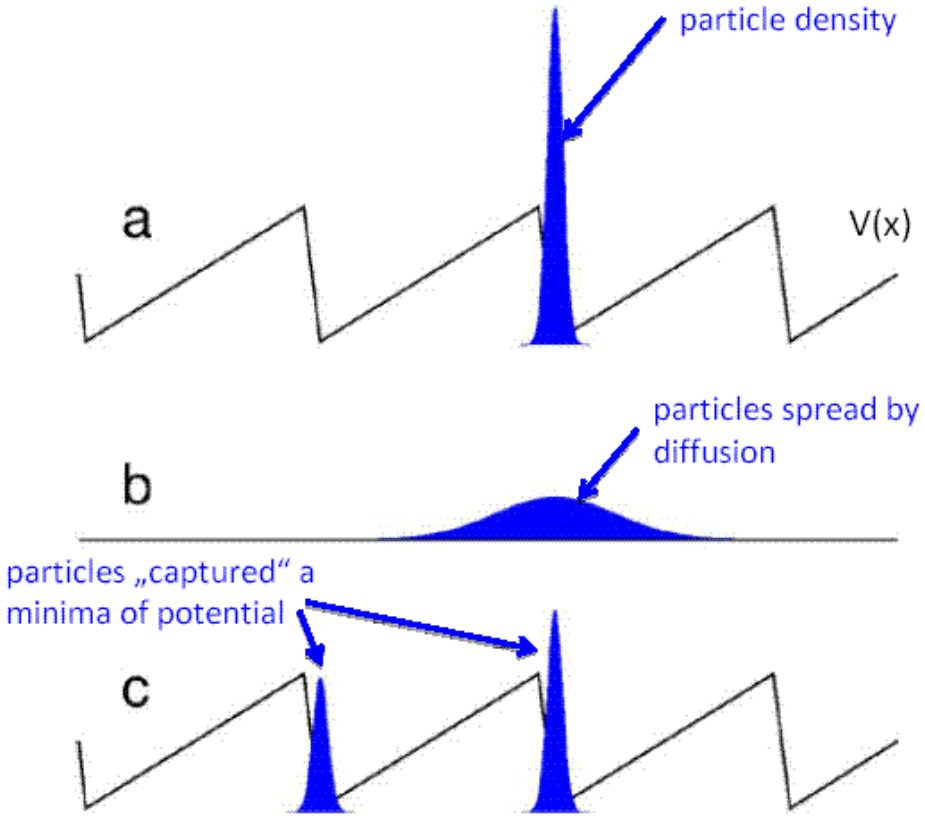

Figure 6.2.1: The flashing ratchet. (a) When the potential is switched on, particles move to minima of the $V(x)$; (b) after the potential is switched off ($V(x,t) = 0$) the particles start diffusing symmetrically; (c) if the potential is switched on again the particles are captured in the minima again; due to the asymmetry of $V(x)$ more a captured in the minimum to the left than that to the right of the original position.

6.3 Rocking ratchet

Another ratchet type is the **rocking ratchet**. It uses a potential of the same shape as the flashing ratchet (Fig. 6.3.1(a)). However, instead of switching the potential on and off it is rocked, i.e. tilted to the left and right (about the same angle). When the potential is tilted to the right (clockwise) by a suitable angle (not too large) particles located at some minimum can move since they always experience a force driving them back to the minimum (see Fig. 6.3.1(b)). On the hand,

tilting to the left (anti-clockwise) by the same angle the particle will move to the left. In this way a current to the left is generated.
Note that the motion in the rocking ratchet is into the opposite direction of that in the flashing ratchet. Particles move in the direction of maximum which is farther away from the original position.

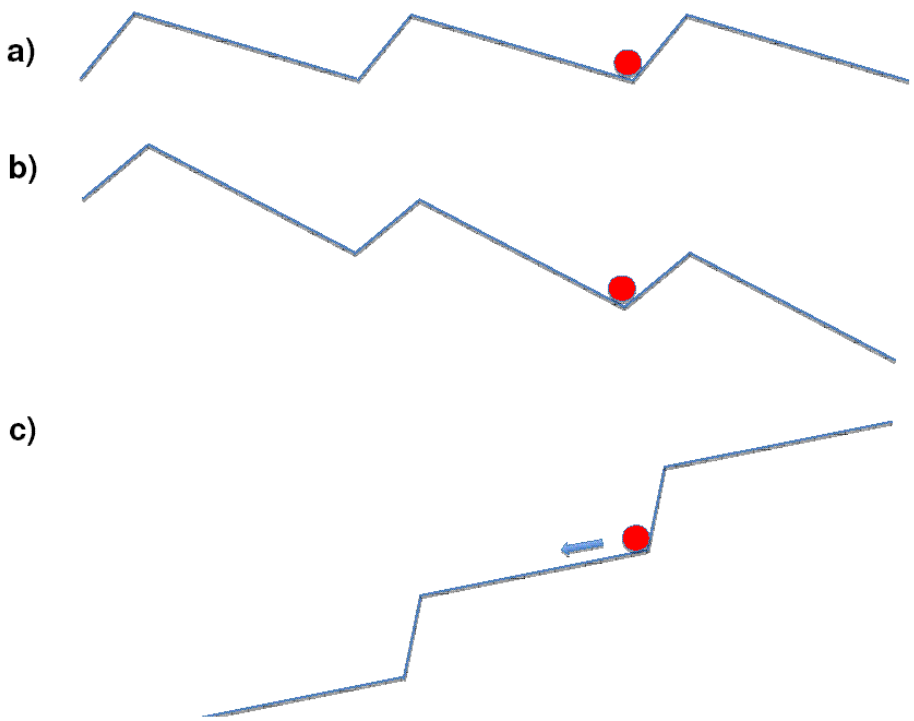

Figure 6.3.1: The rocking ratchet: a) neutral position; b) tilt clockwise: the particle moves to the right: c) tilt anti-clockwise (same angle as in b)): the particle can not move. Periodic clockwise - anti-clockwise (i.e. ”rocking”) tilting leads to a particle motion to the right.

Chapter 7

Universal properties of poker tournaments

In the following we will apply the theoretical framework developed so far to a non-physical problem, namely the chip distribution in poker tournaments. Similar to equilibrium statistical physics[1] our main interest is in **universal properties** of this distribution, i.e. those properties which do not depend on the *details* of the dynamics. This leads to an important simplification: we do not need a very detailed model implementing all poker rules. Instead a simpler one which captures the main features (which is usually the difficult part in modelling!) should be sufficient.

7.1 Poker basics

Poker is a card game[2] where each player gets a number of cards (typically 5). Depending on the type of poker played all or some of the cards are hidden from the other players and can only be seen by their owner. The value of a *hand* of cards is determined by the probability of the specific combination of cards: combinations that occur with a lower probability have a higher value. The players bet on their hand (a process called *bidding*) which brings in an element of psychology, e.g. by betting a large amount he/she can pretend to have a hand with high value forcing others to give in. This is called *bluffing*. For more details about the poker rules we refer to [1] where also a list of poker hands and their values (i.e. probabilities) can be found.
The process of bidding makes poker dynamics much more complex than that of a typical physical system. There are different types of players (aggressive, cautious, bluffing,...). Good players are adaptive and change their strategy during a tournament to adjust it to the strategies of their opponents. This makes the dynamical rules itself time-dependent, similar to time-dependent forces in classical mechanics.
Since we interested only in the universal properties of poker tournaments we neglect many of these details. Of course later one has to check, e.g. by comparing with empirical data, whether these simplifications are really justified!
In a poker tournament each player starts with the same amount of chips (or money). Players who lost all their chips are eliminated from the tournament. The last player remaining is the

[1] Observed typically at phase transitions
[2] Or rather a family of card games with slightly different rules.

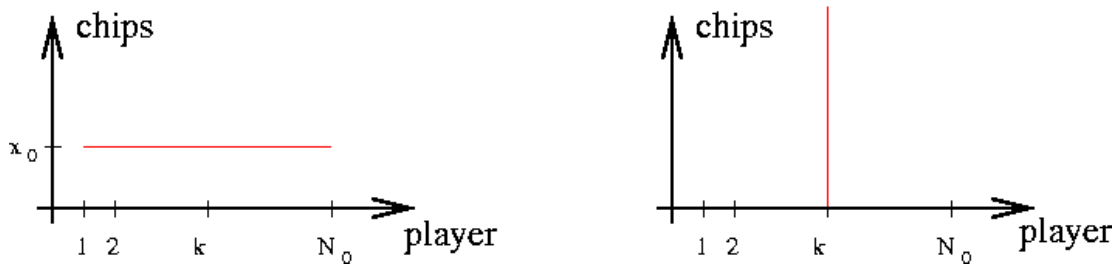

Figure 7.1.1: The initial and final chip distribution in a poker tournament (where player k is the winner).

winner and has won all the chips. The initial and final chip distribution is shown in Fig. 7.1.1. It looks just like the time-reversed version for the density distribution of a diffusion starting from a δ-function initial condition (see also Fig. 2.1.1)! This is already a first indication that a suitable modification of the methods developed for the description of diffusive processes could also be applicable to the present problem!

7.2 Simplified model for poker tournaments

C. Sire [2] has proposed a very simple model for poker tournaments which focusses on the bidding process and allows to predict the dynamics of the chip distribution.
The tournaments starts with N_0 players playing at tables with T players each. Each player initially has x_0 chips. In real tournaments (including Internet games) one has typically $N_0 \sim 10 - 10000$ and $T = 10$.
At each table separate games are played. The dynamics of the bidding process is shown in Fig. 7.2.1. Bidding procedes in a fixed order which is changed every game. The last bidder has to place a minimum bet called *blind* even before receiving his cards. This is to make the game more interesting because otherwise it could happen quite frequently that no one places a bet and there will be no game. All bets placed by the players during a game have to be at least the blind, not less. In so-called "no limit" games there is no upper limit to a bet, i.e. a player can bet all his/her chips at once. This is called *all in*. Of course players can always "pass", i.e. place no bet at all. In this case they are eliminated from that round.
The blind b, i.e. the minimum bet, usually increases exponentially in time in real tournaments:

$$b(t) = b_0 e^{t/t_0}\,, \tag{7.2.1}$$

(e.g. $40 - 60 - 100 - 150 - 200 - 300 - 500 - ...$ every 10-15 minutes). The growth rate t_0 is an important quantity as it allows to control the duration of the tournament: if t_0 is small the increase is faster and the tournament will be over more quickly. Typical values for the initial blind b_0 at the start of the tournament depend on the initial number of chips x_0 and are typically of the order $x_0/b_0 \sim 50 - 100$.

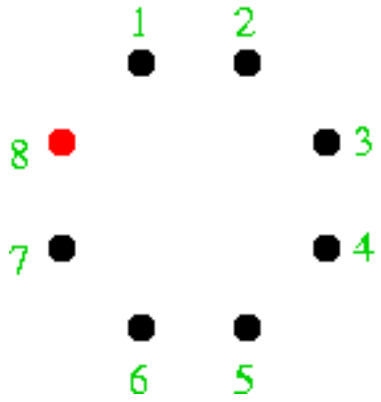

Figure 7.2.1: The (simplified) bidding dynamics used in the model. Green numbers indicate the order in which players put their bets. The red dot indicates the position of the blind which is a forced minimum bet in order to make the game more interesting.

Since the value of a hand depends on the probability of its occurrence it is natural to represent it by just a probability $c \in [0, 1]$. The full model for poker tournaments is then defined by the following rules where t is the time since the beginning of the tournament:

1. The blind bets $b(t) = b_0 e^{t/t_0}$

2. All players receive a hand represented by a random number c which is equally distributed in $[0, 1]$.

3. At each table, according to the bidding scheme in Fig. 7.2.1, players bet b chips with probability $e(c)$ if $0 < c < c_0$. This corresponds to a situation where the player considers his hand not to be too good, with the critical hand value c_0 being the boundary between not so and good hands. $e(c)$ is an evaluation function which is expected to be monotonically increasing: players will play good hands (i.e. larger c) more often than bad ones (i.e. smaller c). The precise form[3] of $e(c)$ turns out to be of no importance for the universal properties [2] we are interested in. Therefore we use

$$e(c) = c^n \tag{7.2.2}$$

   where n is the number of players that have already placed bets, including the blind. Therefore $n \geq 1$. This makes sense because now $e(c)$ can be interpreted as the probability that c is highest among $n + 1$ random cards. It captures the fact that players are more careful betting on bad hands if already some other players have placed bets before them.

4. The first player who has a hand with value larger than the critical hand value c_0 (i.e. $c > c_0$) goes all-in, i.e. bets all the chips he has left. Thus the probability that a player goes all-in is given by $q := 1 - c_0$.

[3]The determination of the optimal form of $e(c)$ is of course a problem of great practical importance for any poker player. It is extremely difficult for $T > 2$ and is studied within the field of game theory.

5. The following players, including the blind, can also go all-in if their $c > c_0$. Otherwise they will pass. In case they have less chips than the first all-in player they only bet this amount.

6. The highest c wins all the pot, i.e. the money bet by all the players in that game. In case nobody has placed a bet, the blind receives his bet b back.

7. Players who have no chips left are eliminated from the tournament.

8. After each round players are redistributed in such a way that the total number of tables is as small as possible, i.e. it is $[N(t)/T] + 1$ where $[.]$ denotes the integer part.

9. The time, or round index, t is increased by one: $t \to t + 1$.

The model contains a number of parameters which we collect in the following table.

parameter	**interpretation**
N_0	initial number players
$N(t)$	number players left at time t
x_0	initial number of chips of each player
T	maximum number of players at each table
b_0	initial blind
t_0	growth rate of blind
$b(t)$	blind at time t
$e(c)$	evaluation function
c_0	critical hand value
$q = 1 - c_0$	all-in probability

Of course many simplifying assumptions have been made:

- each player only receives one card represented by a single number;

- only three type of bids are possible: pass, bid the value of the blind or go all-in, i.e. the allowed bids for each player are 0, $b(t)$ and $x(t)$;

- there are no multiple bidding rounds, i.e. players can not increase their previous bids;

- it is not considered that for good players the evaluation function is time-dependent since during a tournament they ”learn” to interpret the actions of their opponents and change strategy accordingly;

- the element of ”bluffing”, i.e. placing high bets with a bad hand, is completed neglected in the model.

7.3 Analysis of the simplified model

In the following we will consider poker tournaments from three different points of view:

- empirical data from real poker tournaments
- computer simulations of the simplified model
- approximate analytical description of the simplified model

and compare these results. The comparison with the empirical data will tell us if the assumptions made in the derivation of the model are valid or not.

7.3.1 The case without all-in processes ($q = 0$)

First we will consider the (unrealistic) case where no player can go all-in, i.e. bet all his money. This corresponds to the case $q = 0$.
In the spirit of the Langevin equation we can then write down a dynamical equation for the number of chips[4] $x(t)$ of a player at time t:

$$x(t+1) = x(t) + \epsilon(t)b(t) \tag{7.3.1}$$

where $\epsilon(t)$ can be interpreted as a win-loss function: for $\epsilon(t) = -1$ the player loses the blind $b(t)$, for $\epsilon(t) = 1$ he wins $b(t)$. The average of $\epsilon(t)$ is assumed to be $\overline{\epsilon} = 0$ since there is no winning strategy in the mathematical sense for poker and all players are assumed to be equal.
The term $\epsilon(t)b(t)$ in (7.3.1) can be seen in analogy to the random collisions of a Brownian particle. However, here the strength of the collision term is time-dependent.
Since we neglect learning effects of the players, consecutive bids can be considered to be statistically independent. Therefore we deal with a Markov process (without memory).
If a typically value of x, which is of the order of the average $\overline{x}(t)$, is much larger than the value $b(t)$ of the blind, we can take the continuum limit (in time t) of (7.3.1) which leads to the Langevin equation

$$\frac{dx}{dt} = \sigma\eta(t)b(t) \tag{7.3.2}$$

where $\sigma^2 = \overline{\epsilon^2}$ and noise correlations $\langle\eta(t)\eta(t')\rangle = \delta(t-t')$. This is the evolution of a generalized Brownian particle (overdamped motion).
For the derivation of (7.3.2) we take the timestep to be Δt instead of 1 in (7.3.2) and than take the limit $\Delta t \to 0$. Taking into account $x(t+\Delta t) - x(t) \approx \frac{dx}{dt}\Delta t$ we arrive at (7.3.2) with $\sigma\eta(t)\Delta t = \epsilon(t)$.
The number $P(x,t)$ of surviving players with x chips then evolves according to a Fokker-Planck equation

$$\frac{\partial P}{\partial t} = \frac{\sigma^2 b^2(t)}{2}\frac{\partial^2 P}{\partial x^2} \tag{7.3.3}$$

[4]Note that $x(t) \geq 0$.

which can be derived in analogy to Sec. 4. The term $\partial P/\partial x$ is missing since the motion here is overdamped. The Fokker-Planck equation (7.3.3) has to satisfy the absorving boundary condition

$$P(x=0,t)=0 \tag{7.3.4}$$

since only the number of *surviving* players are considered and elimated players should have no influence on the dynamics. In addition we have the initial condition

$$P(x,t=0)=N_0\delta(x-x_0)\,. \tag{7.3.5}$$

The Fokker-Planck equation (7.3.3), e.g. using the method of images, and its solution is given by

$$P(x,t)=\frac{N_0}{\sqrt{2\pi\tau(t)}}\left(e^{-\frac{(x-x_0)^2}{2\tau(t)}}-e^{-\frac{(x+x_0)^2}{2\tau(t)}}\right) \tag{7.3.6}$$

where we have introduced

$$\tau(t)=\frac{1}{2}\sigma^2 b_0^2 t_0\left(e^{2t/t_0}-1\right)\,. \tag{7.3.7}$$

For large times $t\gg\tau$ the distribution of chips becomes scale invariant, i.e. it can be written in the form

$$P(x,t)=\frac{N(t)}{\overline{x}(t)}f\left(\frac{x}{\overline{x}(t)}\right) \tag{7.3.8}$$

where the density of the density of the surviving players is given by

$$\frac{N(t)}{N_0}=\frac{2x_0}{\sqrt{\pi t_0}\sigma b_0}e^{-t/t_0}\,. \tag{7.3.9}$$

This result has an interesting interpretation: according to (7.3.9) the decay constant for the number is t_0, i.e. the same as the growth rate of the blinds! Thus indeed t_0 controls the length of the tournament as we had originally claimed. Note that the equality of the decay constants for the blind and the number of surving players is a specific prediction of the model for $q=0$ which can be tested (in principle) with empirical data!
The total duration of a tournament t_f, obtained by taking $N(t_f)=O(1)$ in (7.3.9), is

$$\frac{t_f}{t_0}\simeq\ln(N_0)-\frac{1}{2}\ln(t_0)+\ln\left(\frac{x_0}{b_0}\right)\,. \tag{7.3.10}$$

It only grows logarithmically with the number of players and the ratio x_0/b_0.
The average number of chips per surving player is given by

$$\overline{x}(t)=\frac{N_0}{N(t)}x_0 \tag{7.3.11}$$

since N_0x_0 is the total number of chips. Explicitly one finds

$$\overline{x}(t)=\frac{\sqrt{\pi t_0}\sigma}{2}b(t)\,, \tag{7.3.12}$$

i.e. the average number of chips is proportional to the blind. For $t_0 \gg 1$ one has $\overline{x}(t)/b(t) \gg 1$ which was just the assumption made in the derivation of the continuum approach. This supports the consistency of our analysis.
The normalized chip distribution

$$f(X) = \overline{x}(t)\frac{P(x,t)}{N(t)} \qquad \text{with } X = \frac{x}{\overline{x}(t)} \tag{7.3.13}$$

is given by

$$f(X) = \frac{\pi}{2} X e^{-\pi X^2/4}, \qquad F(X) = \int_0^X f(Y)dY = 1 - e^{-\pi X^2/4} \tag{7.3.14}$$

which is the well-known **Wigner distribution** which e.g. occurs in quantum chaos or the distribution of nuclear energy levels. This form is truly universal as it is independent of all microscopic parameters x_0, t_0, $b_0 \ldots$
The results obtained from computer simulations of the full model show a very good agreement with the prediction (7.3.14) by our approximate analytical theory, see Fig. 7.3.1.

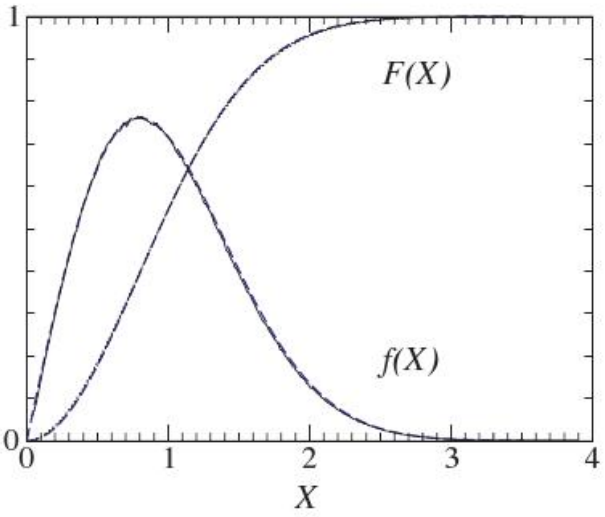

Figure 7.3.1: Normalized chip distribution $f(X)$ and its cumulative sum $F(X)$ for $q = 0$ obtained from computer simulations (full lines) with $N_0 = 10000$, $t_0 = 2000$, $x_0/b_0 = 100$ averaged over 10000 tournaments. The distributions have been extracted at times for which $N(t)/N_0 = 0.5\ 0.3$ and 0.1. For comparison, the Wigner distribution (7.3.14) of the approximate analytical theory is shown (broken lines). It shows almost perfect agreement. (from [2])

7.3.2 The case with all-in processes ($q > 0$)

Next we consider the case $q > 0$ where players might go all-in and bet all the chips they have left. Note that this only becomes relevant compared to the case $q = 0$ if at least two players go

all-in! If only one player goes all-in this is equivalent to betting just $b(t)$ because he can not win more than the blind from each other player.
We consider a fiexed table with T players and assume q to be small. Then the probability for an all-in process (i.e. at at least two players going all-in) is

$$P_{\text{all-in}} = q^2 \frac{T(T-1)}{2} + O(q^3)\,. \tag{7.3.15}$$

The factor q^2 is the probability that two players go all-in and $T(T-1)/2$ the number of possible pairs out of the T players. Processes where more than two players bet all their money are of order q^3 and will be neglected here (since q is assumed to be small).
Each of the two all-in players wins with probability 1/2. Therefore the player with less chips looses with probability 1/2 and gets eliminated from the tournament. Thus we can determine the decay rate for the number of players due to all-in process (averaged over all $\frac{N}{T}$ tables):

$$\left.\frac{dN}{dt}\right|_{\text{all-in}} = -\frac{1}{2}\cdot\frac{N}{T}\cdot P_{\text{all-in}} =: -\frac{N}{t_{\text{all-in}}} \qquad \text{with } t_{\text{all-in}} = \frac{4}{q^2(T-1)}\,. \tag{7.3.16}$$

Sire [2] argues that the optimal choice for the $t_{\text{all-in}}$ (and hence for q) is such that the decay due to all-in processes is the same as t_0, i.e. the decay induced by chip fluctuations due to non-all-in processes (which are of order $b(t)$). The total decay rate will be shown to remain equal to t_0^{-1}. One then has $t_{\text{all-in}} = 2t_0$ since the inverse decay rates add up. The choice for $t_{\text{all-in}}$ can be understood as follows:

- For $t_{\text{all-in}} < 2t_0$ the game would be dominated by all-in processes. Then $x(t)$ quickly becomes large compared to $b(t)$. In such a situation the first player to go all-in takes the risk of being eliminated (if another player also goes all-in) just to win a neglible blind.
- For $t_{\text{all-in}} > 2t_0$ players with few chips will often go all-in because they have a good chance to double their chips.

One would expect that real poker players, on average, self-adjust their q to the optimal value. Thus q is no longer a free parameter of the model, but takes the value

$$q = \sqrt{\frac{2}{(T-1)t_0}}\,. \tag{7.3.17}$$

We now look at the time evolution of the number $P(x,t)$ of remaining players:

$$\frac{\partial P}{\partial t} = \frac{\sigma^2 b^2}{2}\cdot\frac{dP}{dt}\frac{\partial^2 P}{\partial x^2} + \frac{2}{t_0}\left(K(P) - P\right)\,, \tag{7.3.18}$$

where the first term comes from non-all-in processes with pots of the order $b(t)$ (see equation (7.3.3)). In the second term comes the from the all-in processes and the factor $2/t_0$ is the rate of all-in processes involving the player considered. We have introduced the all-in kernel

$$K(P) := \frac{1}{4}P(x/2)\int_{x/2}^{\infty}\frac{P(y)}{N}dy + \frac{1}{2}\int_0^{x/2}P(x-y)\frac{P(y)}{N}dy + \frac{1}{2}\int_0^{\infty}P(x+y)\frac{P(y)}{N}dy\,. \tag{7.3.19}$$

The three terms have the following interpretation: the first term describes the doubling of chips ($x/2 \to x$) by winning against a player who has more chips ($> x/2$). The second term comes from a win against a player with fewer chips, who then gets eliminated from the tournament. Finally, the last term originates in processes where the player looses against a player who has less chips. Note that the factor $1/N$ has to be included since it is assumed that the player only has one opponent who goes all in.
Integrating (7.3.19) over all x we find

$$\int K(P)dx = \frac{3}{4} \tag{7.3.20}$$

since the first two terms add up to 1/2 which is the probability to win in an all-in with two players. Thus the probability to stay in the tournament after an all-in event is 3/4. We can also determine the decay rate due all-in processes. Since the rate of all-in processes is $/t_0$ it is given by

$$\left(1 - \frac{3}{4}\right)\frac{2}{t_0} = t_{\text{all-in}}^{-1}\,. \tag{7.3.21}$$

This is consistent with the assumption made above.
In order to find a solution for (7.3.19) we make the Ansatz (of scaling form)

$$P(x,t) = \frac{\lambda}{\hat{x}^2(t)} f\left(\frac{x}{\hat{x}(t)}\right)\,. \tag{7.3.22}$$

The integral of f is assumed to be normalized to 1, so that

$$N(t) = \frac{\lambda}{\hat{x}(t)}\,. \tag{7.3.23}$$

Inserting this Ansatz into (7.3.19) one obtains an integro-differential equation for which no solution is known. However, the asymptotic behaviour for small and large $X = \frac{x}{\hat{x}(t)}$:

$$f(X) \sim \frac{X}{2} \qquad \text{for } X \ll 1 \tag{7.3.24}$$

$$f(X) \sim 2\mu e^{-}\mu X \qquad \text{for } X \gg 1 \tag{7.3.25}$$

where $\mu \approx 1.562$. Thus the scaling function decays slower than in the case $q = 0$. The result is in good agreement with computer simulations (Fig. 7.3.2)

Finally one can determine the average number L_{N_0} of chip leaders[5] during a tournament with N_0 players. It is found that for both $q = 0$ and $q > 0$ it grows logarithmically,

$$L_{N_0} \propto \ln N_0\,, \tag{7.3.26}$$

which is a behaviour well-known from other models, e.g. models for evolution.

[5]The chip leader is the player with the most number of chips.

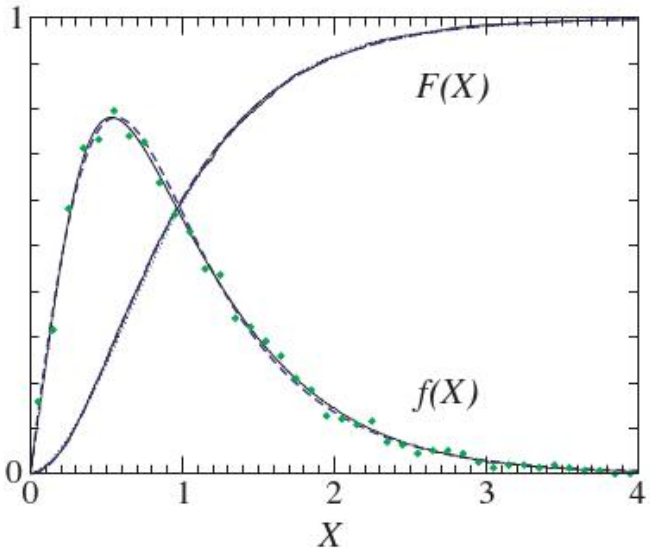

Figure 7.3.2: Normalized chip distribution $f(X)$ and its cumulative sum $F(X)$ for $q > 0$ obtained from computer simulations (full lines). The other parameters are the same as in Fig. 7.3.1. For comparison, the numerical solution of the integro-differential equation for $f(X)$ is shown (dashed lines). The circles are data from 20 real poker tournaments. (from [2])

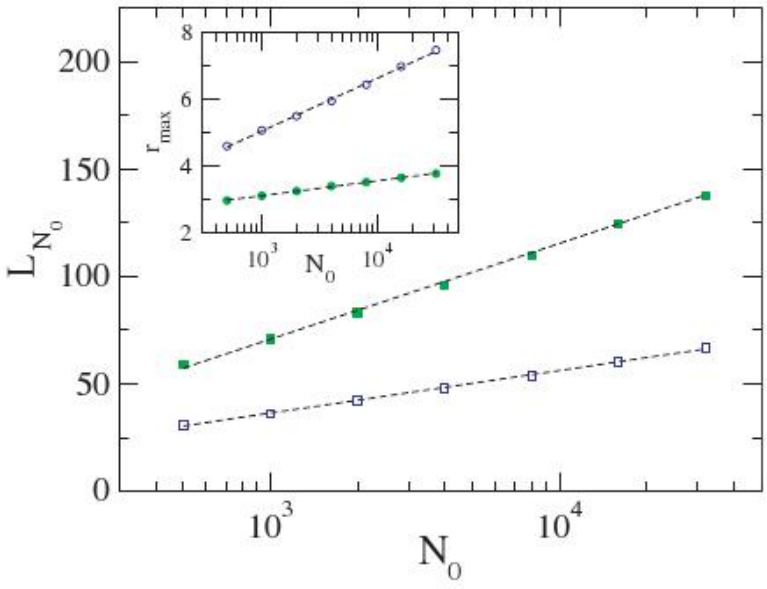

Figure 7.3.3: The average number of chip leaders L_{N_0} grows logarithmically as a function of N_0. Full symbols correspond to the case $q = 0$, open ones to $q > 0$. (from [2])

Chapter 8

Master equation

We have already seen that the master equation can be usefull tool to determine the time evolution of systems governed by stochastic dynamics. In the following we will introduce the general formalism and consider a few examples.

8.1 The master equation in continuous and discrete time

We consider a stochastic process which is defined by transition rates $W(C \to C')$. Here C and C' are states (or configurations) of our system. In the random walk studied in Sec. 5.2 (see eq. (5.2.1)) these states were simply labeled by the position of the particle. For many-particle systems the configuration becomes much larger.

A simple example for a stochastic many-particle system are interacting random walkers on a discrete lattice of L sites (Fig. 8.1.1). Each state can be characterized by the occupation numbers n_j of the sites $j = 1, 2, \ldots, L$ where $n_j = 1$ if site j is occupied by a particle and $n_j = 0$ if it is empty. The full state or configuration of the system is then given by the vector $\mathbf{n} := (n_1, n_2, \ldots, n_L)$.

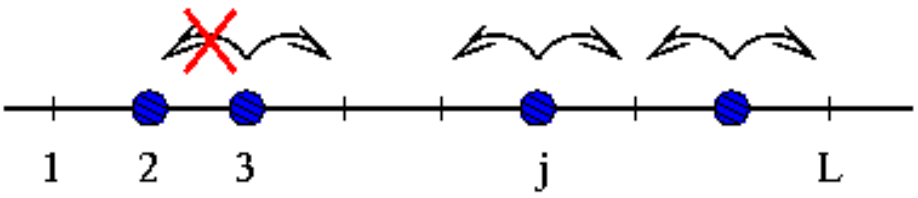

Figure 8.1.1: Interaction random walkers. Particles can move to the left and neighbour sites, but only if these are not occupied by another particle.

8.1.1 Master equation for continuous time

We are now interested in the probability $P(\mathbf{n}, t)$ to find the system in state $\mathbf{n}$ at time t. Its time evolution is determined by the **master equation**

$$\boxed{\frac{\partial P(\mathbf{n}, t)}{\partial t} = \sum_{\mathbf{n}'} w(\mathbf{n}' \to \mathbf{n}) P(\mathbf{n}', t) - \sum_{\mathbf{n}'} w(\mathbf{n} \to \mathbf{n}') P(\mathbf{n}, t)} \tag{8.1.1}$$

which is a rate equation taking into account the number of transitions per time between different states $\mathbf{n}$, $\mathbf{n}'$. The interpretation of (8.1.1) is basically the same as that given for (5.2.1). The first sum collects the **gain terms**, i.e. all transitions from states $\mathbf{n}'$ to the state $\mathbf{n}$ whereas the second sum are the **loss terms**, transitions from $\mathbf{n}$ to some other state $\mathbf{n}'$.
Master equations of the form (8.1.1) are valid for dynamical systems with a continuous time variable t. This can not directly be realized in computer simulations which usually proceeds in discrete steps. Therefore continuous time dynamics is usually realized by the so-called **random-sequential update**. Here in each step a site j is chosen randomly (with equal probability) and then its state is updated according to the transition rules. Note that this usually requires the update of neighbouring site (the target site of a moving particle). This is repeated over and over[1].
The master equation (8.1.1) can be written in a more compact form by introducing the matrix M, sometimes called **Markov matrix**

$$M_{\mathbf{n},\mathbf{n}'} = w(\mathbf{n}' \to \mathbf{n}) - \delta_{\mathbf{n}\mathbf{n}'} \sum_{\mathbf{n}''} w(\mathbf{n} \to \mathbf{n}'') \,. \tag{8.1.2}$$

Then the master equation can be written as a matrix equation

$$\boxed{\dot{\mathbf{P}}(t) = M \cdot \mathbf{P}(t)} \tag{8.1.3}$$

where $\mathbf{P}$ is the vector with the components $P(\mathbf{n}, t)$.
The Markov matrix M has some special properties which follow from the fact that it describes the dynamics of a stochastic system. All non-diagonal elements of M are non-negative since they are transition rates. Furthermore the sum of all elements in each column of M is zero, i.e. $\sum_i M_{ij} = 0$. This is related to the conservation of probabilities. Matrices with these properties are usually called **stochastic matrices** [6]. They should not be confused with *random matrices* which have elements that are stochastic variables!

8.1.2 Master equation for discrete time

Next we consider the case of discrete time dynamics. This is often used in the applications that we will study later in the this course. Instead of being a continuous variable, time proceeds in discrete timesteps Δt:

$$t_n = n\Delta t \qquad (n \in \mathbb{N}) \,. \tag{8.1.4}$$

[1] It is possible (but unlikely for large systems) that in the next step the same site j will be updated again!

The **master equation in discrete time** then takes the form

$$\boxed{P(\mathbf{n}, t + \Delta t) = \sum_{\mathbf{n}'} W(\mathbf{n}' \to \mathbf{n}) P(\mathbf{n}', t)} \tag{8.1.5}$$

where now $W(\mathbf{n}' \to \mathbf{n})$ are no longer transition rates (with the dimension of an inverse time), but dimensionless **transition probabilities**.
The interpretation of (8.1.5) is rather straightforward. The sum on the right-hand-side runs over all possible states $\mathbf{n}'$ at time t. The probability that the state is changed from $\mathbf{n}'$ to $\mathbf{n}$ in the timestep $t \to t + \Delta t$ is then proportional to the probability $P(\mathbf{n}', t)$ of finding the system in state $\mathbf{n}'$ times the transition probability $W(\mathbf{n}' \to \mathbf{n})$. Probability conservation implies

$$\sum_{\mathbf{n}'} W(\mathbf{n} \to \mathbf{n}') = 1 \tag{8.1.6}$$

since the state $\mathbf{n}$ has to evolve in *some* state during the timestep Δt.

One would expect that in the limit of small timesteps $\Delta t \to 0$ one recovers the continuous time master equation (8.1.1). This will be checked now explicitly.
First we expand

$$P(\mathbf{n}, t + \Delta t) \approx P(\mathbf{n}, t) + \frac{\partial P(\mathbf{n}, t)}{\partial t} \Delta t \tag{8.1.7}$$

which we then use in

$$\begin{aligned} \frac{\partial P(\mathbf{n}, t)}{\partial t} &= \lim_{\Delta t \to 0} \frac{P(\mathbf{n}, t + \Delta t) - P(\mathbf{n}, t)}{\Delta t} \\ &= \sum_{\mathbf{n}'} w(\mathbf{n}' \to \mathbf{n}) P(\mathbf{n}', t) - \left[\sum_{\mathbf{n}'} w(\mathbf{n} \to \mathbf{n}') \right] P(\mathbf{n}, t) \end{aligned} \tag{8.1.8}$$

where in the second step we have used the discrete master equation (8.1.5) and the probability conservation (8.1.6) (the term in $[\ldots]$ is equal to 1). By comparison with the continuous time master equation (8.1.1) we find a relation between the transition rates in the continuous case and the transition probabilities in the disrete case:

$$w(\mathbf{n}' \to \mathbf{n}) = \lim_{\Delta t \to 0} \frac{W(\mathbf{n}' \to \mathbf{n})}{\Delta t} \,. \tag{8.1.9}$$

Discrete time dynamics is easy to realize in computer simulations through the so-called **parallel (or synchronous) update**. In contrast to the random-sequential update the dynamical variables are not updated in random order, but all at the same time[2].

[2] Since a computer program usually works sequentially this requires some tricks when writing the computer code.

8.2 Derivation from Chapman-Kolmogorov equation

The master equation is equivalent to the so-called Chapman-Kolmogorov equation for Markov processes, i.e. stochastic processes "without memory" [6]. Compared to the Chapman-Kolmogorov equation the master equation is easier to handle and more natural for applications to physical problems.
Let us first define a Markov process more explicitly by considering the conditional probability density $P_{n-1}(C_n, t_n|C_1, t_1; C_2, t_2, \ldots, C_{n-1}, t_{n-1})$ to find the system at time t_n in state C_n provided that it was in state C_j at time t_j for $j = 1, 2, \ldots, n-1$ with $t_1 < t_2 < \ldots < t_n$. A **Markov process** has no memory which implies that this conditional probability density does not depend on the states at earlier times $t_1, t_2, \ldots, t_{n-2}$, i.e.

$$P_{n-1}(C_n, t_n|C_1, t_1; C_2, t_2, \ldots, C_{n-1}, t_{n-1}) = P_1(C_n, t_n|C_{n-1}, t_{n-1})\,. \tag{8.2.1}$$

The quantity $P_1(C_n, t_n|C_{n-1}, t_{n-1})$ is just the **transition probability**. Thus a Markov process is fully determined by the functions $P(C_n, t)$ and $P_1(C_n, t_n|C_{n-1}, t_{n-1})$.
The **Chapman-Kolmogorov equation**[3] is now given by

$$P_1(C_3, t_3|C_1, t_1) = \sum_{C_2} P_1(C_3, t_3|C_2, t_2) P_1(C_2, t_2|C_1, t_1) \qquad (t_1 < t_2 < t_3)\,. \tag{8.2.2}$$

It expresses the fact that a stochastic process starting in state C_1 at time t_1 reaches the state C_3 at time t_3 via any one of the possible intermediate states C_2 at time t_2 (Fig. 8.2.1).

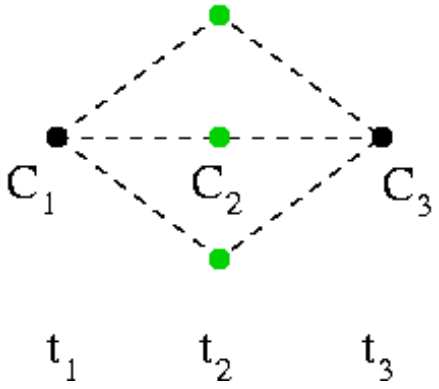

Figure 8.2.1: Illustration of the Chapman-Kolmogorov equation. The stochastic process starting in state C_1 at time t_1 can reach the state C_3 at a later time t_3 via any of the possible intermediate states C_2 at time t_2.

Considering Markov processes which are homogeneous in time, i.e. satisfy

$$P_1(C, t|C_0, t_0) = P(C, C_0, t - t_0)\,, \tag{8.2.3}$$

[3] Its continuous time analogue is sometimes called **Smoluchowski equation** or **coagulation equation**.

one can derive the master equation from the Chapman-Kolmogorov equation. First we expand for small Δt

$$P(C, C_0, t + \Delta t) = \delta(C_2 - C_1) + w(C_0 \to C)\Delta t + O\left((\Delta t)^2\right) \tag{8.2.4}$$

with some function $w(C_0 \to C)$, the transition rate. The δ-function expresses the fact that the probability to stay at the same state has to be 1 for $\Delta t = 0$.
Equation (8.2.4) needs to be slightly corrected in order to satisfy the normalization condition

$$\sum_C P(C, C_0, t) = 1 \tag{8.2.5}$$

which becomes clear when remembering that $P(C, C_0, t)$ is actually the conditional probability $P_1(C, t|C_0, t_0)$.
Equation (8.2.4) and (8.2.5) then imply that

$$P(C, C_0, t + \Delta t) = (1 - \alpha\Delta t)\delta(C_2 - C_1) + w(C_0 \to C)\Delta t + O\left((\Delta t)^2\right) \tag{8.2.6}$$

with

$$\alpha(C_0) = \sum_C w(C_0 \to C)\,. \tag{8.2.7}$$

If we now insert (8.2.6) into the Chapman-Kolmogorov equation (8.2.2) (keeping (8.2.3 in mind!) and taking the limit $\Delta t \to 0$ we obtain

$$\frac{\partial P}{\partial t}(C_3, C_1, t) = \sum_{C_2} \left[w(C_2 \to C_3)P(C_2, C_1) - w(C_3 \to C_2)P(C_3, C_1)\right]\,. \tag{8.2.8}$$

8.3 Stationary state and detailed balance

Generically, the properties of stochastic systems become time-independent in the long-time limit $t \to 0$, i.e. they reach a **stationary state** (or **steady state**)

$$P(\mathbf{n}) = \lim_{t\to\infty} P(\mathbf{n}, t)\,. \tag{8.3.1}$$

It is the solution of the **stationary master equation**

$$\sum_{\mathbf{n}'} w(\mathbf{n}' \to \mathbf{n})P(\mathbf{n}') = \sum_{\mathbf{n}'} w(\mathbf{n} \to \mathbf{n}')P(\mathbf{n}) \tag{8.3.2}$$

which is obtained by putting $\frac{\partial P(\mathbf{n},t)}{\partial t} = 0$ in (8.1.1).
A simple solution of (8.3.2) would be a situation where corresponding terms in the sums of the right- and left-hand-side are equal. This is called **detailed balance**[4]:

$$\boxed{w(\mathbf{n}' \to \mathbf{n})P(\mathbf{n}') = w(\mathbf{n} \to \mathbf{n}')P(\mathbf{n}) \qquad \text{for all } \mathbf{n},\ \mathbf{n}'\,.} \tag{8.3.3}$$

[4]See our previous discussion in Sec. 1.1 and 5.4.

This condition implies for all pairs of states $\mathbf{n}$, $\mathbf{n}'$ the transitions from $\mathbf{n}'$ to $\mathbf{n}$ per unit time must balance the transitions from $\mathbf{n}$ to $\mathbf{n}'$, i.e. there is no probability current flowing between these states. The system then can not be distinguished from its time-reversed variant. Therefore it is clear that not all stationary states satisfy detailed balance! We have already argued in Sec. 1.1 that it is *the* characteristic property of systems in equilibrium.
The detailed balance condition in the form (8.3.3) is not very practical! In order to check whether it is satisfied or not one already needs to know the exact stationary state $P(\mathbf{n})$. This problem is solved by the **Kolmogorov criterion**. It states that a stochastic process fulfills the detailed balance condition (8.3.3) iff for *all* cycles $\mathbf{n}_1 \to \mathbf{n}_2 \to \ldots \to \mathbf{n}_k \to \mathbf{n}_1$ in configuration space the transition rates satisfy

$$w(\mathbf{n}_1 \to \mathbf{n}_2)w(\mathbf{n}_2 \to \mathbf{n}_3)\cdots w(\mathbf{n}_k \to \mathbf{n}_1) = w(\mathbf{n}_1 \to \mathbf{n}_k)w(\mathbf{n}_k \to \mathbf{n}_{k-1})\cdots w(\mathbf{n}_2 \to \mathbf{n}_1)\,. \tag{8.3.4}$$

This criterion no longer involves the stationary solution, but only the transition rates which define the stochastic process!
In order to prove the criterion we first assume that our system satisfies detailed balance and show that (8.3.4) is a necessary condition. After a little rearranging of factors we get for the ration of the left- and right-hand-side of (8.3.4):

$$\frac{w(\mathbf{n}_1 \to \mathbf{n}_2)}{w(\mathbf{n}_2 \to \mathbf{n}_1)} \cdot \frac{w(\mathbf{n}_2 \to \mathbf{n}_3)}{w(\mathbf{n}_3 \to \mathbf{n}_2)} \cdot \ldots \frac{w(\mathbf{n}_k \to \mathbf{n}_1)}{w(\mathbf{n}_1 \to \mathbf{n}_k)} = \frac{P(\mathbf{n}_2)}{P(\mathbf{n}_1)} \cdot \frac{P(\mathbf{n}_3)}{P(\mathbf{n}_2)} \cdot \ldots \frac{P(\mathbf{n}_1)}{P(\mathbf{n}_k)} = 1 \tag{8.3.5}$$

where, in the first step, we have used the detailed balance condition (8.3.3).
The Kolmogorov criterion is also a sufficient condition, i.e. we can derive the detailed balance condition (and the stationary state) from it. We start with an arbitrary state $\mathbf{n}_1$ which has a stationary weight[5] $P(\mathbf{n}_1)$. Then we choose another state $\mathbf{n}_2$ which is connected with $\mathbf{n}_1$, i.e. $w(\mathbf{n}_1 \to \mathbf{n}_2) \neq 0$. For this state we define

$$P(\mathbf{n}_2) = P(\mathbf{n}_1)\frac{w(\mathbf{n}_1 \to \mathbf{n}_2)}{w(\mathbf{n}_2 \to \mathbf{n}_1)}\,. \tag{8.3.6}$$

Iterating this procedure yields

$$P(\mathbf{n}_k) = P(\mathbf{n}_1)\frac{w(\mathbf{n}_1 \to \mathbf{n}_2)\cdots w(\mathbf{n}_{k-1} \to \mathbf{n}_k)}{w(\mathbf{n}_2 \to \mathbf{n}_1)\cdots w(\mathbf{n}_k \to \mathbf{n}_{k-1})}\,. \tag{8.3.7}$$

Since the weight (probability) $P(\mathbf{n}_k)$ should be uniquely determined, it has to be independent of the path taken from $\mathbf{n}_1$ to $\mathbf{n}_k$, i.e. especially we have

$$P(\mathbf{n}_k) = P(\mathbf{n}_1)\frac{w(\mathbf{n}_1 \to \mathbf{n}_k)}{w(\mathbf{n}_k \to \mathbf{n}_1)}\,. \tag{8.3.8}$$

Combining (8.3.7) and (8.3.8) gives the criterion (8.3.4).

[5] *weight* here means an unnormalized probability.

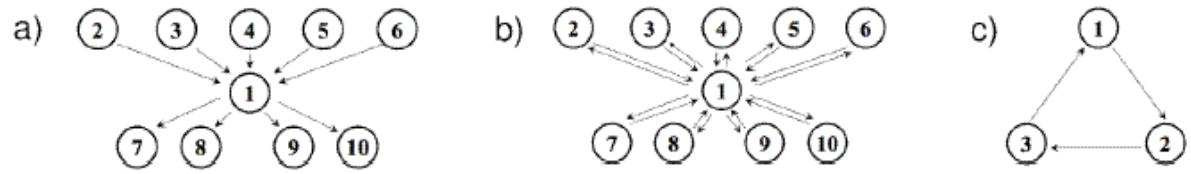

Figure 8.3.1: Illustration of detailed and pairwise balance.

We have already argued that detailed balance is characteristic for systems in thermal equilibrium. Therefore it should be possible to show that the stationary $P(\mathbf{n}_j)$ have in fact the form of Boltzmann factors. This requires the definition of a suitable energy function $E(\mathbf{n})$:

$$E(\mathbf{n}') - E(\mathbf{n}) := \frac{1}{\beta} \ln \frac{w(\mathbf{n}' \to \mathbf{n})}{w(\mathbf{n} \to \mathbf{n}')} \,. \tag{8.3.9}$$

Using (8.3.8) this indeed shows that the stationary probabilities follow a **Boltzmann distribution**

$$P(\mathbf{n}) \propto e^{-\beta E(\mathbf{n})} \,. \tag{8.3.10}$$

In some situations where the strong condition (8.3.3) of detailed balance does not hold, the weaker **pairwise balance condition** (sometimes called **dynamical reversibility**) might be fulfilled. It requires that for each pair of states $\mathbf{n}_1$ and $\mathbf{n}_2$ a unique state $\mathbf{n}_3$ exists so that

$$w(\mathbf{n}_1 \to \mathbf{n}_2) P(\mathbf{n}_1) = w(\mathbf{n}_3 \to \mathbf{n}_1) P(\mathbf{n}_3) \tag{8.3.11}$$

Systems satisfying pairwise balance are not in equilibrium. However, due to the rather simple structure of the condition some general conclusion about such system can be drawn. The differences between the conditions (8.3.3) and (8.3.11) are explained graphically in Fig. 8.3.1.
Consider the ten configurations in Fig. 8.3.1a) labelled by the integers $1, ...10$. If the only transitions into and out of the configuration 1 are those shown in Fig. 8.3.1, then $dP_1/dt = 0$ provided $\sum_{j=2,..,6} w(j \to 1) P_j = \sum_{j=7,..,10} w(1 \to j) P_1$. In other words, for any arbitrary configuration i, stationarity of P_i is satisfied if merely the total probability current into the configuration i is exactly balanced by the total probability current out of i.
In contrast, the equilibrium, which is a special stationary state, requires detailed balance (see Fig. 8.3.1b): the net forward and reverse probability current in beween the members of each pair of configurations must balance exactly. Clearly this is a much stronger condition than what was needed to guarantee a non-equilibrium stationary state. Finally, for many models a pairwise-balance is also satisfied by the non-equilibrium stationary state (see Fig. 8.3.1c): The probability current from configuration 1 to 2 is exactly balanced by that from 2 to 3 so that $dP_2/dt = 0$. Similarly, the probability current 2 to 3 is equal to that from 3 to 1, and so on.

Chapter 9

Quantum formalism

In the following we will develop the formalism for the master equation a little bit further. By introducing a suitable notation certain similarities with quantum mechanics are revealed. This will later allow to use methods originally developed for quantum system to be applied to stochastic systems as well.

9.1 Master equation in quantum formalism

The **quantum formalism** is a reformulation of the master equation that emphasizes similarities with quantum mechanics. To be concrete we consider a discrete stochastic system where the state at each site j ($j = 1, \ldots, L$) is specified by a state variable n_j. In the following, typically n_j will be discrete, e.g. an occupation number. The time-evolution of the probability $P(\mathbf{n}, t)$ to find the system in the configuration $\mathbf{n} = (n_1, \ldots, n_L)$ is determined by the master equation (8.1.1).
First we introduce vectors $|\mathbf{n}\rangle = |n_1, ..., n_L\rangle$ corresponding to the configurations $\mathbf{n}$. These form an orthonormal basis of the configuration space. Then state vectors can be defined by

$$|P(t)\rangle = \sum_{\mathbf{n}} P(\mathbf{n}, t)|\mathbf{n}\rangle \tag{9.1.1}$$

where the coefficients of the basis vectors are the probabilities to find the system in the corresponding basis state. As in quantum mechanics, sometimes it is useful to interpret (9.1.1) as a vector

$$|P(t)\rangle = \begin{pmatrix} P(0,0,\ldots,0) \\ P(0,0,\ldots,1) \\ \vdots \\ P(1,1,\ldots,1) \end{pmatrix} \tag{9.1.2}$$

where we have assumed that $n_j = 0, 1$.
Next we introduce the **stochastic Hamiltonian** $\mathcal{H}$, also known as **generator** of the stochastic

process, **Liouville operator** or **Markov matrix**. It is defined by its matrix elements

$$\boxed{\langle \mathbf{n}'|\mathcal{H}|\mathbf{n}\rangle = \begin{cases} -w(\mathbf{n}\to\mathbf{n}') & \text{for } \mathbf{n}\neq\mathbf{n}' \\ \sum_{\mathbf{n}''\neq\mathbf{n}} w(\mathbf{n}\to\mathbf{n}'') & \text{for } \mathbf{n}=\mathbf{n}' \end{cases},} \tag{9.1.3}$$

i.e. up an overall sign the stochastic Hamiltonian is exactly the Markov matrix introduced in (8.1.2). The off-diagonal elements of $\mathcal{H}$ have to be non-positive and, as for the case of the Markov matrix, the sum of the elements in each column is 0 due to probability conservation, i.e. the stochastic Hamiltonian is also a stochastic matrix.
Using this notation, the master equation (8.1.1) can be rewritten as

$$\boxed{\frac{\partial}{\partial t}|P(t)\rangle = -\mathcal{H}\,|P(t)\rangle\,,} \tag{9.1.4}$$

which has the form of a Schrödinger equation in imaginary time[1].
Despite these similarities with standard quantum mechanics, one important difference has to be emphasized: the stochastic Hamiltonian $\mathcal{H}$ is in general not hermitian, as can be seen from the definition (9.1.3). This is a consequence of the fact that we are dealing with nonequilibrium systems and will have important implications.

9.1.1 An example: Aymmetric random walk

For illustration we will consider the asymmetric random walk. The state space is rather simple. As basis vectors we choose $\langle n|$ which corresponds to the state where the random walker is at site n. Thus the state space is L-dimensional where L is the number of lattice sites. The non-zero transition probabilities are $w(n\to n+1)=p$ and $w(n\to n-1)=q$. The stochastic Hamiltonian is then of the following form:

$$\begin{aligned} \mathcal{H} &= \begin{pmatrix} \sum_{j\neq 1} w(1\to j) & -w(2\to 1) & \cdots & -w(L\to 1) \\ -w(1\to 2) & \sum_{j\neq 2} w(2\to j) & \cdots & -w(L\to 2) \\ -w(1\to 3) & -w(2\to 3) & \ddots & \\ \vdots & \vdots & \ddots & \\ -w(1\to L) & -w(2\to L) & & \sum_{j\neq L} w(L\to j) \end{pmatrix} \\ &= \begin{pmatrix} p+q & -q & 0 & \cdots & -p \\ -p & p+q & -q & \cdots & 0 \\ 0 & -p & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & \\ -q & 0 & \cdots & & p+q \end{pmatrix} \end{aligned} \tag{9.1.5}$$

As one can easily see, the stochastic Hamiltonian of the asymmetric random walk ($p\neq q$) is not hermitian.

[1] Replacing $t\to it$ brings it into the standard Schrödinger form.

The matrix (9.1.5) has a special form. Its matrix elements are constant along each diagonal, i.e. the matrix elements depend only on the difference of the indices ($a_{ij} = a_{i-j}$). Such matrices are called **Toeplitz matrices**. Since it is also cyclic, the stochastic Hamiltonian of the random walk belongs to the class of **circulent matrices**.

9.2 Properties of the stochastic Hamiltonian

For the stationary state, where $\partial P/\partial t = 0$, the master equation (9.1.4) reduces to

$$\mathcal{H}\,|P_0\rangle = 0 \tag{9.2.1}$$

which implies that the stationary state $\langle P_0|$ is an eigenvector of the stochastic Hamiltonian $\mathcal{H}$ with eigenvalue 0.
Does such an eigenvector exist? To prove this we first define a (bra-) vector

$$\langle 0| := \sum_{\mathbf{n}} \langle \mathbf{n}| \,. \tag{9.2.2}$$

In matrix notation this vector would correspond to the row vector $(1, 1, \ldots, 1)$. It is now easily checked that

$$\langle 0|\mathcal{H} = 0 \tag{9.2.3}$$

since the sum in each column of $\mathcal{H}$ is 0. In other words: the vector $|0\rangle$ is a left eigenvector of the stochastic Hamiltonian with eigenvalue 0. Thus also a corresponding right eigenvector $|P_0\rangle$ has to exist. Since generically $\mathcal{H}$ is not hermitian, these vectors are not conjugates of each other, i.e. in general we have

$$|P_0\rangle \neq (\langle 0|)^\dagger \,. \tag{9.2.4}$$

Thus indeed a stationary state exists.
If in addition some further conditions are satisfied one can even show that the stationary state is unique. Then the system is **ergodic** which is a very useful property since e.g. ensemble averages are the same as time averages in the stationary state. We do not give a full mathematical characterization here. The main requirement is that the configuration space can not be decomposed into different subspaces between which there are no transitions possible.

What can be said about the other eigenvalues of the stochastic Hamiltonian $\mathcal{H}$? The special properties of a stochastic matrix allows to draw some rather general conclusions. These are based on **Gershgorin's theorem** which states that the eigenvalues λ_j of a complex $n \times n$-matrix $A = (a_{ij})$ lie in the union of the disks

$$D_j = \left\{ z \in \mathbb{C} \,:\, |z - a_{jj}| \leq \sum_{l \neq j} |a_{jl}| \right\} \,, \tag{9.2.5}$$

i.e. a circle of radius $\sum_{l\neq j} |a_{jl}|$ around a_{jj} (see Fig. 9.2.1).

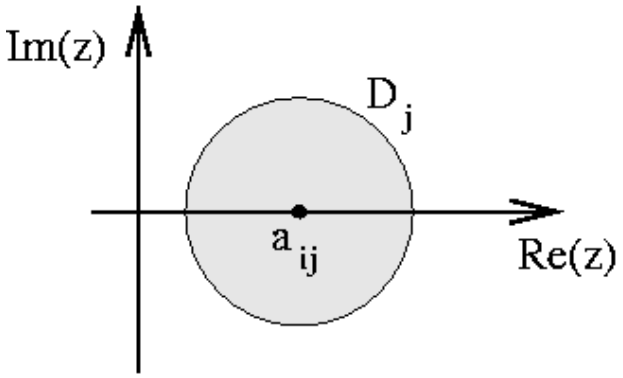

Figure 9.2.1: Illustration of Gershgorin's theorem.

For a stochastic matrix one has $\sum_{l\neq j}|a_{jl}| = -\sum_{l\neq j} a_{jl} = a_{jj}$ since all off-diagonal elements are non-positive and the sum of each column is 0. Thus Gershgorin's theorem tells us[2] that the eigenvalues lie within the disks $D_j = \{z \in \mathbb{C} : |z - a_{jj}| \leq a_{jj}\}$. Especially this means that

$$\mathrm{Re}(\lambda_j) \geq 0\,. \tag{9.2.6}$$

This means that all eigenvalues of a stochastic matrix have a non-negative real part. We will later see that this is has implications for physics since the real parts can be interpreted as relaxation times (see also Sec. 5.2).
Thus we can summarize the general properties of the eigenvalue spectrum of a stochastic matrix: the eigenvalue $E_0 = 0$ exists and all other eigenvalues E_j can be complex but have a non-negative real part, $\mathrm{Re}(\lambda_j) \geq 0$.
In principle one can also show that the elements of the eigenvector to the eigenvalue $E_0 = 0$ are all non-negative so that they can be interpreted as probabilities. This will be done below for the case of discrete time. The present case is then included in the limit of $\Delta t \to 0$.

9.3 Formal solution and averages

In analogy to quantum mechanics one can write down the formal solution of the master equation (9.1.4) as

$$|P(t)\rangle = e^{-\mathcal{H}t}|P(0)\rangle\rangle \tag{9.3.1}$$

where $|P(0)\rangle$ is the initial state of the system at time $t = 0$. Conservation of probability then implies $\langle 0|\exp(-\mathcal{H}t) = \langle 0|$.
Often one is interested in averages

$$\langle\mathcal{A}\rangle(t) = \sum_{\mathbf{n}} \mathcal{A}(\mathbf{n})P(\mathbf{n},t) \tag{9.3.2}$$

[2]Note that A and A^t have the same eigenvalues.

of some observable $\mathcal{A}(\mathbf{n})$ at time t. E.g. $\mathcal{A}(\mathbf{n}) = \mathbf{n}_j$ would give the probability to find a particle at site j if $\mathbf{n}$ is the particle number operator.
Using the left eigenvector $\langle 0|$ we can rewrite this expectation value in the quantum formalims:

$$\langle \mathcal{A} \rangle(t) = \langle 0|\mathcal{A}|P(t)\rangle. \tag{9.3.3}$$

Note that this is different from quantum mechanics where averages are given by $\langle\psi|\mathcal{A}|\psi\rangle$. In the stochastic quantum formalism the elements of the vector $|P(t)\rangle$ are already probabilities whereas in quantum mechanics the *square* of the modulus is interpreted as probability density.
Expanding the initial state $|P(0)\rangle = \sum_\lambda a_\lambda |P_\lambda\rangle$ in terms of the eigenvectors $|P_\lambda\rangle$ with eigenvalues E_λ of $\mathcal{H}$, the expectation values can be rewritten as

$$\langle \mathcal{A} \rangle(t) = \langle 0|\mathcal{A}e^{-\mathcal{H}t}|P(0)\rangle = \sum_\lambda a_\lambda e^{-E_\lambda t}\langle 0|\mathcal{A}|P_\lambda\rangle. \tag{9.3.4}$$

This shows that the behavior for large times t is governed by the low-lying excitations, i.e. the eigenvalues with smallest real parts $\mathrm{Re}(E_\lambda)$. These determine the leading **relaxation times** τ_λ through

$$\tau_\lambda^{-1} = \mathrm{Re}(E_\lambda) \tag{9.3.5}$$

provided that the corresponding matrix element $\langle 0|\mathcal{A}|P_\lambda\rangle$ is non-vanishing.

9.4 Comparison stochastic dynamics vs. quantum mechanics

The following table provides a dictionary for the similarities and differences of stochastic dynamics and quantum mechanics. As example for a quantum mechanical system a quantum spin chain with spin $S = 1/2$ has been used so that the system has two (local) states characterized by $S_j^z = \pm 1/2$. The corresponding stochastic system also has two local states, e.g. the occupation number $n_j = 0, 1$.

	stochastic dynamics	quantum mechanics
local state	$n_j = 0, 1$	$S_j^z = \pm 1/2$
global state	probability $P(n_1, \ldots, n_L; , t)$	wavefunction $\psi(S_1^z, \ldots, S_L^z; t)$
model definition	transition rates $w(\mathbf{n}' \to \mathbf{n})$	energies E_σ (+ matrix elements)
fundamental equation	master equation	Schrödinger equation
state representation	$\lvert P(t)\rangle = \sum_{\mathbf{n}} P(\mathbf{n}, t)\lvert\mathbf{n}\rangle$	vector $\lvert\psi\rangle$ in Hilbert space
Hamiltonian	non-hermitian, stochastic matrix	hermitian
lowest eigenstate	stationary state	ground state
lowest eigenvalue	0	E_0
spectrum	E_α complex, $\mathrm{Re}(E_\alpha) \geq 0$	E_α real
eigenvectors (EV)	right EV $\neq$ left EV	right EV = left EV
averages	$\langle\mathcal{A}\rangle(t) = \langle 0\lvert\mathcal{A}\rvert P(t)\rangle$	$\langle\mathcal{A}\rangle = \langle\psi\lvert\mathcal{A}\rvert\psi\rangle$
excitations	relaxation ${\tau_\lambda}^{-1} \propto \mathrm{Re}(E_\lambda)$	energy gap Δ

9.5 Quantum formalism for discrete time dynamics

The quantum formalism can also be applied for stochastic dynamics with discrete time update (see Sec. 8.1.2). The master equation in discrete time (8.1.5) can be rewritten by introducing a **transfer matrix** $\mathcal{T}$ with elements

$$\mathcal{T}_{\mathbf{n},\mathbf{n}'} := W(\mathbf{n}' \to \mathbf{n}) \tag{9.5.1}$$

instead of a stochastic Hamiltonian $\mathcal{H}$ as in the case of continuous time dynamics. Since the matrix elements of T are probabilities they are non-negative with $\mathcal{T}_{\mathbf{n},\mathbf{n}'} \in [0,1]$. Furthermore conservation of probability implies

$$\sum_{\mathbf{n}'} \mathcal{T}_{\mathbf{n},\mathbf{n}'} = \sum_{\mathbf{n}'} W(\mathbf{n}' \to \mathbf{n}) = 1\,. \tag{9.5.2}$$

In terms of the transfer matrix the master equation takes the form

$$\boxed{|P(t+\Delta t)\rangle = \mathcal{T}|P(t)\rangle\,.} \tag{9.5.3}$$

A formal solution can be obtained by iteration

$$|P(t)\rangle = \mathcal{T}^n|P(0)\rangle \qquad \text{for } t = n\Delta t\,. \tag{9.5.4}$$

The stationary state $|P_0\rangle = \lim_{t\to\infty}|P(t)\rangle$ corresponds to the (right) eigenvector $|P_0\rangle$ of the transfer matrix $\mathcal{T}$ with eigenvalue 1:

$$|P_0\rangle = \mathcal{T}|P_0\rangle\,. \tag{9.5.5}$$

Similar arguments for the existence and uniqueness of the stationary state as in the case of continuous time apply. Now $\langle 0|$ as defined in (9.2.2) is a left eigenvector of $\mathcal{T}$ with eigenvalue 1. The corresponding right eigenvector $\langle P_0|$ is the stationary state. All other eigenvalues of $\mathcal{T}$ have real parts ≥ 1.

Such statements can be proved using the **Perron-Frobenius theorem**. If $A = (a_{ij})$ is a real $N \times N$ matrix with strictly positive entries $a_{ij} > 0$ then there is a real number $\lambda_0 > 0$ which is an eigenvalue of A and any other (possibly complex) eigenvalue λ is strictly smaller than, i.e. $|\lambda| < \lambda_0$. As a consequence the spectral radius of A is equal to λ_0. In compact form the Perron-Frobenius theorem states that a real square matrix with positive entries has a unique largest real eigenvalue with a corresponding eigenvector which has strictly positive components. The theorem can be extended to the case $a_{ij} \geq 0$, but then additional conditions have to be satisfied, i.e. the matrix has to be irreducible.

Chapter 10

Local processes

So far we have made no restrictions in the transition rates $w(\mathbf{n} \to \mathbf{n}')$. In principle transitions between arbitrary states $\mathbf{n}$ and $\mathbf{n}'$ were allowed. In many applications this is not the case, e.g. for interacting and non-interacting random walks where particles are allowed to move to neighbouring sites only. In the following we will study this important case in more detail.

10.1 Processes with nearest-neighbour interactions

In the following we consider a simple but important class of stochastic processes where particles only interact with those on nearest neighbour sites. In addition, the so-called **exclusion principle** is applied, i.e. each site j can be occupied by at most one particle. Therefore the state of each site can be specified by the occupation number $n_j = 0, 1$. Such processes are usually called **2-state models** or **single-species models** [1]

For processes with nearest neighbour interactions most transition rates are actually zero. Only transition rates between configurations $\mathbf{n}$ and $\mathbf{n}'$ that are different only at neighbouring sites j and $j+1$ are non-vanishing, i.e. only those of the form

$$\boxed{w\big((n_1, \dots n_j, n_{j+1}, \dots, n_L) \to (n_1, \dots n'_j, n'_{j+1}, \dots, n_L)\big)\,.} \tag{10.1.1}$$

In the case of processes with nearest-neighbour interactions only, the stochastic Hamiltonian can be written in the form

$$\mathcal{H} = \sum_{j=1}^{L} h_{j,j+1}\,, \tag{10.1.2}$$

i.e. as a sum of local Hamiltonians $h_{j,j+1}$ with nearest-neighbour interactions that act non-trivially only on sites j and $j+1$. Here we have assumed periodic boundary conditions. Generalization to other cases is straightforward. Later we will often consider systems coupled to

[1] The terminology '2-state model' is used in two different ways in the literature. It can either refer to the number of states of a *site* or the number of states of the particle (if it has an internal degree of freedom).

interpretation	process	rate
diffusion to left	$01 \to 10$	Γ^{01}_{10}
diffusion to right	$10 \to 01$	Γ^{10}_{01}
pair annihilation	$11 \to 00$	Γ^{11}_{00}
pair creation	$00 \to 11$	Γ^{00}_{11}
death on left	$10 \to 00$	Γ^{10}_{00}
death on right	$01 \to 00$	Γ^{01}_{00}
birth on left	$00 \to 10$	Γ^{00}_{10}
birth on right	$00 \to 01$	Γ^{00}_{01}
fusion on left	$11 \to 10$	Γ^{11}_{10}
fusion on right	$11 \to 01$	Γ^{11}_{01}
branching to left	$01 \to 11$	Γ^{01}_{11}
branching to right	$10 \to 11$	Γ^{10}_{11}

Table 10.1: Definition and interpretation of rates allowed in a 2-state stochastic process with nearest neighbour interactions. Branching and fusion are sometimes called **coagulation** (or **coalescence**) and **decoagulation**. The diagonal elements $\Gamma^{\tau\tau'}_{\tau\tau'}$ are determined by the condition $\sum_{\sigma,\sigma'} \Gamma^{\tau\tau'}_{\sigma\sigma'} = 0$.

boundary reservoirs. In this case the stochastic Hamiltonian has the form

$$\mathcal{H} = h_1 + \sum_{j=1}^{L-1} h_{j,j+1} + h_L \tag{10.1.3}$$

where h_1 and h_L act only on the first and last site, respectively, and describe the coupling to the corresponding particle reservoirs.

10.2 Reaction-diffusion processes

Next we classify the different local processes that are allowed in the case of nearest-neighbour interactions. This leads to the 12 non-trivial processes shown in the following table: This type of model is also called **reaction-diffusion process**. In terms of the local transition rates $\Gamma^{s'_j s'_{j+1}}_{s_j s_{j+1}}$ for a local transition of sites j and $j+1$ from $(s'_j s'_{j+1})$ to $(s_j s_{j+1})$ the local Hamiltonian in the basis $(|00\rangle, |01\rangle, |10\rangle, |11\rangle)$ is given by

$$h_{j,j+1} = \begin{pmatrix} \Gamma^{00}_{01} + \Gamma^{00}_{10} + \Gamma^{00}_{11} & -\Gamma^{01}_{00} & -\Gamma^{10}_{00} & -\Gamma^{11}_{00} \\ -\Gamma^{00}_{01} & \Gamma^{01}_{00} + \Gamma^{01}_{10} + \Gamma^{01}_{11} & -\Gamma^{10}_{01} & -\Gamma^{11}_{01} \\ -\Gamma^{00}_{10} & -\Gamma^{01}_{10} & \Gamma^{10}_{00} + \Gamma^{10}_{01} + \Gamma^{10}_{11} & -\Gamma^{11}_{10} \\ -\Gamma^{00}_{11} & -\Gamma^{01}_{11} & -\Gamma^{10}_{11} & \Gamma^{11}_{00} + \Gamma^{11}_{01} + \Gamma^{11}_{10} \end{pmatrix}. \tag{10.2.1}$$

A **particle-hopping model** is a special type of reaction-diffusion processes that conserves the number of particles, at least in the bulk[2]. Then only those $\Gamma_{s_j s_{j+1}}^{s'_j s'_{j+1}}$ with $s_j + s_{j+1} = s'_j + s'_{j+1}$ are non-zero, i.e. Γ_{10}^{01}, Γ_{01}^{10}. Our main example for a 2-state particle hopping will be the **Totally Asymmetric Simple Exclusion Process (TASEP)** which has $\Gamma_{10}^{01} = 0$, $\Gamma_{01}^{10} = p$.

10.3 Matrix representation of stochastic Hamiltonians

The stochastic Hamiltonian $\mathcal{H}$ in (10.1.2) is the sum of operators $h_{j,j+1}$ acting on sites j and $j+1$ only. The basis vectors of the state space are $|n_1, \ldots, n_L\rangle$ which mathematically are given by a tensor product (see Sec. 10.3.1) of the basis vectors on each site:

$$|n_1, \ldots, n_L\rangle = |n_1\rangle \otimes |n_2\rangle \otimes \cdots \otimes |n_L\rangle \,. \tag{10.3.1}$$

The local Hamiltonian $h_{j,j+1}$ then acts in the following way:

$$h_{j,j+1}|n_1, \ldots, n_L\rangle = |n_1\rangle \otimes \cdots |n_{j-1}\rangle \otimes \left(\hat{h}_{j,j+1}|n_j, n_{j+1}\rangle\right) \otimes |n_{j+2}\rangle \otimes \cdots |n_L\rangle \,, \tag{10.3.2}$$

where $\hat{h}_{j,j+1}$ is an operator that "lives" in the product space of sites j and $j+1$.
In a matrix representation the state space of a site j is two-dimensional and thus the state space of a system of L sites would be 2^L-dimensional. The stochastic Hamiltonian is then represented by a $2^L \times 2^L$-dimensional matrix whereas $\hat{h}_{j,j+1}$ corresponds to a 4×4-matrix[3].

10.3.1 Definition of tensor product

We give a brief remainder the description of multilinear mappings by tensor products of matrices.
The **tensor product** of two vector spaces V and W with basis vectors $\{v_1, \ldots, v_N\}$ and $\{w_1, \ldots, w_M\}$, respectively, is given by the cartesian product

$$V \otimes W := \{(v_i, w_j) | i = 1, \ldots, N;\ j = 1, \ldots, M\} \tag{10.3.3}$$

The tensor product of a $n_1 \times n_2$ matrix $A = (a_{ij})$ and a $m_1 \times m_2$ matrix B is defined as the following $(n_1 m_1) \times (n_2 m_2)$ matrix:

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n_2}B \\ a_{21}B & a_{22}B & \cdots & a_{2n_2}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{n_1 1}B & a_{n_1 2}B & \cdots & a_{n_1 n_2}B \end{pmatrix} . \tag{10.3.4}$$

Here $a_{ij}B$ stands for the submatrix obtained by multiplying all elements of B by a_{ij}.

[2] **Bulk** refers to the part of a system which is (far) away from the boundaries.

[3] Note that $h_{j,j+1}$ would also correspond to a $2^L \times 2^L$-matrix since it acts in the full state space. However, most of its elements are trivial.

Chapter 11

Solution methods

In the following we again consider for simplicity a 2-state model in one dimension with nearest-neighbour interactions. We will be interested in systematic methods to solve the master equation for such systems. However, as with many-body quantum systems, exact solutions are possible only in a few cases, e.g. for special parameter values of models. If one is interested in analytical results, in such cases one has to rely on approximative methods.

11.1 Exact methods

11.1.1 Bethe Ansatz

The Bethe Ansatz is probably the most successful method for the exact solution of interacting many particle systems. It has been developed by Hans Bethe in 1931 in his celebrated solution of the one-dimensional isotropic Heisenberg model of interacting quantum spins. The method is closely related to the concept of **integrability**, i.e. the existence of an infinite number of (mutually commuting) conservation laws, and the occurrence of non-diffractive scattering. With hindsight to the quantum formalism described in detail in Sec. 9 it is natural to try application of the method also to stochastic systems.
As the name already indicates, a special form of the probability vector $P(n_1, \ldots, n_L)$ is assumed as an "Ansatz". In fact it is more convenient to use the position x_j of occupied sites (i.e. those with $n_l = 1$) as parameters. For a state with N occupied sites $x_1, \ldots, x_N$ (i.e. $n_{x_j} = 1$) the typical structure of the **Bethe Ansatz** is

$$P(x_1, \ldots, x_N) = \sum_{\{\sigma_1,\ldots,\sigma_N\}} A_{\{\sigma_1,\ldots,\sigma_N\}} \prod_{j=1}^{N} z_{\sigma_j}^{-x_j} \tag{11.1.1}$$

The sum is over all permutations σ of the numbers $(1, \ldots, N)$ and the $A_{\{\sigma_1,\ldots,\sigma_N\}}$ are unknown amplitudes. The structure of the Bethe Ansatz is similar to a Slater determinant which corresponds to the case $A_{\{\sigma_1,\ldots,\sigma_N\}} = \mathrm{sign}(\sigma)$.

The parameters z_j are also unknown. Sometimes they are expressed by (quasi-)momenta k_j such that $z_{\sigma_j}^{-x_j} = e^{ik_{\sigma_j}x_j}$. They have to satisfy a system of N coupled nonlinear equations which arise as consistency conditions when applying periodic (or other) boundary conditions.
The applicability of the Bethe Ansatz imposes serious restrictions on the form of the interactions and therefore only few models can be treated by this method. Especially the Ansatz works only in one dimension. Its application to stochastic systems has been reviewed e.g. in [8].

11.1.2 Matrix-product Ansatz

A very powerful method for the determination of stationary solutions of the master equation is the so-called **matrix-product Ansatz (MPA)**. A rather extensive review can be found in [9]. For a system with open boundaries the probabilities $P(\mathbf{n})$ in the stationary state can be written in the form

$$\boxed{P(n_1,\ldots,n_L) = \frac{1}{Z_L}\langle W| \prod_{j=1}^{L} [n_j D + (1-n_j)E] \, |V\rangle \,.} \tag{11.1.2}$$

For periodic boundary conditions the MPA takes the translational invariant form

$$P(n_1,\ldots,n_L) = \frac{1}{Z_L} \operatorname{Tr} \left(\prod_{j=1}^{L} [n_j D + (1-n_j)E] \right). \tag{11.1.3}$$

Z_L is a normalization constant that can be calculated as

$$Z_L = \langle W|C^L|V\rangle, \qquad \text{with} \qquad C = D + E \,. \tag{11.1.4}$$

In (11.1.2), (11.1.3) E and D are matrices and $\langle W|$ and $|V\rangle$ are vectors characterizing the boundary conditions.
The Ansatz (11.1.2) has a simple interpretation. It implies the translation of a configuration of the system into a product of matrices. Starting with site 1, for each empty site ($n_j = 0$) a factor E appears and for each occupied site ($n_j = 1$) a factor D. In this way, one would have e.g.

$$P(1,1,0,0,1,0) \sim D \cdot D \cdot E \cdot E \cdot D \cdot E \cdot = D^2E^2DE \,. \tag{11.1.5}$$

The boundary conditions are taken into account by multiplication with suitable $\langle W|$ and $|V\rangle$, i.e. a combination of matrix elements of this product.
The explicit form of the matrices E and D and the boundary vectors has to be determined from the condition that (11.1.2) or (11.1.3) solves the master equation (8.1.1). This leads in general to algebraic relations between the matrices E and D and the boundary vectors $\langle W|$ and $|V\rangle$. We will later consider this in more detail for some important cases.
One can show that the stationary states of one-dimensional stochastic processes are generically of matrix-product form [7]. However, so far an explicit determination of the matrices from the algebra has been possible in only a few cases.

11.2 Approximate methods

11.2.1 Mean-field approximation

Mean-field theories (MFT) neglect all correlations between state variables. Formally, the **mean-field approximation (MFA)** corresponds to factorizing the probability $P(\mathbf{n})$ into single-site contributions $P_1(n_j)$:

$$\boxed{P(n_1, \ldots, n_L) \approx P_1(n_1)P_1(n_2)\cdots P_1(n_L)} \tag{11.2.1}$$

Inserting this Ansatz into the master equation[1] simplifies the problem considerably and leads to an equation which only involves the function P_1. For translationally invariant systems P_1 is independent of the site j. It amounts in approximately expressing the transition rates $w(\mathbf{n} \to \mathbf{n}')$ through the single-site contributions $P_1(n_j)$.
Alternatively one could write down the *exact* master equation for P_1. However, this requires to take into account correlations with neighbouring sites, i.e. quantities like $P_2(n_j, n_{j+1})$ or $P_3(n_{j-1}, n_j, n_{j+1})$ in the case of nearest-neighbour interactions. These larger P_n have then to be expressed approximately as products of P_1 similar to (11.2.1).
Sometimes quantities of interest can be calculated in MFA without requiring formal calculations. An example is the current in the TASEP introduced in Sec. 10.2 which is a model of interacting random walks with unidirectional particle motion. In the stationary state the current has to be independent of the position because the continuity equation

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \vec{J} = 0 \tag{11.2.2}$$

reduces to $dJ/dx = 0$. Formally the current between sites j and $j+1$ is given by

$$J_j = p\langle n_j(1 - n_{j+1})\rangle \tag{11.2.3}$$

because a configuration $\mathbf{n}$ only contributes to the current between these sites if j is occupied and $j+1$ empty. If this is the case a current will flow, i.e. a particle moves from j to $j+1$, with probability p per time.
Since the MFA neglects correlations, the current is approximately given by

$$J = J_j \approx p\langle n_j\rangle\langle 1 - n_{j+1}\rangle = p\rho(1-\rho) \tag{11.2.4}$$

where $\rho = N/L$ is the particle density (i.e. the probability that the site is occupied) if there are N particles on a lattice of L sites.

Note that often a configuration can be characterized by different variables. The main example in this course are particle-hopping models. Instead of occupation numbers $n_j = 0, 1$ or particle

[1] In most applications, MFA is used for the steady-state properties. But in principle also the full dynamics could be studied with time-dependent P_1.

positions x_j, any (translational-invariant) state can equivalently be characterized by the interparticle distances (gaps) d_j, also called headway in traffic engineering. The mean-field theories for n_j and d_j will usually lead to different results. In order to distinguish these different approximations one sometimes speaks of **site-oriented (SOMF)** and **particle-oriented mean-field approach (POMF)**. Since the latter is used quite frequently in the context of traffic modelling, the terminology **car-oriented MFT (COMF)** is more standard and will be used in the following.
There is a large class of models for which MFT even gives the exact result. These models are said to have a **product-measure** or **factorized steady-state**. One example is the TASEP with continuous time dynamics, i.e. for this model the current (11.2.4) is indeed exact. However, this is no longer true if we apply discrete time dynamics!

11.2.2 Cluster approximation

MFT can be systematically extended to take into account short-range correlations. This leads to the **cluster approximation** which is also known under the name **local structure theory**, n**-step Markovian approximation** and goes back to the **probability path method**.
The n-cluster approximation treats a cluster of n neighbouring sites exactly. The 1-cluster approximation is identical to the (site-oriented) mean-field approach, see (11.2.1). The 2-cluster approximation corresponds to the factorization

$$P(n_1, \ldots, n_L) \propto P_2(n_1, n_2) P_2(n_2, n_3) \cdots P_2(n_{L-1}, n_L)\,. \tag{11.2.5}$$

For periodic boundary conditions an additional factor $P_2(n_L, n_1)$ appears.
For larger clusters one has the freedom to choose different overlaps of neighbouring cluster. For instance there are two different 3-cluster approximations

$$P(n_1, \ldots, n_L) \propto P_3(n_1, n_2, n_3) P_3(n_2, n_3, n_4) \cdots \tag{11.2.6}$$

called $(3, 2)$-cluster approximation, and

$$P(n_1, \ldots, n_L) \propto P_3(n_1, n_2, n_3) P_3(n_3, n_4, n_5) \cdots \tag{11.2.7}$$

called $(3, 1)$-cluster approximation. In general there are $n - 1$ different n-cluster approximations (n, m), where m is the overlap between neighbouring clusters. For fixed size n, the quality of the approximation typically increases with the overlap m. Fig. 11.2.1 shows a graphical representation of the cluster decomposition.
The probabilities corresponding to clusters of different sizes are not independent, but related through the **Kolmogorov consistency conditions**:

$$\begin{aligned} P_{n-1}(n_1, \ldots, n_{n-1}) &= \sum_{\tau} P_n(\tau, n_1, \ldots, n_{n-1}) \\ &= \sum_{\tau} P_n(n_1, \ldots, n_{n-1}, \tau)\,, \end{aligned} \tag{11.2.8}$$

e.g. $\sum_\tau P_2(n_1, \tau) = P_1(n_1)$. These relations give strong conditions especially for the 2-cluster approximation of particle hopping models. There it also provides a relation with physical parameters, e.g. the density ρ in translational invariant systems through $P_1(1) = \rho$ and $P_1(0) = 1 - \rho$.

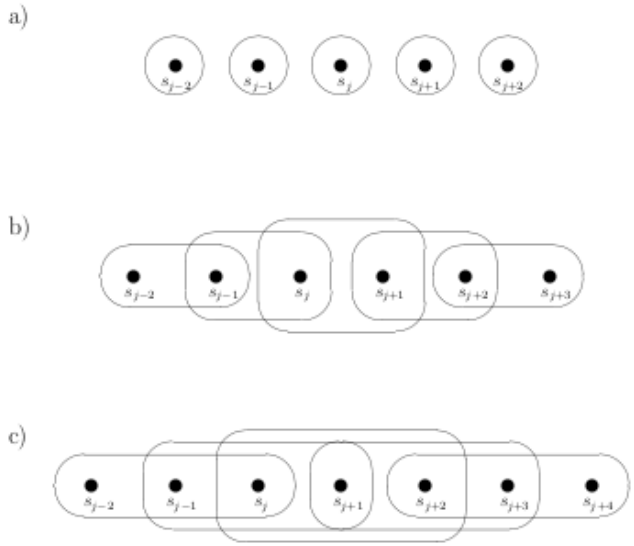

Figure 11.2.1: Graphical representation of the $n-$cluster approximation for a) $n = 1$ (i.e. mean-field theory), b) $n = 2$, and c) $n = 3$ (more specifically the $(3, 2)$-cluster approximation). Each circle corresponds to one factor in the factorization.

Chapter 12

Zero-Range Process

We now consider a particle hopping model which does not satisfy the exclusion principle. Here an arbitrary number of particles is allowed on each site. However it will turn later out that this model contains particle hopping models with exclusion, like the TASEP, as special cases.

12.1 Definition of the Zero-Range Process

The ZRP is characterized by its ultra-local transition rules. Originally it was introduced in mathematics by Spitzer in 1970. Meanwhile many general and rigorous results are known for it coming from mathematics. In physics, the importance and relevance of the ZRP was only realized widely almost 30 years later.

As in the TASEP, the zero-range process describes the dynamics of N (classical and indistinguishable) particles on a lattice of L sites. We will consider one-dimensional lattice with periodic boundary conditions in the following, but in principle many results can be generalized to any lattice structure. In contrast to the TASEP, particles in the zero-range process are not subject to any exclusion rule. Each lattice site j can be occupied by an arbitrary number n_j of particles. Using random-sequential dynamics, in each update step a particle is chosen at random, say on site j, which is allowed to move to any other site l. The essential point is now that the corresponding hopping rate w_{jl} is *independent* of the state of the target site l, i.e. the number of particles present at that site. It only depends on the number of particles n_j on the starting site j. Therefore the rate can conveniently be denoted by $u(n_j)$.

This property motivates the name **zero-range process (ZRP)** since the range of interaction can be interpreted to be zero. Obviously this only works if the number of particles allowed at each site is not restricted. Any restriction would imply that the hopping rate can not be independent of the state of the target site since one has to check whether there are less particles than the maximally allowed number. If this number is reached, the hopping rates from any other site to this site vanish.

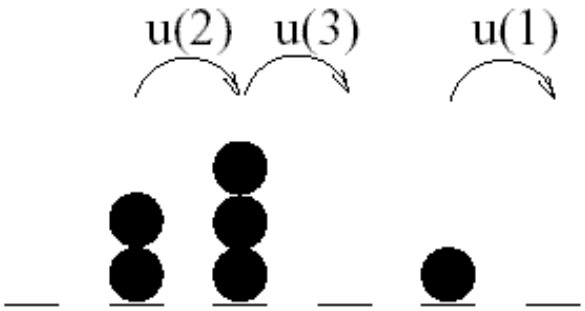

Figure 12.1.1: Dynamics of the zero-range process: Each site of the lattice can be occupied by an arbitrary number of particles. The departure rate $u(n)$ of a particle depends only on the number n of particles at the departure site, not on that at the target site.

12.2 Solution of the Zero-Range Process

The ZRP has the important property of being exactly solvable. In fact its steady state is given by a simple factorised form, where the probability $P(\mathbf{n})$ of finding the system in a configuration $\mathbf{n} = (n_1, n_2, \ldots, n_L)$ is given by a product of (scalar) factors $f(n_l)$:

$$P(\mathbf{n}) = \frac{1}{Z_{L,N}} \prod_{l=1}^{L} f(n_l) \; . \tag{12.2.1}$$

The normalisation $Z_{L,N}$, which is sometimes called *(canonical) partition function*, ensures that the probabilities for all configurations containing N particles sum up to one:

$$Z_{L,N} = \sum_{\mathbf{n}} \prod_{l=1}^{L} f(n_l) \, \delta\left(\sum_{l=1}^{L} n_l - N \right) \; , \tag{12.2.2}$$

where the Kronecker-δ guarantees that only configurations containing N particles are included. In order to prove the factorisation property (12.2.1) and determine the factors $f(n_l)$ we insert this as an Ansatz into the stationary master equation:

$$\sum_{l=1}^{L} u(n_{l-1}+1) P(\cdots, n_{l-1}+1, n_l-1, \cdots) = \sum_{l=1}^{L} u(n_l) P(\mathbf{n}) \; . \tag{12.2.3}$$

In the steady state the probability current due to hops into a particular configuration $\mathbf{n}$ (lhs of (12.2.3)) balances the probability current due to hops out of the same configuration (rhs of (12.2.3)). The choice $u(0) = 0$ takes into account that a site has to be occupied before a particle can move away from it[1]. Now we substitute the factorised probability (12.2.1) into the master

[1] It is also assumed that $P(n_1, \ldots, n_L) = 0$ if at least one of the arguments is negative.

equation (12.2.3) and equate each term in the sum separately. After cancelling common factors this leads to the conditions

$$u(n_{l-1}+1)f(n_{l-1}+1)f(n_l-1) = u(n_l)f(n_{l-1})f(n_l)\ , \tag{12.2.4}$$

which can be rewritten in the form

$$u(n_{l-1}+1)\frac{f(n_{l-1}+1)}{f(n_{l-1})} = u(n_l)\frac{f(n_l)}{f(n_l-1)} \stackrel{!}{=} \text{constant}\ , \tag{12.2.5}$$

for all values of l. The equality of the first two terms implies that they are constant since the first term depends only on n_{l-1} whereas the second one depends only on n_l. The constant can be set equal to unity without loss of generality (by rescaling time and thus the rates) so that we obtain the iteration

$$f(n_l) = \frac{f(n_l-1)}{u(n_l)}\ . \tag{12.2.6}$$

Setting $f(0) = 1$, again without loss of generality, this iteration is solved by

$$f(n) = \prod_{j=1}^{n}\frac{1}{u(j)} \qquad \text{for} \quad n > 0\ . \tag{12.2.7}$$

Thus we have shown that the factorisation (12.2.1) indeed holds if we choose the factors according to (12.2.7). This is also true for other types of update.
Alternatively one could also start with prescribing the stationary state by defining the function $f(n)$ and then determine the corresponding hopping rates through

$$u(n) = \frac{f(n-1)}{f(n)}\ . \tag{12.2.8}$$

This shows that the ZRP gives a complete account of all the possible steady state behavior in particle-conserving models with factorised steady state.
The solution for the case of discrete time dynamics can be obtained in an analogous way. The stationary probabilities factorize as in (12.2.1) with the weight functions

$$f(n) = \begin{cases} 1-u(1) & (n=0) \\ \frac{1-u(1)}{1-u(n)}\prod_{j=1}^{n}\frac{1-u(j)}{u(j)} & (n \geq 1)\,. \end{cases} \tag{12.2.9}$$

Knowing the explicit form (12.2.1), (12.2.7) of the steady state probabilities one can calculate all stationary properties of the ZRP, at least in principle.

One important quantity that helps to understand the structure of the stationary state is the probability $p(n)$ that a given site contains n particles. Fixing the occupation of site $j = 1$ to be n one has

$$p(n) = \sum_{n_2,\ldots,n_L} P(n, n_2 \ldots n_L)\,\delta\left(\sum_{l=2}^{L} n_l - (N-n)\right)\ , \tag{12.2.10}$$

where the Kronecker-δ takes into account that the remaining sites contain only $N - n$ particles. Using the factorization (12.2.1) one finally obtains

$$p(n) = f(n)\frac{Z_{L-1,N-n}}{Z_{L,N}} \ . \tag{12.2.11}$$

Note that this probability $p(n)$ to find n particles at a given site is different from the single-site weight $f(n)$. This is a consequence of the fact that the constraint of fixed particle number induces correlations between sites, although the steady state factorizes.

Chapter 13

TASEP I: periodic boundary conditions

The **Totally Asymmetric Simple Exclusion Process (TASEP)** introduced in Sec. 10.2 is a particle-hopping model (see Fig. 13.0.1) which serves as paradigm not only for many transport or traffic models, but is also of great importance for the general theory of nonequilibrium processes. Here it is of similar relevance as the Ising model for equilibrium statistical mechanics and phase transitions. One of its attractive properties is the fact that many exact results can be obtained. First we consider the simpler case of periodic boundary conditions which leads to a translational-invariant stationary state.

13.1 Mapping to ZRP

Let us first consider the TASEP with periodic boundary conditions and random-sequential dynamics. In the following, L and N refer to the total number of sites and particles, respectively. The state of each site i is characterized by the occupation number n_i such that $n_i = 0$ if the site is empty and $n_i = 1$ if it is occupied by a particle.
In the following we will determine the stationary state of the TASEP exactly. This is achieved by mapping it onto a zero-range process for which we have already derived the steady-state distribution in Chapter 12.2.
In this mapping the particles of the TASEP become the sites in the ZRP. The particles in the ZRP picture are then given by empty sites ('holes') in the TASEP picture: the empty sites between particles j and $j + 1$ in the TASEP become the particles on site j in the ZRP (see Fig. 13.1.1).

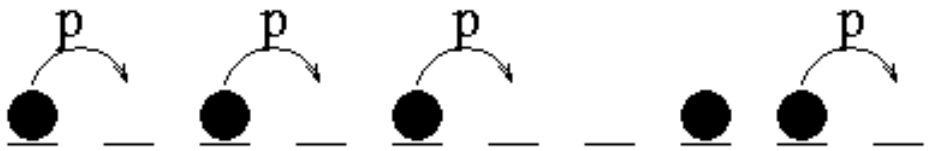

Figure 13.0.1: Dynamics of the (totally) asymmetric simple exclusion process: particles move to empty right neighbour sites with rate p.

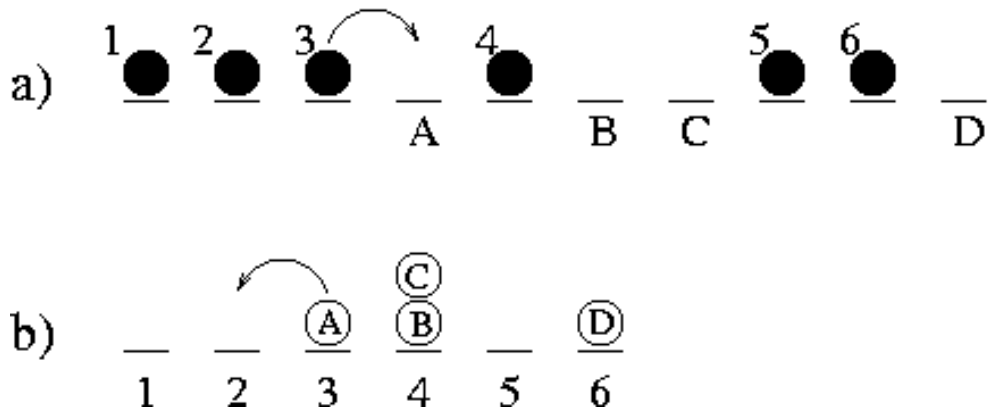

Figure 13.1.1: Mapping of the TASEP onto the ZRP for $L = 10$ and $N = 6$. Particle in the TASEP correspond to sites in the ZRP (indicated by numbers). Empty sites in the TASEP become particles in the ZRP (indicated by letters). The arrows show a possible local transition with probability $u(1)$.

Thus we introduce a new dynamical variable d_j which is just the number of (consecutive) empty sites in front of particle j. In traffic engineering this quantity is often called **headway**. Obviously it satisfies the conservation law

$$\sum_{j=1}^{N} d_j = L - N \,. \tag{13.1.1}$$

Using this mapping the TASEP with L sites and N particles is equivalent to a ZRP with N sites and $L - N$ particles. Note that the particles in the ZRP hop to the left when the particles in the TASEP move to the right.
This mapping holds for time-continuous dynamics and for parallel as well. The hopping rates in the ZRP are given by

$$u(d_j) = p \qquad \text{for } d_j \geq 1 \,. \tag{13.1.2}$$

Using the exact results for the ZRP we know that the stationary probability distribution is of product form (see (12.2.1)) with weight functions

$$f(n) = \frac{1}{p^n} \tag{13.1.3}$$

for continuous time (i.e. random-sequential update) and

$$f(n) = \begin{cases} 1-p & (n=0) \\ \left(\frac{1-p}{p}\right)^n & (n \geq 1) \end{cases} \tag{13.1.4}$$

for discrete time dynamics (i.e. parallel update).
By applying the approximate methods introduced in Sec. 11.2 to the TASEP with periodic boundary conditions one finds some interesting results. It turns out that for the case of continuous time

dynamics already the simple mean-field approximation gives the exact result. Note that this MFA applies to the occupation numbers as dynamical variables whereas by mapping to the ZRP the headways become the relevant variables. So in this case there are no correlations between different sites. In contrast, for discrete time dynamics the MFA is not exact. Instead it turns out that the 2-cluster approximation gives the exact steady state. That means that for the discrete time case there are correlations between neighbouring sites. This shows that the choice of the update procedure (discrete vs. continuous time or random-sequential vs. parallel update) can have a strong influence on the properties of the stationary state.

13.2 Fundamental diagram

Using the results of the previous section we now can calculate the expectation values of various variables.
The simplest expectation value is the local density $\rho_j = \langle n_j \rangle$. Since the number of particles N is conserved and the stationary state shows translational invariance we must have

$$\rho_j = \langle n_j \rangle = \frac{N}{L} =: \rho \tag{13.2.1}$$

where ρ is usually called **global density**.
Another interesting quantity is the so-called **fundamental diagram** which is probably the most important observable in traffic engineering. The fundamental diagram gives the density dependence of the flow, i.e. the function $J(\rho)$. The **current** (or **flow**) J is related to the average velocity $\bar{v}$ of all particles by the **hydrodynamic relation**

$$J = \rho \bar{v} \,. \tag{13.2.2}$$

For the TASEP we have already seen (eq. (11.2.3)) that the current is determined by

$$J = p\langle n_j(1 - n_{j+1}) \rangle \tag{13.2.3}$$

in terms of occupation numbers. Alternatively we can also express it by the headway,

$$J = p\rho \sum_{d \geq 1} \langle P(d) \rangle = p\rho \left(1 - P(d = 0)\right) \tag{13.2.4}$$

where $P(d)$ is the stationary probability distribution of the headways d_j. Eq. (13.2.4) is easy to understand since a particles can only move if its headway is larger than 0 which is satisfied with probability $1 - P(d = 0)$ in such a case it will move with rate p. Thus $p(1 - P(d = 0))$ is the average velocity of a particle and the hydrodynamic relation gives (13.2.4).
In the case of continuous time dynamics (random-sequential update) we have already seen that the exact fundamental diagram is given by

$$\boxed{J(\rho, p) = p\rho(1 - \rho)} \tag{13.2.5}$$

since MFA is exact in this case. For discrete time dynamics (parallel update) it can be shown that the exact fundamental diagram is of the form

$$\boxed{J(\rho,p)=\frac{1}{2}\left[1-\sqrt{1-4p\rho(1-\rho)}\right]} \tag{13.2.6}$$

Fig. 13.2.1 shows a comparison of this exact result with that of the mean-field approximation which gives the exact result for continuous time. One sees that the MFA strongly underestimates the correct flow for the parallel udpate (Fig. 13.2.1). This shows the importance of correlations which can be described as **particle-hole attraction** (see Excercise 14): in the case of parallel dynamics the probability to find an empty site in front of a particle is increased compared to the completely random case, i.e. $P(1,0) > P(1)P(0)$.

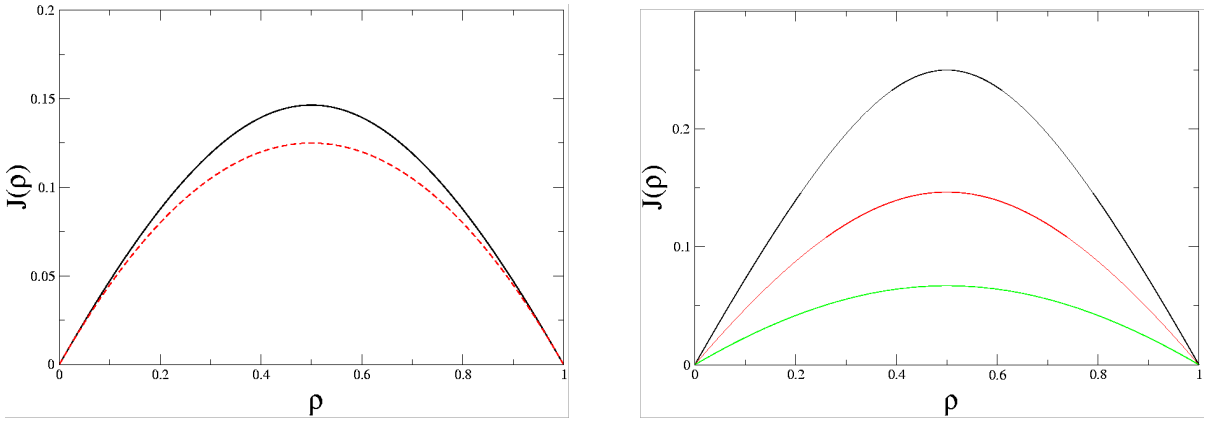

Figure 13.2.1: Flow-density relation for the TASEP with parallel dynamics. Left: Comparison of the exact result and mean-field theory (broken line) for $p = 0.5$. The mean-field result is exact for the case of random-sequential dynamics. Right: Comparison of exact results for $p = 0.25$, $p = 0.5$ and $p = 0.75$ (from bottom to top).

Another important observation is the **particle-hole symmetry** of the fundamental diagram in both cases. It is symmetric with respect to the density $\rho = 1/2$, i.e.

$$J(\rho) = J(1-\rho)\,. \tag{13.2.7}$$

The current does not change if all particles are replaced by empty cells ('holes') and vice versa (**particle-hole transformation**). This is of course expected from the definition of the dynamics since every time a particle moves to the right, a hole moves to the left.

Chapter 14

TASEP II: boundary-induced phase transitions

We now study the TASEP with open boundary conditions by assuming that the first and the last site are coupled to particle reservoirs (Fig. 14.0.1). If the first site is empty ($n_1(t) = 0$) a particle is inserted with an input rate α. If the last site is occupied ($n_L(t) = 1$) the particle is removed with rate β. Due to this change of boundary conditions the system is no longer translational invariant. In addition, the total particle number is not conserved. This makes the theoretical treatment much more difficult. However, we will see that the exact solution for the stationary state can still be obtained for values of the system parameters α, β and p. In the following we will consider the case $p = 1$. As we have seen in Excercise 13 this is no restriction if we are interested in the stationary state since p can be changed by rescaling of time.

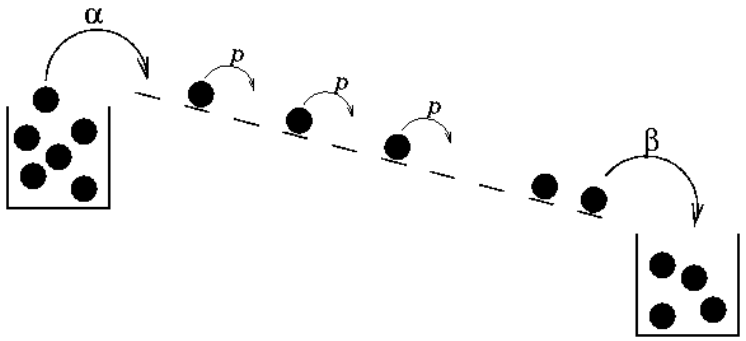

Figure 14.0.1: Definition of the TASEP with open boundary conditions.

14.1 Solution by matrix-product Ansatz

For the solution of the TASEP with open boundary conditions we apply the matrix-product Ansatz

$$P(n_1,\ldots,n_L) = \frac{1}{Z_L}\langle W|X_{n_1}\cdots X_{n_L}|V\rangle \tag{14.1.1}$$

where for an empty site the matrix $X_0 = E$ is used and for an occupied site the matrix $X_1 = D$. Z_L is a normalization factor which can be expressed as

$$\begin{aligned} 1 &= \sum_{\{n_j\}} P(n_1,\ldots,n_L) = \frac{1}{Z_L}\langle W|\Big[\sum_{\{n_j\}} X_{n_1}\cdots X_{n_L}\Big]|V\rangle \\ &= \frac{1}{Z_L}\langle W|\,(X_0+X_1)^L\,|V\rangle = \frac{1}{Z_L}\langle W|C^L|V\rangle \end{aligned} \tag{14.1.2}$$

where we have used the multinomial expansion and introduced the matrix

$$C = D + E\,. \tag{14.1.3}$$

In a similar way one can express the **density profile**, i.e. the average occupation of site j, by the matrices:

$$\rho_j = \langle n_j\rangle = \frac{1}{Z_L}\langle W|C^{j-1}DC^{L-j}|V\rangle \tag{14.1.4}$$

where one has to take into account that only configuration with site j occupied contribute to the sum. Therefore the C at site j has to be replaced by the factor D. Finally the current between site j and $j+1$ is given by

$$J_{j,j+1} = \langle n_j(1-n_{j+1})\rangle = \frac{1}{Z_L}\langle W|C^{j-1}DEC^{L-j-1}|V\rangle \tag{14.1.5}$$

since hopping between these sites is only possible if site j is occupied (factor D) and site $j+1$ is empty (factor E). In the stationary state the equation of continuity $\dot\rho + J' = 0$ tells us that $J' = 0$ so that the current is the same everywhere in the system, i.e. independent of j.
Inserting the Ansatz (14.1.1) into the master equation yields conditions on the so far unknown matrices D, E and the vectors $\langle W|$, $|V\rangle$:

$$\boxed{\begin{aligned} DE &= D+E\,, \\ \alpha\langle W|E &= \langle W|\,, \\ \beta D|V\rangle &= |V\rangle\,. \end{aligned}} \tag{14.1.6}$$

One can easily convince oneself that these are necessary conditions: The relation $DE \propto C = E + D$ is the simplest way to achieve a constant current

$$\begin{aligned} J = J_{j,j+1} &= \frac{1}{Z_L}\langle W|C^{j-1}DEC^{L-j-1}|V\rangle = \frac{1}{Z_L}\langle W|C^{j-1}(D+E)C^{L-j-1}|V\rangle \\ &= \frac{1}{Z_L}\langle W|C^{L-1}|V\rangle = \frac{Z_{L-1}}{Z_L} \end{aligned} \tag{14.1.7}$$

as required by the equation of continuity. Similar arguments apply to the boundary currents:

$$J_l = \alpha(1-\rho_1) = \frac{\alpha}{Z_L}\langle W|EC^{L-1}|V\rangle = \frac{1}{Z_L}\langle W|C^{L-1}|V\rangle = \frac{Z_{L-1}}{Z_L} = J\,, \quad (14.1.8)$$

$$J_r = \beta\rho_L = \frac{\beta}{Z_L}\langle W|C^{L-1}D|V\rangle = \frac{1}{Z_L}\langle W|C^{L-1}|V\rangle = \frac{Z_{L-1}}{Z_L} = J \quad (14.1.9)$$

where we have used the boundary equations from (14.1.6).
One also can show that the conditions (14.1.6) are sufficient for (14.1.1) being a solution of the stationary master equation, but this requires more complicated calculations.
In order to calculate expectation values one can use use explicit representations for D, E, $\langle W|$ and $|V\rangle$. What is the form of these representations? First we assume that the matrices E and D commute. Then we have

$$\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)\langle W|V\rangle = \langle W|D+E|V\rangle = \langle W|DE|V\rangle = \langle W|ED|V\rangle = \frac{1}{\alpha\beta}\langle W|V\rangle\,, \quad (14.1.10)$$

where we have used (14.1.6) several times, so that we find the condition

$$\alpha+\beta = 1 \quad (14.1.11)$$

for which representations with commuting E and D exist. Then E and D can be chosen as real numbers e, d (one-dimensional representations) and and explicit solution is given by (see Excercise 15)

$$e = \frac{1}{\alpha}\,, \qquad d = \frac{1}{\beta}\,. \quad (14.1.12)$$

The matrix-product Ansatz (14.1.1) then factorizes and takes the form of a mean-field approximation. The density profile ρ_j on the line (14.1.11) is flat, i.e. independent of j.
Next we consider the case where E and D do not commute. In that case the representations are infinite-dimensional. This can be shown in the following way: first it is clear that no vector $|X\rangle$ with $E|X\rangle = |X\rangle$ can exist. Otherwise we would have

$$D|X\rangle = DE|X\rangle = (D+E)|X\rangle = D|X\rangle + E|X\rangle = D|X\rangle + |X\rangle \quad (14.1.13)$$

so that $|X\rangle = 0$. If we assume now that E and D have finite-dimensional representations, then the matrix[1] $I-E$ would be invertible since it has no eigenvalue 1. Therefore we can use $DE = D+E$ to express D in the form

$$D = E\,(E-I)^{-1} \quad (14.1.14)$$

which implies that D and E commute. This is a contradiction to our original assumption.
So we see that the algebraic relations (14.1.6) either have one-dimensional representations in the case $\alpha+\beta=1$ or infinite-dimensional one in all other cases. Some explicit represenations are given in Excercise 15.

[1] Here I denotes the unit matrix.

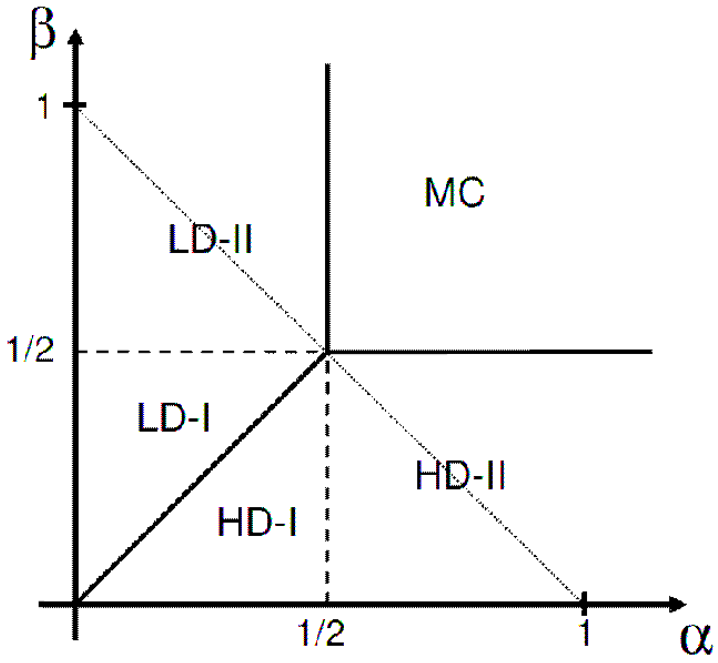

Figure 14.2.1: Exact phase diagram for the TASEP in continuous time for $p = 1$. The two high- and low-density sub-phases differ in the decay of the density profile to its bulk value from the boundaries. On the thin dotted line $\alpha + \beta = 1$ the mean-field approximation becomes exact (corresponding to a one-dimensional representation of the matrix-product Ansantz) and the density profile becomes flat.

14.2 Phase diagram

The phase diagram consists of three different phases, the **high-density, low-density** and **maximal current phase** (Fig. 14.2.1). They can be distinguished by the dependence of the current on the system parameters. Using (14.1.7) the current is found to be

$$J = \begin{cases} 1/4 & \text{for } \alpha \geq 1/2,\ \beta \geq 1/2 \\ \alpha(1-\alpha) & \text{for } \alpha < 1/2,\ \beta > \alpha \\ \beta(1-\beta) & \text{for } \beta < 1/2,\ \alpha > \beta \end{cases} . \tag{14.2.1}$$

In the low-density phase $\alpha < 1/2$, $\beta > \alpha$ the current only depends on the input rate and is independent of β. In the high-density phase $\beta < 1/2$, $\alpha > \beta$ it only depends on β and is independent of α. Finally, in the maximal current phase $\alpha, \beta \geq 1/2$ it is independent of both boundary rates. Here it reaches the maximum current 1/4 which can be realized in the periodic system, i.e. the maximum of the fundamental diagram[2].
Why do we find three different phases? This can be understood by a simple argument. In the stationary state the current through the system has to be the same everywhere due the equation

[2]Remember that we have set $p = 1$.

phase	$J(\alpha,\beta)$	$\rho_{L/2}$	ρ_1	ρ_L	left end	right end
LD-I	$\alpha(1-\alpha)$	α	α	$\frac{\alpha(1-\alpha)}{\beta}$	α	$e^{-j/\xi}$
LD-II	$\alpha(1-\alpha)$	α	α	$\frac{\alpha(1-\alpha)}{\beta}$	α	$j^{-3/2}e^{-j/\xi_\alpha}$
HD-I	$\beta(1-\beta)$	$1-\beta$	$1-\frac{\beta(1-\beta)}{\alpha}$	$1-\beta$	$e^{-j/\xi}$	$1-\beta$
HD-II	$\beta(1-\beta)$	$1-\beta$	$1-\frac{\beta(1-\beta)}{\alpha}$	$1-\beta$	$j^{-3/2}e^{-j/\xi_\beta}$	$1-\beta$
MC	$1/4$	$1/2$	$1-\frac{1}{4\alpha}$	$\frac{1}{4\beta}$	$\frac{1}{2\sqrt{\pi}}j^{-1/2}$	$-\frac{1}{2\sqrt{\pi}}j^{-1/2}$

Table 14.1: Characteristic properties of the phases in the TASEP with random-sequential dynamics ($p = 1$). It shows the current J, the bulk density $\rho_{L/2}$, the densities ρ_1, ρ_L at the chain ends and the asymptotic decay of the density profile near the left (input) and right (output) end of the chain. For $p \neq 1$ the current is given by $pJ(\alpha/p, \beta/p)$.

of continuity. There are three different mechanisms which in principle may limit the current, namely the efficiency of the input (corresponding to the low-density phase), the output (corresponding to the high-density phase) and the transport in the bulk (corresponding to the maximal current phase). The low-density phase can also be called **input-controlled phase**, and high-density phase **output-controlled phase**. In contrast to these two **boundary-controlled phases**, the maximal current phase is a **bulk-controlled** phase. Here the current through the system has reached the maximum of the fundamental diagram of the corresponding ASEP with periodic boundary conditions.
The density profile can also be calculated analytically. We already know that on the line $\alpha+\beta = 1$ it is flat, i.e. $\rho_j = \text{const.}$, independent of j. In general, the **bulk density** far away from both boundaries is of special interest. One finds that, except for the maximum current phase, the density profile becomes flat in the bulk and shows a dependence on j only in a region near the boundaries. There the behaviour can be characterized by two independent length scales ξ_α and ξ_β determined by the input and output rate, respectively. Explicitly these length scales are given by

$$\xi_\sigma = -\ln[4\sigma(1-\sigma)] \qquad \text{with } \sigma = \alpha, \beta\,. \tag{14.2.2}$$

There is a third length scale ξ that becomes relevant which is given by

$$\frac{1}{\xi} = \left| \frac{1}{\xi_\alpha} - \frac{1}{\xi_\beta} \right| . \tag{14.2.3}$$

In the maximal current phase the density profile does not become flat even in the bulk. Instead an algebraic behaviour $\rho_j \propto j^{-1/2}$ is observed.
The main features of the different phases are summarized in Table 14.1.

14.3 Boundary-induced phase transitions

In nonequilibrium system like the TASEP a new kind of phase transitions can be observed which is usually not found in equilibrium. This has first been discovered by Krug in 1991 for the

TASEP with $\beta = 1$, but the phase diagram in Sec. 14.2 shows that this is a rather generic feature of driven diffusive systems. In contrast to "normal" equilibrium systems, the state in the bulk of nonequilibrium systems can change qualitatively if the boundary conditions (here: the rates α, β) are changed. Such transitions are called **boundary-induced phase transitions**.
In equilibrium this is not possible, at least for systems with short-ranged interactions. In a d-dimensional system the boundary contribution to the free energy F is of the order L^{d-1} where L is the typical length of the system. The bulk contribution, on the other hand, is proportional to the volume, i.e. of the order L^d. Therefore,

$$F = F_{\text{Bulk}} + F_{Bound} \sim L^d + L^{d-1} \to F_{\text{Bulk}} \qquad (\text{for } L \to \infty)\,, \tag{14.3.1}$$

so that in the thermodynamic limit only the bulk terms contribute and the precise form of the boundary conditions becomes irrelevant.

Returning to the TASEP, we have seen that the behavior of the density profiles is characterized by three length scales ξ_α, ξ_β and ξ. They describe how far the boundary conditions can be "felt" away from the boundary. If one of them becomes infinite, the boundaries influence the whole system and a boundary-induced phase transition occurs (at least in principle). At the transition from the high-density phase to the maximal current phase, ξ_β diverges. Similarly at the boundary between the low-density and maximal current phase, ξ_α is divergent. The density profiles, which can be considered as a kind of order parameter, change in a continuous way when passing through these phase boundaries so that the transition can be specified as a phase transition of second order. At the boundary between the high- and low-density phase both ξ_α and ξ_β are finite, but ξ becomes infinite. Here the density changes discontinously when passing through the phase boundary and the transition is of first order. The density profile on the boundary $\alpha = \beta \leq 1/2$ is linear with $\rho_1 = \alpha$ and $\rho_L = 1 - \alpha$. However, making a snapshot of a simulation one observes a different type of density profile, namely **phase separation** into a low density region at the input end of the system and a high density region at the output end. These two regions are separated by a sharp **shock** or **domain wall** (Fig. 14.3.1). This shock diffuses through the system, i.e. it performs a symmetric random walk with average velocity $\langle v_S \rangle = 0$. Average over all possible shock positions than leads to the linear density profile decribe above.

14.4 Extremal principle

Much information about the system with open boundary conditions can be derived if one knows only the fundamental diagram (flow-density relation) $J(\rho)$ of the system with periodic boundary conditions! If we assume that we know the densities ρ_l and ρ_r of the open system at the left and right boundaries, respectively, the current $J(\rho_l, \rho_r)$ in the open system is determined by the **extremal principle**

$$\boxed{J(\rho_l, \rho_r) = \begin{cases} \max_{\rho \in [\rho_r, \rho_l]} J(\rho) & \text{for } \rho_l > \rho_r \\ \min_{\rho \in [\rho_l, \rho_r]} J(\rho) & \text{for } \rho_l < \rho_r \end{cases}}\,. \tag{14.4.1}$$

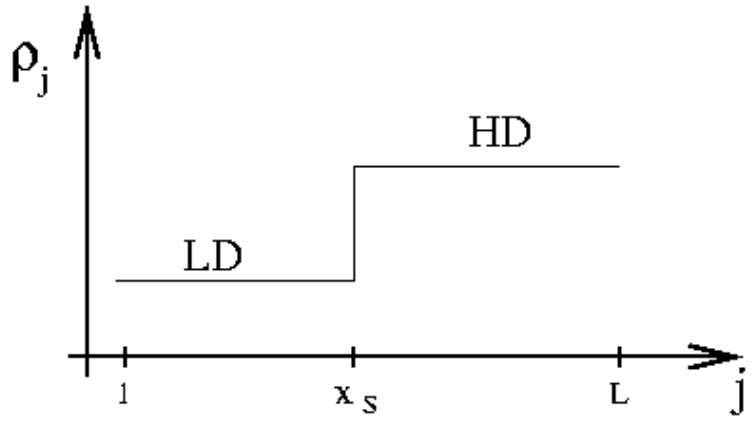

Figure 14.3.1: Snapshot of the density profile of the TASEP on the boundary between the high- and low-density phase ($\alpha = \beta \leq 1/2$). It shows phase separation into a LD and a HD region separated by a sharp shock (domain wall) at x_S.

The boundary densities are controlled by the reservoirs. For the TASEP we have $\rho_l = \alpha$ and $\rho_r = 1 - \beta$.
The extremal principle can also be interpreted as **steady state selection principle**: the state realized in an open system is selected by the boundary conditions from the steady states of the periodic system.
As an example for the use of the extremal principle we consider the case where $\rho_r < \rho_l < 1/2$. Therefore $J(\rho_l, \rho_r) = \max_{\rho \in [\rho_r, \rho_l]} J(\rho)$. Since both densities are smaller than 1/2 we are in the monotonically increasing part of the fundamental diagram and thus $J(\rho_l, \rho_r) = J(\rho_l)$. Expressing this by the boundary rates we finally have $J(\alpha, \beta) = J(\alpha) = \alpha(1 - \alpha)$ which is exactly the current in the low-density phase (see Table 14.1). This is indeed correct because $\rho_r < \rho_l < 1/2$ corresponds to $\alpha < 1/2$ and $\beta > 1/2$. Similar considerations can be made for any other combination of ρ_r and ρ_l, e.g. $\rho_r < 1/2 < \rho_l$ leads to a point in the maximal current phase with $J(\alpha, \beta) = 1/4$.
Even if the fundamental diagram is not analytically known like in the case of the TASEP, the extremal principle is extremely useful. It allows e.g. to predict, how many different phases the system with open boundary conditions will exhibit. This number of phases is closely related to the number of local extrema in the fundamental diagram. Since the TASEP fundamental diagram has only one local maximum, it shows three phases. In the exercises we will study the case of three local extrema (two maxima separated by a local minimum). In this situation one finds seven phases in the open system.

14.5 Derivation of the extremal principle

The justification of the extremal principle does not depend on the microscopic details of the dynamics. It is therefore expected to hold under rather general conditions.

For the derivation of the extremal principle two velocities are relevant which both can be determined from the fundamental diagram of the periodic system. The first is the **shock velocity**

$$v_S = \frac{J_2 - J_1}{\rho_2 - \rho_1} \tag{14.5.1}$$

which is the velocity of a shock separating a region of density ρ_1 with current J_1 from a region of density ρ_2 and current J_2 (Fig. 14.5.1). It can be derived from the equation of continuity applied to the shock region. The effective current into this area is $J_2 - J_1$. In time Δt the particle number near the shock changes by $(J_2 - J_1)\Delta t$ which leads to a motion of the shock. The change of particle number due to the shock motion, on the other hand, is $(\rho_2 - \rho_1)v_S\Delta t$. Combining these two expressions then gives (14.5.1).

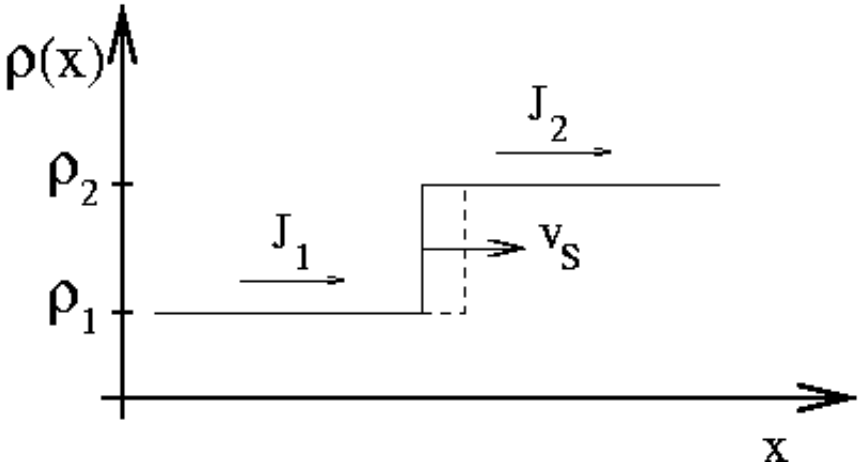

Figure 14.5.1: Derivation of the shock velocity. If the current J_2 is larger than J_1 the shock will move to the right (broken line). Applying the continuity equation to this area gives v_S.

The second velocity is the **collective velocity** defined by

$$v_c = \frac{dJ}{d\rho}\,. \tag{14.5.2}$$

It is the velocity at which a small perturbation of the density moves (Fig. 14.5.2).
The two velocities can be easily determined from the fundamental diagram (Fig. 14.5.3), even when its analytical form is not known. v_S is the slope of the secant line through the points (ρ_1, J_1) and (ρ_2, J_2). Since $v_c(\rho)$ is the derivative of the fundamental diagram it is the slope of the tangent at $(\rho, J(\rho))$. For the TASEP (in continuous time) we have $J(\rho) = \rho(1-\rho)$ and therefore

$$v_S = \frac{\rho_2(1-\rho_2) - \rho_1(1-\rho_1)}{\rho_2 - \rho_1} \qquad \text{and} \qquad v_c = 1 - 2\rho\,. \tag{14.5.3}$$

We now return to the justification of the extremal principle. Imagine that we prepare the system in an initial state which has a shock just as in Fig. 14.5.1. If $v_S > 0$ the shock will move to the

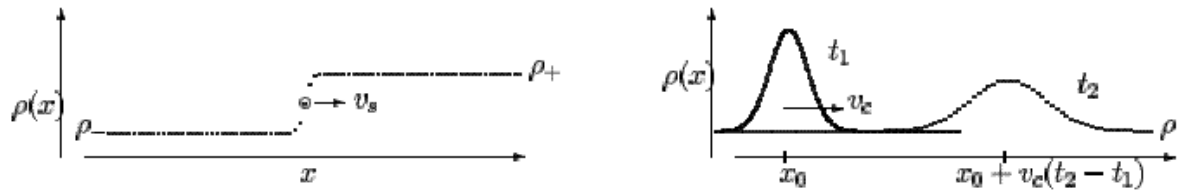

Figure 14.5.2: Illustration of shock velocity (left) and collective velocity (right).

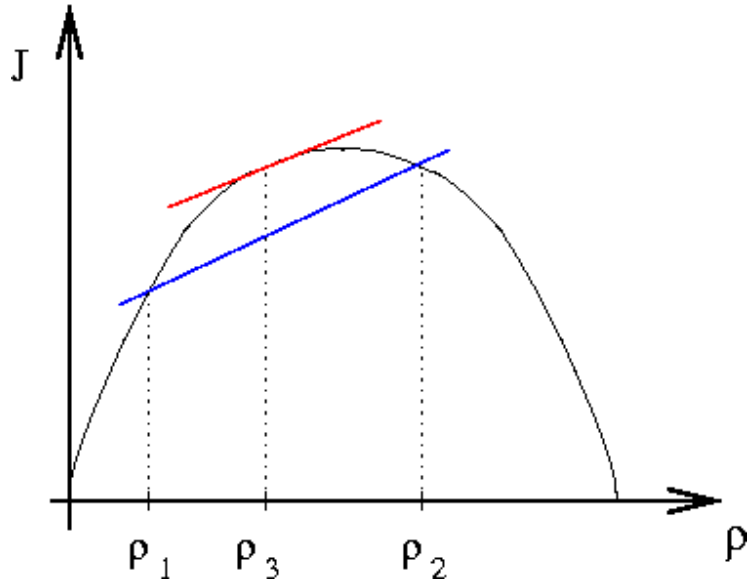

Figure 14.5.3: Shock and collective velocity can be determined from the fundamental diagram. v_S is the slope of the secant line (blue) through (ρ_1, J_1) and (ρ_2, J_2). $v_c(\rho)$ is the slope of the tangent (red) at $(\rho, J(\rho))$.

right and finally reach the right boundary. The system is then left in a state of density ρ_1. On the other hand, if $v_S < 0$ the shock will move to the left and the system reaches the density ρ_2. Although this argument is not exact (at least for discrete systems) it tells us that we can expect a phase transition when $v_S = 0$ where the two regions can coexist indicating a first order transition.

The collective velocity controls the stability of shocks. It is responsible for the second order transitions into the maximal current phases. The reason is that it changes sign at density $\rho = 1/2$ (compare with (14.5.3)).

The collective velocity allows also to understand why in a system with a fundamental diagram of the form as in Fig. 14.5.3 only upward shocks (where the density increases in the direction of motion) like in Fig. 14.5.1 are stable. The shock motion corresponds to a transport of particles from the high density region to the low density region leading to new density profile as in Fig. 14.5.4. The areas move with the respective collective velocities. In the case of a downward

shock (in which the density decreases in the direction of motion) and a fundamental diagram like Fig. 14.5.3 these velocities lead to a further transport of particles from the high to the low density region and finally the complete "melting" of the shock. One can easily convince oneself that for an upward shock the situation is different. Here the two collective velocities point backwards into the direction of the shock location which implies that such a "melting" of the shock will not grow. Instead it leads to a healing of the shock fluctuation and its stabilization.

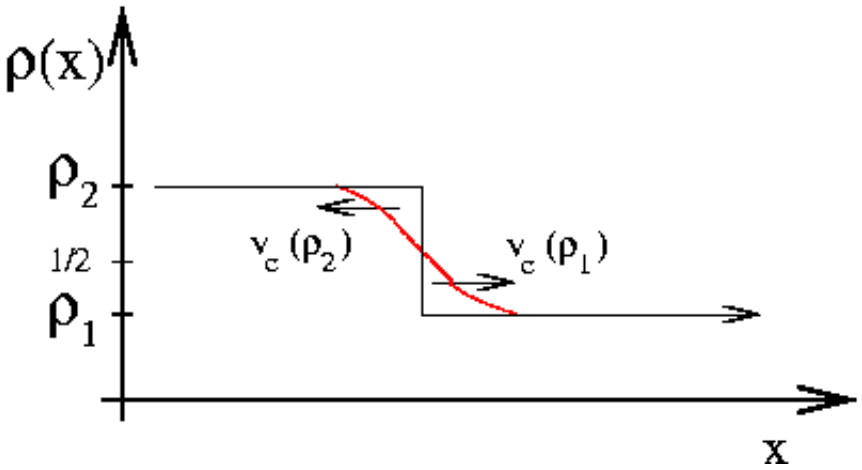

Figure 14.5.4: Downward shocks are unstable in TASEP-like systems. The shock moves by "melting" (red) where the two regions move with the collective velocities $v_c(\rho_i)$. For a downward shock their behavior leads to a further melting of the shock.

Chapter 15

Vehicular traffic

15.1 Empirical results

15.1.1 Observables and data analysis

Let us first define some characteristic quantitative features of vehicular traffic. The **flow** J, which is sometimes also called **flux** or **current** or, in traffic engineering, **traffic volume**, is defined as the number of vehicles crossing a detector site per unit time. The maximum possible flow is called the **capacity** (of the road). The distance from a selected point on the leading vehicle to the same point on the following vehicle is defined as the **distance-headway** (or **spatial headway**). The **time-headway** is defined as the time interval between the departures (or arrivals) of two successive vehicles recorded by a detector placed at a fixed position on the highway. The distributions of distance-headways and time-headways are regarded as important characteristics of traffic flow. For example, larger headways provide greater margins of safety whereas higher capacities of the highway require smaller headways.

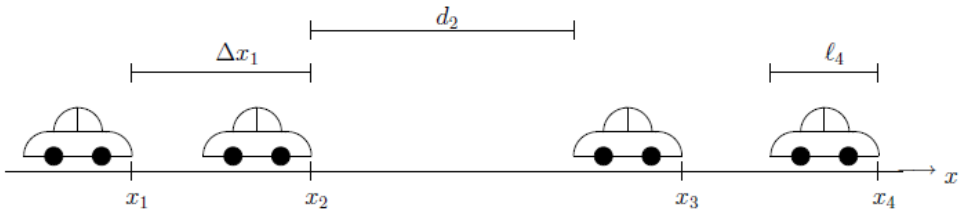

Figure 15.1.1: Definition of the quantities used to describe the configuration of vehicles on a one-dimensional road. The cars move from left to right and their label is increasing in driving direction (downstream). The length of car j is ℓ_j, its distance-headway d_j and $\Delta x_j = x_{j+1} - x_j$ the difference of the vehicle positions x_j.

15.1.2 Spontaneous traffic jams

What makes the study of traffic jams so interesting is that jams often seem to appear from nowhere (apparently without obvious reasons) suddenly on crowded highways. These so-called **phantom jams** are formed by *spontaneous* fluctuations in an otherwise streamlined flow. Its occurrence is a typical example of a **collective phenomenon** which can not be understood by just looking at the behaviour of one particle: each driver wants to drive as fast as possible so jams should never occur! Spontaneous jam formation is this a typical **emergent phenomenon** which requires to look at the whole system. This is why methods of physics are well suited for studying traffic problems since there we know many examples of such collective phenomena in many-particle systems, e.g. magnetism, superconductivity et.
Direct empirical evidence for this **spontaneous formation of jams** was presented by Treiterer by analyzing a series of aerial photographs of a multi-lane highway. In Fig. 15.1.2 the picture from Treiterer is redrawn. Each line represents the trajectory of an individual vehicle on one lane of the highway. The space-time plot (i.e., the trajectories $x(t)$ of the vehicles) shows the formation and propagation of a traffic jam. In the beginning of the analysed time vehicles are well separated from each other. Then, due to fluctuations, a dense region appears which, finally, leads to the formation of a jam. The jam remains stable for a certain period of time but, then, disappears again without any obvious reason. This figure clearly establishes not only the spontaneous formation of traffic jam but also shows that such jams can propagate upstream (i.e. backwards, opposite to the direction of flow of the vehicles). Moreover, it is possible that two or more jams coexist on a highway.
Wide jams have some characteristic features. The upstream velocity and, therefore, the outflow (mean values of flux, density and velocity) from a jam is approximately constant. The outflow from a jam and the velocity of the jam fronts are now regarded as two important empirical parameters of highway traffic which can be used for calibrating theoretical models.

Some traffic engineers, however, still argue that spontaneous jam formation does not occur and that jams are always induced by bottlenecks. Usually these are road inhomogeneities such as ramps or sharp bends, or temporary disturbances through lane changes etc. Once a jam has occurred due to such a (temporary) bottleneck it might not dissipate immediately causing most drivers to wonder about the reason for the congestion.

15.2 Models of highway traffic

15.3 Nagel-Schreckenberg model

We consider a road divided into L cells which can contain at most one vehicle. Until not stated otherwise we will assume periodic boundary conditions for simplicity. On the street, N vehicles are moving so that their density is given by $\rho = N/L$.
In the **Nagel-Schreckenberg (NaSch) model**[1], the speed v of each vehicle can take one of the

[1]The model was developed by Kai Nagel and Michael Schreckenberg in 1992 when both were at the University

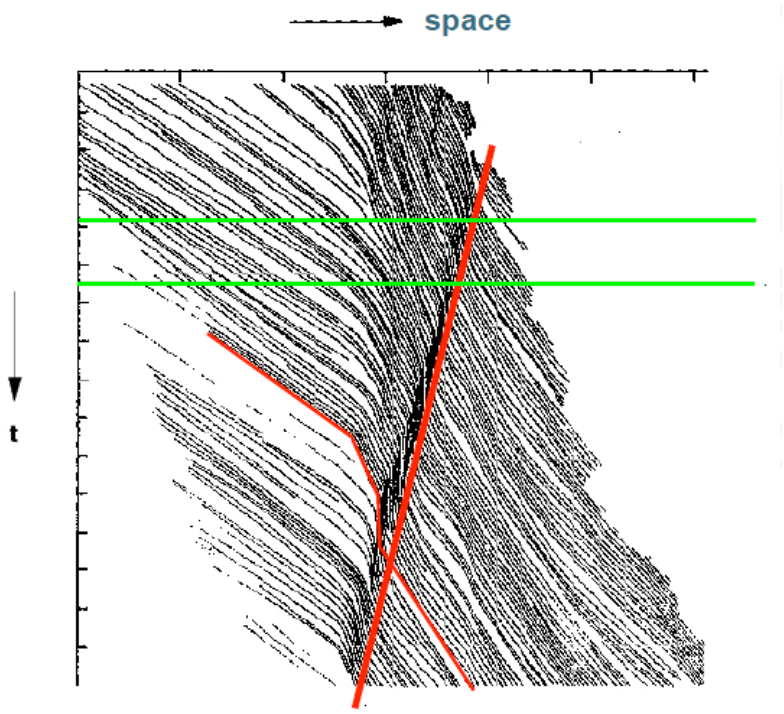

Figure 15.1.2: Trajectories of individual vehicles on a single lane of a multi-lane highway. The trajectories were drawn from aerial photographs. During the analyzed time-interval the spontaneous formation, propagation and dissolution of a jam has been observed. The trajectories have been determined from aerial photographs which give the vehicles positions at certain times (green horizontal lines). The diagonal red line indicates the jam front which moves at constant speed opposite to the driving direction of the vehicles. The discontinuous trajectories correspond to vehicles changing the lane.

$v_{\max}+1$ allowed *integer* values $v = 0, 1, ..., v_{\max}$. Suppose, x_n and v_n denote the position and speed, respectively, of the n-th vehicle. Then, $d_n = x_{n+1} - x_n - 1$, is the (spatial) headway of the n-th vehicle at time t, i.e. the number of empty cells in front of this car (see Fig. 15.1.1) At each time step $t \to t+1$, the arrangement of the N vehicles on a finite lattice of length L is updated *in parallel* according to the following rules:

NaSch1: *Acceleration.* If $v_n < v_{\max}$, the speed of the n-th vehicle is increased by one, but v_n remains unaltered if $v_n = v_{\max}$, i.e.

$$v_n \to \min(v_n + 1, v_{\max})$$

NaSch2: *Deceleration (due to other vehicles).* If $v_n > d_n$, the speed of the n-th vehicle is reduced to d_n, i.e.,

$$v_n \to \min(v_n, d_n)$$

of Cologne.

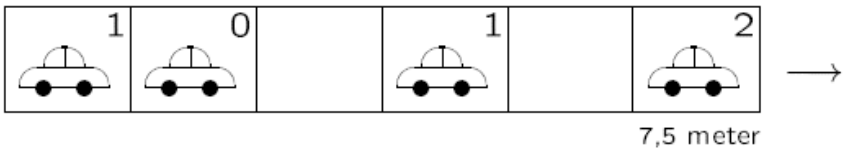

Figure 15.3.1: A typical configuration in the NaSch model. The number in the upper right corner is the speed of the vehicle.

NaSch3: *Randomization.* If $v_n > 0$, the speed of the n-th vehicle is decreased randomly by unity with probability p but v_n does not change if $v_n = 0$, i.e.,

$$v_n \to \max(v_n - 1, 0) \quad \text{with probability } p$$

NaSch4: *Vehicle movement.* Each vehicle is moved forward according to its new velocity determined in *(NaSch1)–(NaSch3)*, i.e.

$$x_n \to x_n + v_n$$

The rules can be summarized in a more formal and concise way:

$$\boxed{x_n(t+1) = x_n(t) + \max\left[0, \min\left\{v_{\max}, v_n(t)+1, d_n(t)\right\} - \eta_n(t)\right] .} \tag{15.3.1}$$

Here $\eta_n(t)$ is a Boolean random variable which takes the value 1 with probability p and 0 with probability $1-p$ and accounts for the randomization step. The term given by the maximum is the new velocity $v_n(t+1)$. Note that the rules can be expressed solely in terms of the positions x_n by using $d_n(t) = x_{n+1}(t) - x_n(t) - 1$ and $v_n(t) = x_n(t) - x_n(t-1)$. This also shows that the configuration at time $t+1$ depends both on that at time t and $t-1$.
The NaSch model is a minimal model in the sense that all the four steps are necessary to reproduce the basic features of real traffic. However, additional rules need to be formulated to capture more complex situations. The rule (NaSch1) reflects the general tendency of the drivers to drive as fast as possible, if allowed to do so, without crossing the maximum speed limit. The rule (NaSch2) is intended to avoid collision between the vehicles. The randomization in (NaSch3) takes into account the different behavioral patterns of the individual drivers, especially, nondeterministic acceleration as well as overreaction while slowing down; this is crucially important for the spontaneous formation of traffic jams. Even changing the precise order of the steps of the update rules stated above would change the properties of the model. E.g. after changing the order of rules (NaSch2) and (NaSch3) there will be no overreactions at braking and thus no spontaneous formation of jams.
The NaSch model may be regarded as stochastic CA. In the special case $v_{\max} = 1$ it reduces to the TASEP with parallel dynamics. The deterministic limit $p = 0$ is then equivalent to the CA rule 184 in Wolfram's notation.

Why should the updating be done in *parallel*, rather than in random sequential manner, in traffic models like the NaSch model? In contrast to a random-sequential update, parallel update can lead to a chain of overreactions. Suppose, a vehicle slows down due the randomization step. If the density of vehicles is large enough this might force the following vehicle also to brake in the deceleration step. In addition, if p is larger than zero, it might brake even further due to rule (NaSch3). Eventually this can lead to the stopping of a vehicle, thus creating a jam. This mechanism of spontaneous jam formation is rather realistic and cannot be modeled by the random-sequential update.
The update scheme of the NaSch model is illustrated with a simple example in Fig. 15.3.2.
Space-time diagrams showing the time evolutions of the NaSch model demonstrate that no jam is present at sufficiently low densities, but spontaneous fluctuations give rise to traffic jams at higher densities (Fig. 15.3.3(a)). From the Fig. 15.3.3(b) it should be obvious that the *intrinsic stochasticity* of the dynamics, arising from non-zero p, is essential for triggering the jams. For a realistic description of highway traffic, the typical length of each cell should be about 7.5 m which is the space occupied by a vehicle in a dense jam. When $v_{\max} = 5$ each time step should correspond to approximately 1 sec of real time which is of the order of the shortest relevant timescale in real traffic, namely the reaction time of the drivers.

15.4 Generalizations of the NaSch model

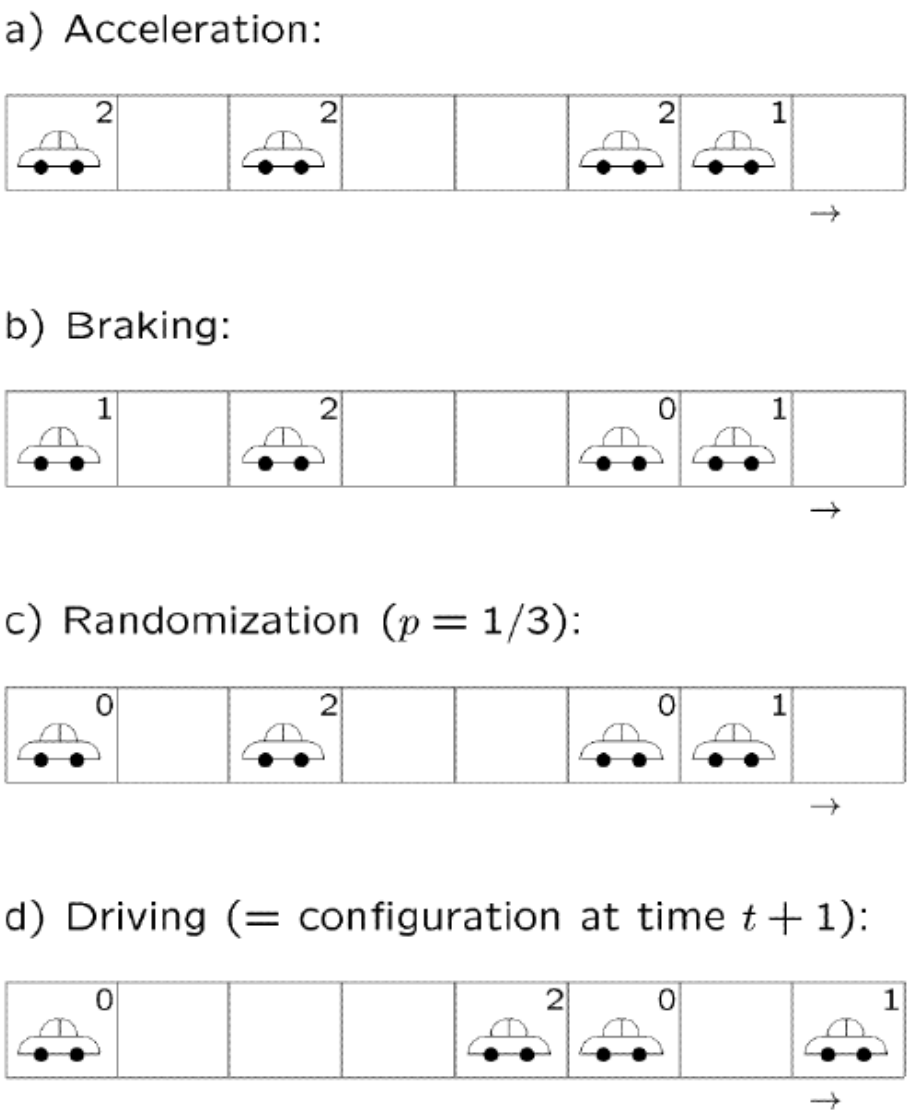

Figure 15.3.2: Step-by-step example for the application of the update rules. We have assumed $v_{\max} = 2$ and $p = 1/3$. Therefore on average one third of the cars qualifying will slow down in the randomization step.

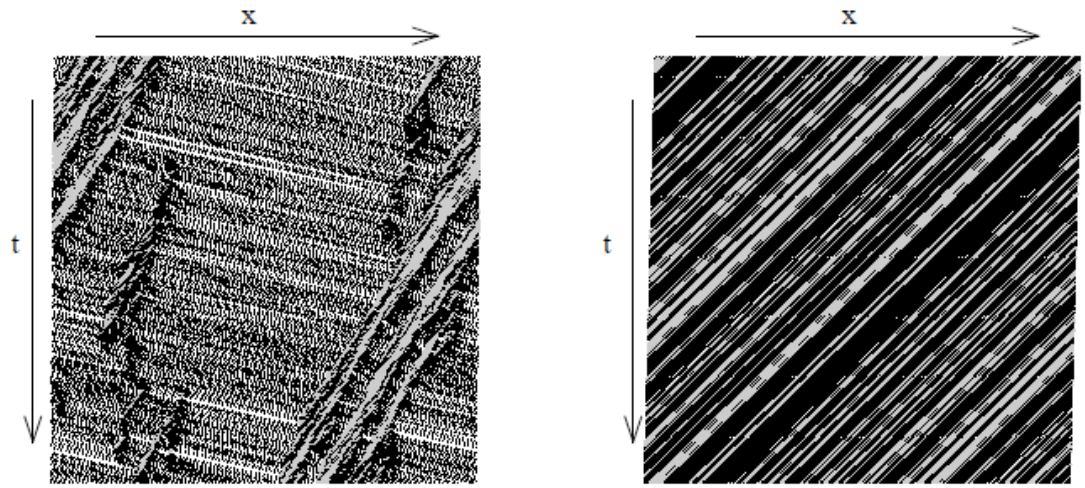

Figure 15.3.3: Typical space-time diagrams of the NaSch model with $v_{\max} = 5$ and (a) $p = 0.25, \rho = 0.20$, (b) $p = 0, \rho = 0.5$. Each horizontal row of dots represents the instantaneous positions of the vehicles moving towards right while the successive rows of dots represent the positions of the same vehicles at the successive time steps.

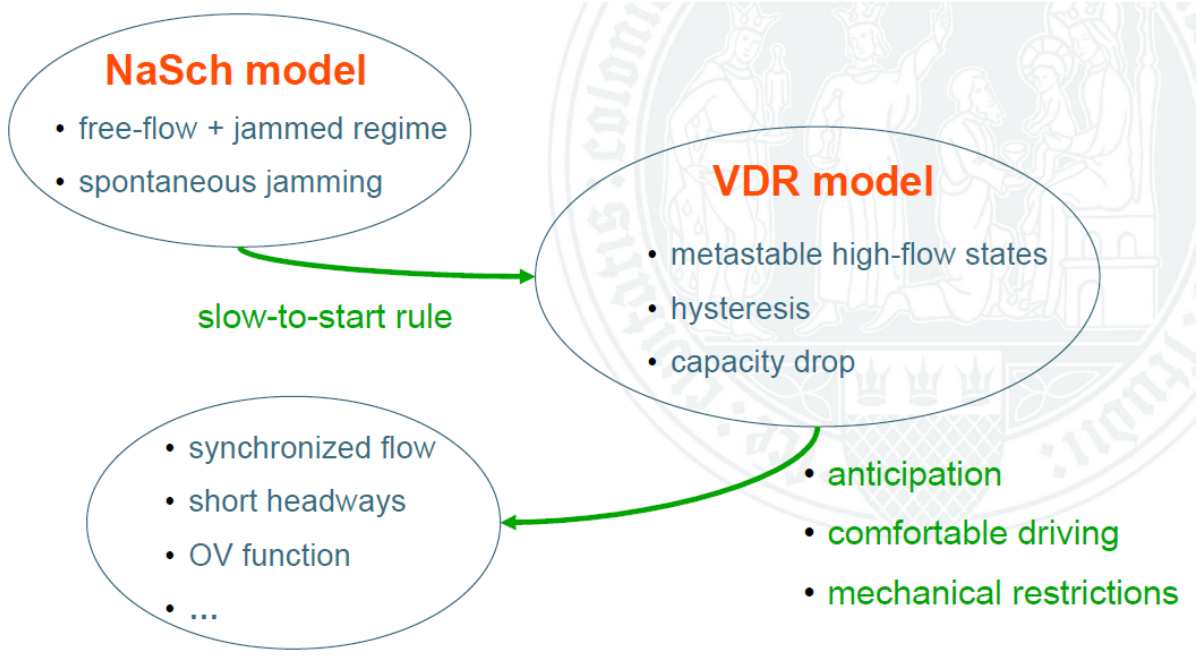

Figure 15.4.1: Generalizations of the Nagel-Schreckenberg model.

Chapter 16

Biological transport

16.1 Traffic on ant trails

16.2 Intracellular transport

Chapter 17

Pedestrian dynamics

17.1 Empirical results

17.2 Modeling approaches

17.3 Applications

Bibliography

[1] Source: Wikipedia

[2] C. Sire: *Universal statistical properties of poker tournaments*, J. Stat. Mech. (2007) P08013 (available from `http://arxiv.org/abs/physics/0703122`)

[3] P. Reimann: *Brownian motors: noisy transport far from equilibrium*, Rev. Mod. Phys. 361, 57 (2002) (available from `http://arxiv.org/abs/cond-mat/0010237`)

[4] R. Feynman: *The Feynman Lectures on Physics*, Vol. 1, Chapter 46 (Addison-Wesley, 1963)

[5] T.R. Kelly, I. Tellitu, J.P. Sestelo: *In search of molecular ratchets*, Angew. Chem. Int. Ed. Engl. 36, 1866 (1997)

[6] N.G. van Kampen: *Stochastic Processes in Physics and Chemistry* (North Holland)

[7] A. Schadschneider, D. Chowdhury, K. Nishinari: *Stochastic Transport in Complex Systems* (Elsevier, 2010)

[8] G.M. Schütz: *Exactly solvable models for many-body systems far from equilibrium*, in *Phase Transitions and Critical Phenomena, Vol. 19*, (Eds.: C. Domb and J. L. Lebowitz), Academic Press, 2001

[9] R. Blythe, M.R. Evans: *Nonequilibrium steady states of matrix product form: a solver's guide*, J. Phys. A40, R333 (2007)

[10] P.L. Krapivsky, S. Redner, E. Ben-Naim: *A Kinetic View of Statistical Physics* (Cambridge University Press, 2010)

[11] V. Privman (Ed.): *Nonequilibrium Statistical Mechanics in One Dimension* (Cambridge University Press, 1997)