The Einstein Smoluchowski Equation in the One Dimensional Exclusion Process

Steven Berghout

Utrecht University

August 26, 2013

Abstract

The Einstein-Smoluchowski equation is a major result in the study of diffusion. For systems with anomalous diffusive behaviour it is not always valid, however it does hold in the one dimensional exclusion process. This is shown by Landim, Olla and Volchan. In order to discuss their method the reader is first introduced in the theory of Markov chains. Then some definitions and methods to rigorously determine hydrodynamic limits will be introduced. Finally we discuss the outline of their proof that relies on these notions.

Contents

Preface

This thesis was written as part of the masters programmes for Mathematical Sciences and Theoretical Physics at the University of Utrecht. I would like to thank Gerard Barkema and Roberto Fernandez both for giving up part of their time supervising my thesis. But also for the chance they thereby gave me to conclude my master programmes at a moment that it looked like I would have to abandon one of them. Furthermore I would like to thank Gerard Barkema for being very involved, even while supervising a large number of students, and also for giving a fun introduction in a subject completely new to me. I would like to thank Roberto Fernandez for helping me patiently when I had trouble with parts of the literature and for his supportive feedback.

Steven Berghout

1 Diffusion and the Einstein relation

1.1 Normal diffusion

Diffusion is a process causing substances to redistribute themselves, such that there is a nett transport from areas with a higher concentration to areas with a lower concentration. Modern research on the subject started in the the first half of the nineteenth century. Early work was done by Thomas Graham who, between 1828 and 1833, investigated diffusion in gasses. However until 1855 there was no fundamental law of diffusion available. This was solved by Adolf Fick who formulated the law below that now bears his name, although he formulated it differently:

$$\vec{J} = -D\vec{\nabla}\rho. \tag{1}$$

Fick's law describes the evolution of a dilute tracer substance in a background. The left-hand side is the flux of this substance while the right-hand side is a multiple of the gradient of the density of this tracer substance. The coefficient D is called the diffusion constant and depends on the system at hand. The law is a phenomenological one and can be used, if combined with some other basic equations, to derive a mathematical framework for diffusion. An important equation in this context is the continuity equation

$$\frac{\partial \rho}{\partial t} = -\vec{\nabla} \cdot \vec{J}. \tag{2}$$

This basically states that the change in density of a substance equals the divergence of the flux. Additional terms can be added if particles can be created or destroyed, but if we assume this is not the case the two equations combine to give Fick's second law

$$\frac{\partial \rho}{\partial t} = D\Delta\rho, \tag{3}$$

where $\Delta = \vec{\nabla} \cdot \vec{\nabla}$ is the Laplacian. This is a differential equation for the density of the tracer substance. The complexity of the solution largely depends on the initial particle distribution, however quite some progress can be made before this is going to matter. A method that delays the usage of the initial distribution is solving the equation using Green's functions. The idea behind this method is to solve the equation

$$\left(\frac{\partial}{\partial t} - D\Delta\right)G(r, t) = \delta^{(d)}(r)\delta(t). \tag{4}$$

Note that this quantity must be interpreted in the context of distribution theory to make sense. For now the dimension of space, d, will not be specified. If one has a solution to this equation then the time dependent density is given, in terms of the initial density profile, by

$$\rho(t, u) = \int G(r - r_0, t_0)\rho(r_0, 0)dr_0. \tag{5}$$

The Green's function that solves the above differential equation is given by

$$G(u, t) = \frac{e^{-\frac{r^2}{4Dt}}}{(4\pi Dt)^{d/2}}. \tag{6}$$

In order to derive this result one can write the Dirac delta functions as a Fourier transform:

$$\delta^{(d)}(x)\delta(t) = \frac{1}{(2\pi)^{d+1}} \int_{\mathbb{R}^d} \int_{\mathbb{R}} e^{i(px+ts)} dp ds. \tag{7}$$

Assume then that the Green's function can also be written as a Fourier transform of some unknown $\tilde{G}(p, s)$. If we then let the differential operator act on the Green's function we can exchange the order of differentiation and integration and perform the derivation explicitly. This yields a solution $\tilde{G}(p, s) = \frac{1}{is+Dp^2}$. To obtain the expression for G itself the inverse Fourier transform can be used. That integral can be performed by a smart choice of polar coordinates such that $p \cdot x = ||p||||x|| \cos(\theta)$ where θ is also the azimuthal angle in the coordinate system. Then the angle integrals can be performed, the remaining integrals require a contour integration. There are also other methods to approach this problem, for example one can expand the density in terms of eigenmodes of the differential operator. More general forms of the diffusion equation, where the diffusion coefficient can depend on position, can also be approached through Green's function methods [3]. A fundamentally different approach to diffusion is
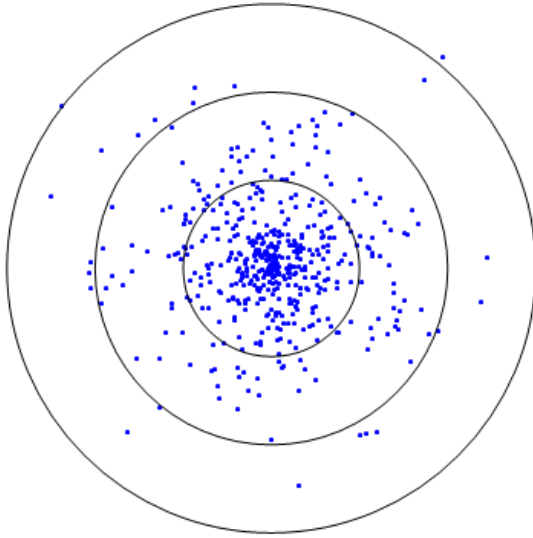

Figure 1: An illustration of the distribution of particles, diffusing normally, when all start out in the origin. The circles represent $1, 2$ and 3 unit lengths

offered if one starts out from the study of random walks and Brownian motions. Brownian motion is a process named after Robert Brown who observed the

behaviour of pollen grains in water. Such particles move through the liquid driven by the impulse transfer from the even smaller water molecules. Brownian motion itself is much more general than just the physical process of particles in a background substance. It even finds applications in the social sciences and economics. Major work on this subject was done by Albert Einstein. In [1] he determines the mean squared displacement of the particles, for large times, in a one dimensional system, to be equal to $< R^2(t) >= 2Dt$, with D the diffusion coefficient. In a symmetric system the mean displacement $< R(t) >$ itself is zero, as is the velocity, thus this is indeed the important quantity at hand. His approach can be generalised to three dimensions where the mean squared displacement can be written as

$$< R^2(t) >= 6Dt, \tag{8}$$

where D is the diffusion coefficient and the factor 6 can be replaced by $2d$ if one wishes to work in a different dimension. Assuming that the underlying mechanism, being the momentum transfer from particles in the background substance, he derived numerical predictions for the mean squared displacement of a Brownian particle. In the process he arrived at an equation that would allow for an experimental determination of the Avogadro constant. These results were indeed used in an experimental verification of the atomic model by Jean Baptiste Perrin who later received a Nobel prize for it. Nowadays a more abstract, mathematical approach exists to Brownian motion. In this context one can describe it as a stochastic process B_t, taking values in $\mathbb{R}$ with $t \in \mathbb{R}$, satisfying the following requirements

1. B_t is almost surely continuous.

2. $B_t - B_s$ has a Gaussian distribution with mean zero and variance $t - s$.

3. $B_0 = 0$

It is possible to realize a Brownian motion, or Wiener process as it is sometimes called, as a limit on random walks. Suppose that $\{X_i : i \in \mathbb{N}\}$ is a set of iid Bernouilli trials of parameter $p = \frac{1}{2}$. Then the following random variables $S_N = \sum_{i=1}^{N} X_i$ form a simple random walk, by definition $S_0 = 0$. To let this become a Wiener process one would need to rescale the lattice on which the random variables take values. However, simply rescaling by a factor of N would cause all dynamics to disappear almost surely. This can be stated more formally as

$$\lim_{N \to \infty} \mathbb{P}\left(\left| \frac{S_N}{N} \right| > \delta \right) = 0, \tag{9}$$

for all $\delta > 0$. This is a direct consequence of the central limit theorem. In order to have nontrivial dynamics we can rescale time in such a way that the dynamics does not disappear nor blow up. Since the mean squared displacement in this systems scales with t we have a logical candidate for how to rescale the process.

Indeed the following random variable is a Brownian motion.

$$B_t = \lim_{N \to \infty} \frac{S_{\lfloor Nt \rfloor}}{\sqrt{N}} \tag{10}$$

It might come as a surprise that different rescalings exist that can lead to processes that neither vanish nor blow up. Consider the quantity $\limsup_{N \to \infty} \frac{S_N}{N \log(\log(N))}$, this is almost surely equal to $\sqrt{2}$. Brownian motion is indeed another approach to diffusion then Fick's law, however, it is still very much related. In fact it is a realisation of a system that behaves according to Fick's laws, since they are solved by a Gaussian distribution with variance scaling as $2Dt$. And for a large number of identical Brownian particles we expect the macroscopic density to correspond with their probability distributions. In a sense Brownian motion can be considered a microscopic description of certain diffusive processes whereas Fick's law gives a macroscopic description. As it turns out such a relation between a microscopic and macroscopic system can be key in solving questions about these systems. This feature will be a major theme throughout this text. A powerful tool in the study of Brownian motion is the Langevin equation, a stochastic differential equation used to describe the behaviour of such a particle:

$$m \frac{\partial^2 x}{\partial t^2} = -\gamma \frac{\partial x}{\partial t} + m\xi(t). \tag{11}$$

In this equation m is the mass of the particle, γ is a friction coefficient that arises due to the background fluid and $m\xi(t)$ a random variable representing the force on the particle. The precise properties of the random variable are determined by the system at hand. A common choice is that of white noise, meaning that ξ is Gaussian and $< \xi(t) >= 0$, $< \xi(t)\xi(t') >= \frac{\lambda}{d}\delta(|t-t'|)$, where λ is some parameter. From this equation we can estimate both the short and the long term behaviour of a Brownian particle quite easily as we will see now. If one multiplies both sides of the Langevin equation with $x(t)$ and then takes the expectation of the random variables on both sides of the equation the result is

$$m \frac{\partial < x\dot{x} >}{\partial t} = m < \dot{x}^2 > -\gamma < x\dot{x} >, \tag{12}$$

where it is used that the average $< x(t)\xi(t) >$ vanishes. If the background medium is in equilibrium then, by the equipartition theorem, the Kinetic energy of a Brownian particle is given by $m < \dot{x}^2 > /2 = k_B T/2$. Let $c = \frac{\gamma}{m}$ then, with the above equalities the equation can be reduced to

$$\left(\frac{\partial}{\partial t} + c \right) < x\dot{x} >= \frac{k_B T}{m}. \tag{13}$$

A solution to this equation is given by

$$< x\dot{x} >= c_1 e^{-ct} + \frac{k_B T}{\gamma}. \tag{14}$$

Using the boundary condition that the mean square displacement is zero at $t = 0$ we get $c_1 = \frac{k_B T}{\gamma}$, then by using $< x\dot{x} >= \frac{1}{2}\frac{\partial}{\partial t} < x^2 >$ and explicit integration we get

$$< x^2 >= \frac{2k_B T}{\gamma}\left(t - \frac{1}{c}(1 - e^{-ct})\right). \tag{15}$$

For very short times $t \ll \frac{1}{c}$ a second order Taylor expansion shows that $< x^2 >\propto t^2$ whereas for large time scales $t \gg \frac{1}{c}$ we have $< x^2 >\propto t$. The long term relation is of course expected, and the short term describes the particle moving at its thermal speed $\sqrt{\frac{k_B T}{m}}$.

1.2   Einstein-Smoluchowski relation

In the previous paragraph we referred to an article by Einstein published in 1905. This article was important in that it provided a testable prediction for the atomic theory of heat. However in the process he also derived an expression for the diffusion coefficient. This expression, the Einstein-Smoluchowski relation, is an important example of how a fluctuation in a system in equilibrium leads to similar behaviour as a small force acting on the system. In 1906, independent of Einstein, Marian Smoluchowski also published an article on Brownian motion. Let us, for now, follow the approach offered in [1] to see how Einstein arrived at this important relation. Consider a system of finite volume entirely occupied by a solvent. Now if one separates the system in two parts, by a wall permeable to this solvent, and dissolves another substance that cannot pass this wall in one side of the system then the wall will experience an osmotic pressure. This was a known phenomenon in that time. However if one would add small suspended particles to the system instead of the dissolved substance then the theory at that time did not predict an osmotic pressure. Einstein then made the important realisation that, if matter was made up of atoms and molecules, then there should be an osmotic pressure in both systems. The only difference, after all, was the size of the particles put in the solvent. Based on this idea the osmotic pressure can then be determined to be (approximately) equal to

$$p = \frac{RT}{N}\nu \tag{16}$$

with ν the density of the added particles, R is the gas constant, T denotes temperature and N denotes Avogadro's number. Suppose that in the previous system the wall is removed. Then we can let a force be applied to the suspended particles, for example an electromagnetic force. After a while the system will reach an equilibrium state. At this point the effect of diffusion and the force acting on the system should cancel out. The resulting equations can then be combined with Fick's law to give an expression for the diffusion coefficient in terms of the response of the system to a force. Here we will derive this equation in a slightly different way using an approach by Kubo [4]. Consider the effect of a force F on a particle suspended in a medium with a friction coefficient β.

For a small force we assume that there exists a terminal drift velocity v when the driving force and the force due to friction cancel. Using Stokes' law we can write frictional force as $F_f = m\beta v$, and therefore we associate to a driving force F a terminal velocity of $v = \frac{F}{m\beta}$. Mobility is a quantity usually defined as the coefficient that determines the response of a particle to a small force:

$$\mu = \lim_{F\downarrow 0} \frac{v(F)}{F}. \tag{17}$$

So in our case we have $\mu = \frac{1}{m\beta}$. Assume now that in a system with suspended Brownian particles a force is introduced via a potential field $U(x)$. The drift velocity of the Brownian particles, ignoring other effects, would be $v = -\frac{\nabla U}{m\beta}$. The flux due to diffusion is given by Fick's law, combining these results gives us the following expression for the nett flux

$$j(x) = -D\nabla\rho - \frac{\nabla U}{m\beta}\rho. \tag{18}$$

In equilibrium the nett flux must vanish everywhere. Due to the Boltzmann statistics we also know that the density must scale as $\rho(x) \propto e^{-\frac{U(x)}{k_B T}}$. Substituting this expression leaves us with

$$j(x) = \left(\frac{D}{k_B T} - \frac{1}{m\beta} \right) \nabla U \rho. \tag{19}$$

This directly implies that the Einstein relation holds: $D = \mu k_B T$. The interesting thing about this relation is that a quantity used to describe the behaviour of a substance in the absence of an external force actually tells you how that substance would, up to first order, react to a force.

Since the diffusion coefficient and mobility are so closely related any formalism used to determine one of the two immediately gives an expression for the other. The Green Kubo formalism gives us such a new expression for D. Let $x(t)$ denote the position of a particle as a function of t and let $v(t) = \frac{\partial x}{\partial t}(t)$, then

$$
\begin{aligned}
D &= \lim_{t\to\infty} \frac{1}{2dt} < |x(t) - x(0)|^2 > \\
&= \frac{1}{2d} \lim_{t\to\infty} \frac{d}{dt} < |r(r) - r(0)|^2) > \\
&= \lim_{t\to\infty} \frac{1}{d} < (r(t) - r(0)) \cdot v(t) > \\
&= \lim_{t\to\infty} \frac{1}{d} \int_0^t d\tau < v(\tau) \cdot v(t) > \\
&= \lim_{t\to\infty} \frac{1}{d} \int_0^t d\tau < v(0) \cdot v(t-\tau) > \\
&= \frac{1}{d} \int_0^\infty dt < v(0) \cdot v(t) >
\end{aligned}
\tag{20}
$$

This allows to calculate the diffusion coefficient as an autocorrelation of velocity and so we can do the same for mobility

$$\mu = \frac{1}{dk_BT} \int_0^\infty <v(0) \cdot v(t)> dt. \tag{21}$$

1.3 Anomalous diffusion and the exclusion process

If one considers a phenomenological description of the behaviour of a substance dissolved or suspended in another medium as the definition of diffusion then Fick's law is not necessarily correct in all diffusive systems. In the case of normal diffusion the square displacement, for large times, scales with t. However some diffusive systems do not have this property and scale as $<R^2(t)>\propto t^\alpha$ for some parameter $\alpha \neq 1$. The case $\alpha > 1$ is called super-diffusion while the case $\alpha < 1$ is called sub-diffusion. In order to generalise to this wider class of diffusive systems we also have to have a definition of the diffusion coefficient that is independent of Fick's law. In normal diffusion we saw that the diffusion equation implied that

$$D = \lim_{t \to \infty} \frac{<|x(t) - x(0)|^2>}{2dt}. \tag{22}$$

It is this relation with the mean squared displacement that we will use as a definition for the diffusion coefficient. An example of a sub-diffusive process is the motion of the middle monomers of polymers. Consider a phantom Rouse polymer, a polymer consisting of a chain of beads connected by harmonic strings with no other interactions with itself. The beads themselves are subject to a generalised version of the Langevin equation. The force on a monomer particle in this system can be written as

$$\phi(t) = -\int_0^t \mu(t - t')v(t')dt' + \xi(t), \tag{23}$$

where ξ is a random noise term, v the velocity of the monomer. The effect of the term containing μ is such that it lets the current force on the monomer depend on its previous velocity, therefore it is called a memory kernel. These memory effects lie at the basis of the anomalous behaviour since it means that there is a negative feedback from moving in any direction. In the particular case of the phantom Rouse polymer this results in a mean squared displacement of a monomer scaling with $t^{1/2}$. Eventually this behaviour breaks down because the diffusion of the polymer itself becomes dominant, meaning that eventually the mean squared displacement starts to scale with t. The order of magnitude of the time when normal diffusion takes over is given by the terminal relaxation time τ of the polymer in question. This kind of behaviour is common for a larger variety of polymers and the parameters corresponding to the different models can be found in [5]. Superdiffusion can be realised by so called Lévy flights, being random walks where the step size is heavy-tailed. Another system where superdiffusion occurs is in the growth of admixture clouds in a turbulent

11

atmosphere. This growth scales with $t^{3/2}$, this was discovered by Richardson in 1926 [6]. This phenomenon is explained by considering the diffusion coefficient to be a system size dependent quantity, namely $D \propto r^{4/3}$ [7]. Given the existence of these systems that have anomalous diffusive behaviour one could ask what properties of normal diffusion are conserved for these systems. In particular the derivation of the Einstein relation relied heavily on Fick's law and it is thus not a priori clear that it should hold in these systems. As it turns out its validity fails in some systems [8], even though many systems with anomalous behaviour do satisfy the equation. The exclusion process in one dimension is an example of an anomalous system for which the Einstein relation holds. This process can be described by multiple random walks on $\mathbb{Z}$ where the particles cannot occupy the same position or pass each other. Intuitively we can already see that the motion of a tagged particle in this system could well be anomalous. If a particle moved a distance δx to the right then it seems more likely that the next particle to its right is closer then the closest particle to its left. Since the particles cannot pass each other this causes a negative feedback on the motion of the particle. The behaviour of a tagged particle in this system is treated extensively in the literature. The most interesting result is the proof that the probability distribution of $\frac{X_t - X_0}{t^{1/4}}$ converges, for one of the simplest varieties where particles jump with rate 1 per unit of time, to a Gaussian with variance [10]

$$\frac{1}{1-\alpha}\sqrt{\frac{2}{\pi}}. \tag{24}$$

Here α is a parameter that gives the probability for any of the sites to be occupied in the original distribution. Since the variance in the position of the particle is immediately related to the mean squared displacement of the particle we obtain an expression for the diffusion coefficient in this system. In this system we would expect that on very short time scales, until the memory effects kick in, the mean squared displacement scales with t rather then $t^{1/2}$. This is confirmed in research on the short term behaviour of this system [9]. In [9] the short term behaviour is investigated in detail and a Taylor expansion of the mean squared displacement is given up to fifth order in time. The subject of the next chapters will now be to discuss the results in [11], showing that the Einstein relation indeed holds in this system.

2 Introduction to Markov processes

The problem of the validity of the Einstein relation has been rigorously solved for a variety of particle systems, amongst those is the exclusion process. Since this text aspires to discuss this proof, some of the necessary background has to be discussed. The first goal will be to get the tools needed to precisely define such discrete particle systems. The most natural way to precisely define models like the exclusion process is in the form of a stochastic process.

Definition 2.1. *A stochastic process is a collection of random variables $\{X_j, j \in J\}$ indexed by a set J, defined on a probability space $(\Omega, \mathcal{F}, P)$.*

The probability space that will eventually be used is a set of functions, where each function will sent a time parameter t to the state of the system at time t. We can then let the random variables correspond to the relevant aspects of the state the system is in, while the index set is our time line. The choice of sigma-algebra is less relevant to the interpretation of the system, but it will require some thought in order to have a well defined process. In this chapter we will mainly look at the theory of Markov processes, since this lies at the basis of the remainder of this text. The first step in this direction is considering one very interesting property of the exclusion process, it lacks a memory of what happened in the past. To clarify, if we know the current state of the system, then the past of our system is irrelevant for its future evolution. This property, often called the Markov property, leads us to the subset of stochastic processes called Markov processes. In this chapter we will introduce several such processes of increasing complexity, this will then lead towards the definition of a Markov process. This approach has the advantage that the definition, that otherwise might seem to have a lot of cumbersome details, looks a lot more natural. In using this approach we follow [16] closely and the theorems in this chapter can be found, sometimes in a slightly more general form, in this book.

2.1 The simple random walk

To get a feeling for what we are working with, let us consider a basic example of a stochastic process that is somewhat similar to the exclusion process. The example is the discrete-time, symmetric random walk on $\mathbb{Z}$. Let $\{X_i : i \in \mathbb{N}\}$ denote a collection of independent random variables with a probability distribution given by $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}$. Then define $S_n = \sum_{i=1}^{n} X_i$ and $S_0 = 0$. The collection $\{S_n : n \in \mathbb{N}_0\}$ forms a stochastic process on the space $\mathbb{Z}$. One interpretation of S_n is that it denotes the position of a particle after n jumps on a one-dimensional lattice. Since this system is closely related to the exclusion process it is worthwhile to further investigate some of its properties. First we can derive its distribution since it follows from symmetry and the following expression

$$\mathbb{P}(|S_n| = k) = \binom{n}{k} \frac{1}{2^n}$$

As expected this process has normal diffusive behaviour at large scales. To see this assume $n \gg k$. Since, for large n, the standard deviation is relatively small, the probability distribution is dominated by values around its mean. These values correspond with a k much smaller then n so the assumption is reasonable. Making this assumption then leads to the approximation

$$\mathbb{P}(|S_n| = k) \approx \sqrt{\frac{2}{\pi n}} e^{-k^2/2n}.$$

Now one can choose a step size and the time between two steps, denote them by x_0 and t_0 respectively. The nett displacement of the particle is given by $x = x_0 k$ and the time it took to get there is $t = n t_0$. This gives us an expression for the position distribution of the particle at time t in a continuum limit:

$$p_t(x) \approx \frac{1}{\sqrt{4\pi Dt}} e^{-x^2/4Dt}.$$

Here $D = \frac{x_0^2}{2t_0}$ equals the diffusion coefficient. This derivation is in no way rigorous, however it does provide an idea of how one could relate a discrete particle system to a continuous probability density. In the next chapter a precise definition of what is meant by the continuum limit of such a system will be given. Afterwards a technique will be discussed to define and prove continuum limits of a discrete particle system. As mentioned before the exclusion process obeys the Markov property, so does the example above. This can be made more precise by looking at the interdependence of elements of the set $\{S_n\}$. Suppose we know the value of S_n then clearly S_{n+1} is either equal to $S_n + 1$ or $S_n - 1$ with equal probabilities. If instead we are given a more general distribution for S_n, then the distribution of S_{n+1} still follows. Just assign for each point of $\mathbb{Z}$ the probability to be occupied based on the probability its neighbors were occupied. So the distributions of the random variables $\{S_m : m < n\}$ are irrelevant for the distribution of S_{n+1}, if we know the distribution of S_n. We can immediately generalise it for stochastic processes of the form $\{X_n : n \in \mathbb{N}_0\}$ on a countable probability space Ω, it can then be formulated as

$$\mathbb{P}(X_{n+1} = k_{n+1} : X_n = k_n, ..., X_0 = k_0) = \mathbb{P}(X_{n+1} = k_{n+1} : X_n = k_n) \quad (25)$$

where $k_0, ..., k_{n+1} \in \Omega$. One way to work with this kind of process, sometimes suitable for simple chains, is by writing each step the system takes as a matrix acting on a vector. This can be done if for all states $x, y \in \Omega$ the transition is time homogeneous (independent of n). The system then gives rise to a well defined matrix given by

$$p(x, y) = \mathbb{P}(X_{n+1} = y : X_n = x).$$

Now define vectors $\vec{a}^{(i)} = (\mathbb{P}(X_i = x_1), ..., \mathbb{P}(X_i = x_n))$. The vector $\vec{a}^{(i)}$ then contains all the information about the distribution of X_i. By using the following relation between two subsequent vectors we can express the distribution of X_n

in terms of the distribution of X_i for any $i < n$:

$$\vec{a}_j^{(i)} = \sum_{k=1}^{n} p(x_k, x_j)\vec{a}_k^{(i-1)}.$$

This method however, although straightforward, quickly becomes tedious if the system is large.

2.2   Markov chains

The discrete time, time-homogeneous Markov chain on a countable state space Ω is a generalisation of the example of the previous paragraph. It also brings us one step closer to the exclusion process. Let $p(x,y)$ be a stochastic matrix, meaning $p(x,y) \geq 0$ for all $x, y \in \Omega$ and

$$\sum_{y \in \Omega} p(x,y) = 1$$

Definition 2.2. *A collection $\{X_n, n \in \mathbb{N}_0\}$ of random variables on a probability space $(\Omega, \mathcal{F}, P)$ is a Markov chain with transition probability $p(x,y)$, $x, y \in \Omega$, if for all $n \geq 0$ and $k_0, ..., k_n \in \Omega$ we have*

$$\mathbb{P}(X_{n+1} = k_{n+1} : X_n = k_n, ..., X_0 = k_0) = p(k_n, k_{n+1})$$

An immediate, but nontrivial, problem with this definition is whether a Markov chain exists if a transition matrix and initial state are given. In the case of a chain of finite length $m \geq 0$ on a countable state space S, it can easily be shown that it does. In fact such a process can be constructed explicitly. Define $\Omega = S^{m+1}$, then we can think of the elements $\omega = (s_0, ..., s_m) \in \Omega$ as functions from $\{0, ..., m\}$ to S or, equivalently, as paths through the state space. Given an initial state, $X_0 = x_0$ almost surely, we can write down the probability for any given element of Ω:

$$\mathbb{P}(\omega) = \mathbb{1}_{\{x_0\}}(s_0)p(s_0, s_1)...p(s_{m-1}, s_m).$$

From this we can define a chain $\{X_n : 0 \leq n \leq m\}$ via

$$X_n(\omega) = s_n.$$

A generalization of this approach to an infinite chain is nontrivial, because the probability of most $\omega \in \Omega$ will vanish. In order to avoid this problem we must assign probabilities to smartly chosen subsets of Ω instead of single elements. This boils down to making a smart choice for our sigma-algebra $\mathcal{F}$. Some events we should include in $\mathcal{F}$ are those where we fix the states for the first m steps. These events correspond precisely, in a natural way, to events in the finite chain. Therefore, if we denote these events as $F_k = \{\omega \in \Omega : s_0 = k_0, ..., s_m = k_m\}$, with $\{k_0, ..., k_m\} \in S$, we can write down the probabilities explicitly. The sets $\{F_k\}$ are called cylinder sets and based on what we found before:

$$\mathbb{P}(F_k) = \mathbb{1}_{\{x_0\}}(k_0)p(k_0, k_1)...p(k_{m-1}, k_m).$$

The question now is whether this is sufficient to get the sigma-algebra and probability measure that we want. The answer follows from Kolmogorov's extension theorem.

Theorem 2.1. *Suppose S is a complete separable metric space, $I = \{i_1, i_2, i_3, ...\}$ a countable index set, and $\Omega = S^I$, the space of functions ω from I into S. Let $\mathcal{F}$ be the product sigma-algebra on Ω, which is by definition the smallest sigma-algebra that contains all sets of the type $\{\omega : \omega(i) \in B\}$ for Borel sets $B \subset S$ and $i \in I$. Suppose that for each n we are given a probability measure μ^n on the space S^n. Assume that the collection μ^n is consistent in this sense: for each n and Borel set $A \subset S^n$,*

$$\mu_{n+1}(\{(s_1, ..., s_{n+1}) \in S^{n+1} : (s_1, ..., s_n) \in A\}) = \mu_n(\{(s_1, ..., s_n) \in A\})$$

Then there exists a probability measure P on $(\Omega, \mathcal{F})$ such that for all finite n,

$$\mu_n(A) = \mathbb{P}(\{\omega \in \Omega : (\omega(i_1), ..., \omega(i_n)) \in A\})$$

This theorem guarantees that we can define a probability measure $\mathbb{P}^x$ on Ω leading to a chain with initial condition $X_0 = x$ a.s. and transition probabilities $p(x, y)$. If we want a more general starting condition where X_0 has a probability distribution μ we can define a probability measure on Ω via

$$\mathbb{P}^\mu(A) = \sum_{x \in S} \mu(x) \mathbb{P}^x(A)$$

for all $A \in \mathcal{F}$.

2.3 The continuous time Markov chain

Until now our attention was restricted to discrete time Markov chains, however this restriction is too strong for the exclusion process as we described it. Instead we want that particles can jump at any moment in time. So the aim is to construct a stochastic process $\{X_t : t \geq 0\}$ taking values in a countable state space S that reflects this behaviour. More precisely we require that there must exist a countable collection of times t_i at which the system jumps from one state to the next. In between those times X_t must be a constant, also the time intervals (t_i, t_{i+1}) should have independent negative exponentially distributed length δt_i. Finally we must have a distribution to determine the state after a jump. This last requirement can, once more, be satisfied by using a stochastic matrix $p(x, y)$ with $x, y \in S$. Once again we want this matrix to be time-independent. The jump rate of the process could depend on the state but for the exclusion process this dependence can be dropped as well. One of the most intuitive ways to construct this process is by taking a discrete time Markov Chain, as we had before, and a collection of appropriately determined times $\{t_i\}$ when the particle should jump. Then define the new process to take the same values as the discrete process in between jumps and to take the ith jump at t_i. Let us try to do this in a more precise manner. Let S be our state

space and $x \in S$ a given initial state. Denote by $p(x, y)$ a stochastic matrix describing transitions in S and let R be the rate at which the system jumps. Now let $\{Y_n : n \in \mathbb{N}_0\}$ be a discrete time Markov chain on the probability space $(\Omega, \mathcal{F}, \mathbb{P}^x)$ as we encountered it before. Let $(\tau_j)_{j \in \mathbb{N}_0}$ be a sequence of exponentially distributed i.i.d. random variables with mean $\mathbb{E}\tau_j = 1$. The latter random variables are defined on a probability space that we shall denote as $(\Omega_\tau, \mathcal{F}_\tau, \mathbb{P}_\tau)$. The probability space of the continuous time Markov chain is now given by

$$(\mathbf{\Omega}, \mathbf{F}, \mathbf{P}^x) = (\Omega \times \Omega_\tau, \mathcal{F} \otimes \mathcal{F}_\tau, \mathbb{P}^x \otimes \mathbb{P}_\tau).$$

We can then define the holding times as $\sigma_n = \frac{\tau_n}{R}$ and from these the random variables $T_i = \sum_{j=0}^{i-1} \sigma_j$. This leads to the following definition for X_t:

$$X_t = Y_n \text{ for } T_n \leq t < T_{n+1}, n \in \mathbb{N}_0.$$

One possible complication would be that X_t is ill-defined when T_n does not go to ∞ for $n \to \infty$. However for the constant rate R that we use T_n diverges almost surely. The memorylessness of this process can be formulated as follows. Let $p_t(x, y) = \mathbf{P}^x(X_t = y)$, then for any collection of times $0 \leq t_0 < t_1 < ... < t_n$ and states $x_0, x_1, ..., x_n$ it can be proven that

$$\mathbf{P}^x(X_{t_0} = x_0, ..., X_{t_n} = x_n) = p_{t_0}(x, x_0) p_{t_1 - t_0}(x_0, x_1) ... p_{t_n - t_{n-1}}(x_{n-1}, x_n).$$

From this we immediately get the Markov property in the following form:

$$\mathbf{P}^x(X_{t_n} = x_n : X_{t_0} = x_0, ..., X_{t_{n-1}} = x_{n-1}) = p_{t_n - t_{n-1}}(x_{n-1}, x_n),$$

whenever this is well defined.

2.4 Skorokhod spaces

The construction of a continuous time Markov chain is rather intuitive once you know how to construct a discrete time Markov chain. However, by using this intuition, a difference in notation between the two processes was introduced. For the discrete time chain we defined a probability space $\Omega = S^{\mathbb{N}_0}$, whereas for the continuous time chain we merged two independent stochastic processes and defined our probability space as the product of the two. It turns out that the continuous time chain can be formulated in a similar way, so let us see how this is done. First we need a probability space that is similar to $S^{\mathbb{N}_0}$. Since $S^{\mathbb{N}_0}$ can be seen as the space of functions from $\mathbb{N}_0$ to S, we will consider functions (or paths if you like) from $\mathbb{R}_+$ to S. However we will be putting some continuity restrictions on the functions in this space, to reflect the behaviour of the jumps. Notice that in the construction of the continuous time chain we chose to let a jump work in the following way. If a particle jumps at time t then shortly before that time it will occupy its old position while at t itself it already occupies its next position. Therefore an appropriate restriction would be to take the set of Càdlàg functions (from $\mathbb{R}_+$ to S) as our probability space. Those are the functions that are continuous from the right and whose left limit exists for all
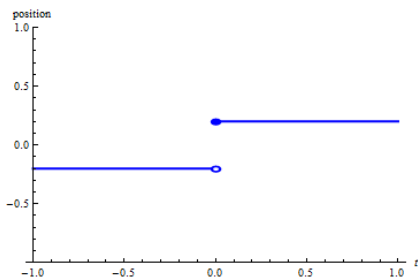
17

Figure 2: An plot of a (part of a) Cádlág function, this particular example is constant in between jumps

points in the domain. This reflects jumps the way we defined them before. In figure 2 we see a particular example of such a function, in that case even more adapted to our situation since we also have that the paths of the particles are constants in between jumps. Now denote the set of Càdlàg functions from $\mathbb{R}_+$ to S by $D(\mathbb{R}_+, S)$, the sigma-algebra $\mathcal{F}_S$ on this set will be the one generated by the coordinate mappings that send $\xi \in D(\mathbb{R}_+, S)$ to its value at time t, $\xi(t)$. So $\mathcal{F}$ is the smallest sigma-algebra such that the coordinate maps are measurable. Now let $X_. = (X_t : 0 \leq t < \infty)$ be a D_S valued random variable, then we get the measure of the probability space by $\mathbb{P}^x(A) = \mathbf{P}^x(X_. \in A)$, for all $A \in \mathcal{F}$. Here the measure $\mathbf{P}^x$ is the one we used before to describe the continuous time Markov process with initial state x. For all $x \in S$ we will denote the expectation of a random variable Y under $\mathbb{P}^x$ by $\mathbb{E}^x Y$. Adjusting to this description we will also introduce a different notation for the Markov property. First define the transition probabilities as

$$p_t(x, y) = \mathbb{P}^x(\xi(t) = y),$$

with $x, y \in S$, $t \in \mathbb{R}_+$ and $\xi \in D(\mathbb{R}_+, S)$. Now we can define a set of useful maps by $(\tau_t \xi)(s) = \xi(t+s)$, they are called shift maps. On an event these shifts act naturally via

$$\tau_t(A) = \{\xi \in D(\mathbb{R}_+, S) : \tau_t \xi \in A\}.$$

This allows a very intuitive expression of the Markov property. First recall that in the discrete case we introduced a sigma-algebra generated by cylinder events, the continuous time counterpart is the sigma-algebra $\mathcal{F}_t$ generated by $\{\xi(s) : 0 \leq s \leq t\}$. The events in the generating set constrain only what happens up to time t. In this notation the Markov property becomes

$$\mathbb{P}^x(\theta_t^{-1}A : \mathcal{F}_t)(\xi) = \mathbb{P}^{\xi(t)}(A),$$

for all $x \in S$, $A \in \mathcal{F}$ and $\mathbb{P}^x-$almost every $\xi \in D(\mathbb{R}_+, S)$. The left hand side can be read as the probability that, given what happened up to time t, A happens from that time onward. The right hand side can then be interpreted as the probability that A happens when starting the process in the point $\xi(t)$. So

18

when the process is at time t in a certain state $\xi(t)$ we can treat it as a new process starting in $\xi(t)$. Later on convergence of measures on this space will be needed so the topology matters. The usual topology on the set of Càdlàg functions in this context is the Skorokhod topology. Equipped with this topology the space is called a Skorokhod space. This is a metrizable and separable topology, intuitively two paths remain close if space and time are only changed a little. This idea is illustrated in figure 3 where the blue and red functions are
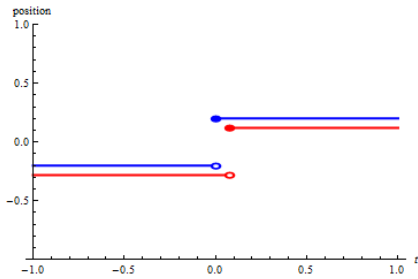

Figure 3: Two Cádlág functions can be close even if they are not at all times spatially close

closer then they would have been under the supremum norm since a relatively small variation in time will make the spatial difference between the two functions everywhere small. Of particular importance are Skorokhod spaces with functions defined on $[0, T]$, these will be used to study the behaviour of the relevant processes for finite times. Now we can define a norm on this space that behaves in the way described above. Let $D([0, T], M)$ be a Skorokhod space, where M is a metric space. Then let Λ denote the space of strictly increasing continuous functions $f : [0, T] \to [0, T]$. Then let $||f|| = \sup_{t \in [0,T]} |f(t)|$ be the uniform norm on $C([0, T], M)$. The Skorokhod metric is now given by

$$d(f, g) = \inf_{\lambda \in \Lambda} \max\{||\lambda - \mathbf{1}||, ||f - g \circ \lambda||\}. \qquad (26)$$

The function λ corresponds to the spatial variation and the element $||\lambda - \mathbf{1}||$ measures how much time is changed. The element $||f - g \circ \lambda||$ is almost the distance between the two functions under the supremum norm, except that the distance can be decreased by reparametrizing time. Balancing these two results in the norm that we want. Some important properties can be found in the third lecture of [14]. In particular the convergence of measures on the Skorokhod space is of importance in the theory. Most of this however is used in lemma's of which the proofs will be omitted in this text.

2.5 Markov processes

The definition of a Markov process can be stated in a single notation, such that it covers both the discrete and continuous time examples we have seen. This notation can easily be used to cover quite a large collection of Stochastic

processes that all share the properties we expect a Markov chain to have. Let Y be a separable metric space and $D(\mathbb{R}_+, Y)$ the space of Càdlàg functions ω from $\mathbb{R}_+$ into Y. Since Y is a metric space it is a topological space and therefore we can equip Y with the Borel sigma-algebra. $D(\mathbb{R}_+, Y)$ also has a separable metric and we can therefore define $\mathcal{F}$ to be the Borel sigma-algebra. Let $X_. = (X_t : t \geq 0)$ be the process defined by $X_t(\omega) = \omega(t)$, with $\omega \in D_Y$. As before let $\mathcal{F}_t$ denote the sigma-algebra generated by $\{X_s : 0 \leq s \leq t\}$. For all $t \geq 0$ define $\tau_t : D(\mathbb{R}_+, Y) \to D(\mathbb{R}_+, Y)$ as the shift map $(\tau_t \omega)(s) = \omega(s+t)$.

Definition 2.3. *A Markov process is a collection of probability measures $\{\mathbb{P}^x : x \in Y\}$ on $D(\mathbb{R}_+, Y)$ such that*

1. *$P^x(\{\omega \in D(\mathbb{R}_+, Y) : \omega(0) = x\}) = 1$*

2. *For any $A \in \mathcal{F}$ the map $x \to \mathbb{P}^x(A)$ is measurable on Y*

3. *$\mathbb{P}^x(\theta_t^{-1} A : \mathcal{F}_t)(\omega) = \mathbb{P}^{\omega(t)}(A)$ for $\mathbb{P}^x-$almost every ω, for all $x \in Y$ and $A \in \mathcal{F}$*

The first requirement ensures we have proper initial conditions. The second requirement basically guarantees that the process behaves nicely. It is not something we checked for the previously given processes, but it does hold in those cases. The third requirement corresponds to the Markov property. Some nice properties can be given immediately. Firstly, we can choose a much more general starting condition. Given a probability measure μ on Y we can use it as the initial condition for a Markov process on D_Y. In this case we can put the following probability measure on the probability space:

$$\mathbb{P}^\mu(A) = \int_Y \mathbb{P}^x(A)\mu(dx),$$

with $A \in \mathcal{F}$. Secondly we can define transition probabilities $p(t, x, B) = \mathbb{P}^x(X_t \in B)$ for $t \geq 0$, $x \in Y$ and B a Borel subset of Y. As a consequence of the Markov property these transition probabilities satisfy the Chapman-Kolmogorov equations

$$p(s+t, x, B) = \int_Y p(s, y, B)p(t, x, dy).$$

In words this says that the probability of going from x to the set B in time $s + t$ equals the probability of going from x to y in time t and then from y to B in time s integrated over all $y \in Y$. Clearly a property we would expect of a transition probability.

2.6 Generators and semigroups

Now that we have the basic framework to describe the exclusion process we will introduce yet another way to describe a Markov process. It turns out that it is often more practical to write down the infinitesimal behaviour of a Markov process then to directly use a description in terms of the above probability

measures. This is done by looking at the expectation of a function on the state space as a function of time. Assume we have a countable state space S, let f be a bounded function on S and let R be a fixed jump rate. We define the generator of a Markov process as the linear operator L satisfying

$$(Lf)(x) = R \sum_{y \in S} p(x, y)(f(x) - f(y)).$$

Of course one can generalise to position dependent jump rates, something that will be needed later. This operator corresponds to the description given above because of the following theorem.

Theorem 2.2. *Assume the rate $R > 0$, $\xi \in D_S$ and let f be a bounded function on S. Then the following continuity property holds*

$$\lim_{t \to 0} \sup_{x \in S} |\mathbb{E}^x[f(\xi(t))] - f(x)| = 0$$

Furthermore L acts like a differentiation operator

$$\lim_{t \to 0} \sup_{x \in S} \left| \frac{\mathbb{E}^x[f(\xi(t))] - f(x)}{t} - Lf(x) \right| = 0$$

This theorem links L to another useful operator, or rather a collection of operators. Given a bounded measurable function f on S we define the operator $S(t)$, for $t \geq 0$ as

$$S(t)f(x) = \mathbb{E}^x[f(X_t)].$$

This definition can equivalently be written as

$$S(t)f(x) = \int_Y f(y)p(t, x, dy).$$

The last equality shows that, for all $t \geq 0$, $S(t)$ acts linearly on functions in its domain. This set of operators satisfies for $s, t \geq 0$ and $x \in Y$

$$S(t + s)f(x) = S(t)S(s)f(x).$$

This identity, together with $S(0) = \mathbb{1}$ makes this set a semigroup. Since $||S(t)f||_\infty \leq ||f||_\infty$, for all f in the domain of $S(t)$, the operators are contractions. The correspondence with the infinitesimal generator of a Markov process follows from the second identity in theorem 2.2 which can be written as

$$\lim_{t \to 0} \sup_{x \in S} \left| \frac{(S(t)f)(x) - f(x)}{t} - Lf(x) \right| = 0.$$

A direct correspondence with the Markov process can also be found. The probability measures P^x are uniquely determined by $S(t)$. This follows because, given $\{S(t)\}$, the probabilities $p(t, x, dy)$ are uniquely determined. From these probabilities the finite dimensional distributions of measures P^x can be computed by repeated integration. That is, for any $n \in \mathbb{N}_0$ we can obtain the quantity

$$\mathbb{E}^x \Phi(X_{t_1}, X_{t_2}, ..., X_{t_n})$$

for any bounded measurable function Φ on Y^n and any $0 \leq t_1 < t_2 < ... < t_n$. These finite dimensional distributions of $\mathbb{P}^x$ can be uniquely extended to the whole sigma-algebra of D_Y, as an application of the so called π-λ-theorem. Many useful theorems regarding these semi-groups and generators can be found in [16]. In this chapter we will only discuss its action on measures. The above analysis only holds if, for a given generator, a corresponding Markov process actually exists. This is not the subject of this text and whenever a Markov process is introduced its existence will be assumed. Moving on, we will restrict ourselves to a relatively nice subset of Markov processes. Denote the space of bounded continuous functions on Y by $C_b(Y)$

Definition 2.4. *A Markov process is a Feller process if $C_b(Y)$ is closed under the action of $S(t)$.*

This translates to transition probabilities as $p(t, x, dx)$ being a weakly continuous function of x for all $t \geq 0$. The exclusion process is an example of a Feller process, so this is an appropriate restriction. For a Feller process we can let the semigroup act on probability measures on Y via its action on functions. Define this action via

$$\int f d[\mu S(t)] = \int S(t) f d\mu$$

for all $f \in C_b(Y)$ and $\mu \in M_1(Y)$, the set of probability measures on Y. This can also be written as

$$\mu S(t)(B) = \int P^x[X_t \in B]\mu(dx)$$

For all Borel measurable subsets B of Y. Note that we therefore can, for all $t \geq 0$, interpret $\mu S(t)$ as the probability distribution of X_t with μ as the initial distribution. Invariant distribution are then characterized by $\mu = \mu S(t)$. We will finish this chapter with an invariance criterion

Theorem 2.3. *Let $D(L)$ be the domain of L, a generator of a strongly continuous contraction semigroup $S(t)$ corresponding to a Markov process X_t. The probability measure μ is invariant for this process iff*

$$\int L f d\mu = 0$$

for all $f \in D(L)$.

In fact we can make the statement stronger, instead of $f \in D(L)$ we can require that the integral vanishes for all f in a core of L. Here a core is defined as a linear subspace of the domain of L, with the property that the graph of L equals the closure of the graph of L restricted to the core. So this adjustment makes the application of this theorem easier by restricting the requirements to a smaller subset of functions. In this text the invariant measures are used to study the long term behaviour of discrete particle systems.

3 Hydrodynamic limits and the entropy method

The proof of the Einstein relation for the exclusion process relies on taking the continuum limit of discrete particle systems. Meaning that the discrete system, in a sense, converges to a system that can be described by a time dependent density profile on the real numbers. With the tools of the previous chapter a rigorous framework can be constructed to let a discrete particle system converge to such a system. After discussing the definitions we will work towards the entropy method, which can be used to prove the hydrodynamic behaviour of a variety of particle systems. The content of this chapter is largely based on three sources, namely [11], [12], [14] and for most of the definitions en theorems in this chapter one can find more information in those texts. This chapter should provide sufficient background to understand the underlying methods for the next, where the formalism introduced here will be applied. Since we will skip many technical aspects of the proof for the Einstein relation we will likewise not go into full detail here. In one aspect this text is, at least for while, more general then needed in that $d \geq 1$ dimensional systems are discussed.

3.1 The zero range process

In this paragraph some of the background for hydrodynamic limits will be discussed, then a new particle system will be defined. That particle system will, in later paragraphs, be used to demonstrate the outline of the entropy method. This system will also play an important part in proving the Einstein relation. The general idea in taking hydrodynamic limits is considering a particle system on a lattice and letting the distance between sites scale as $1/N$, while accelerating time by an N dependent factor, where $N \in \mathbb{N}$. The comparison between a distribution for the particle system and a particle density on a smooth manifold will be made by how the two integrate functions, a notion that will be made precise later. Most of the work will be done on the torus rather then $\mathbb{R}^d$, this has the advantage of easy boundary conditions and a finite number of particles in the discrete systems. Of course any system on the torus corresponds to a periodic system on $\mathbb{R}^d$. Let $\mathbb{T}_N = \mathbb{Z}/n\mathbb{Z}$ be the one dimensional discrete torus with N points and let $\mathbb{T}_N^d = (\mathbb{T}_N)^d$ be its d-dimensional counterpart. Addition and subtraction on the tori will be inherited from addition and subtraction on $\mathbb{Z}$. The discrete systems will often be called the microscopic system, while the smooth torus is called the macroscopic system. Figure 4 illustrates, for the exclusion process, the idea of how particles can be distributed on a torus corresponding to a continuous density function. In this case increasing N would make the discrete system a better and better approximation of the continuous density profile. All of course at a single moment in time. An additional constraint we put on the treated processes is translation invariance of $p(x, y)$, such that $p(y - x) = p(x, y)$ can be considered to be a function of merely one variable $z = y - x$. Furthermore we demand that p is finite and positive and be chosen such that the process is irreducible. Loosely speaking the last requirement says that, if unhindered, any particle can reach any site of the system in a finite
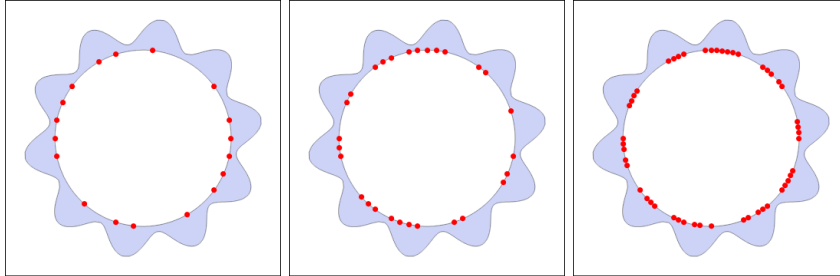
Figure 4: Particles distributed on a torus for different numbers of sites, in each case they are associated to a density function on the torus given by a sinus

number of steps. Note that, at this point, jumps are not restricted to neighboring sites. One particle system with such properties is the zero range process on the torus. Remember that the goal is to obtain information about the limiting behaviour of the exclusion process, however this is not the process of which the hydrodynamic limit will be taken in a usual way. Instead we will do this for the zero range process, a process that can be obtained from the exclusion process by a transformation that will be defined later. Roughly speaking this process describes piles of particles at each point of $\mathbb{T}_N^d$, a particle can jump to another pile at a rate that depends on the number of particles in the original pile. Using the formalism introduced in the previous chapter the zero range model can be defined more precisely. Let $\Omega = \mathbb{N}_0^{\mathbb{T}_N^d}$ be the state space consisting of functions assigning to each vertex of a one dimensional torus a number of particles. Let $p : \mathbb{T}_N^d \to \mathbb{R}$ be the translation-invariant jump probability. If we denote by η_t the configuration at time $t \geq 0$, then the generator of the process is given by

$$(Lf)(\eta) = \sum_{x,y} g(\eta(x))p(y)\left(f(\eta^{x,y}) - f(\eta)\right), \tag{27}$$

where $\eta^{x,y}$ is the state you get from η by moving a single particle from position x to position y. Formally:

$$\eta^{x,y}(z) = \begin{cases} \eta(x) - 1 & \text{when } z = x \\ \eta(y) + 1 & \text{when } z = y \\ \eta(z) & \text{otherwise} \end{cases} \tag{28}$$

The factor $p(y)$ in the generator represents the possible, translation invariant, jumps, while $f(\eta^{x,y}) - f(\eta)$ accounts for the change in f, the remaining factor $g(\eta(x))$ gives the jump rate. That is, given a number of particles at one site there is a rate at which a jump will be made away from that site. The dependence of this jump rate on the number of particles at a site is determined by $g : \mathbb{N}_0 \to \mathbb{R}$. Some restrictions have to be put on g:

1. $|g(k+1) - g(k)| \leq a_0$ for some a_0 and all $k \geq 0$

2. $\sup_{k \geq 0} g(k) \leq a_1$

24

3. $g(k+1) \geq g(k)$ for all $k \geq 0$

Note that the non-interacting zero range process where $g(k) = k$, consisting particles of particles that perform a free random walk with transition matrix p, does not fall under this definition. The zero range process that we get from the aforementioned transformation has $g(k) = \mathbb{1}_{[1,\infty)}$ so for this system these restrictions are fine. When checking the Einstein relation a small force will be applied on the tagged particle. The corresponding adjustment in the zero range process results in a slightly different behaviour for the jump rates at a certain site. For now such anomalies will be ignored. What will be important, however, are the invariant measures of the system. It can easily be checked that the following set of product measures $\{\bar{\nu}_\alpha^N\}_\alpha$ having marginals κ_α given by

$$\kappa_\alpha(k) = \begin{cases} \frac{1}{Z(\alpha)} \frac{\alpha^k}{g(1)...g(k)} & \text{when } k \geq 0 \\ \frac{1}{Z(\alpha)} & \text{when } k = 0 \end{cases} \tag{29}$$

is invariant. If $0 \leq \alpha < \limsup_k g(k)$ the factor $Z(\alpha)$ is finite and uniquely determined by imposing normalisation. In the case of non-interacting particles, $g(k) = k$, these are Poisson measures of parameter α and then $Z(\alpha) = e^\alpha$. In all cases a density can be assigned to these measures by $\rho(\alpha) = \mathbb{E}_{\bar{\nu}_\alpha}(\eta(0))$, since the density is a strictly increasing function of α it is injective. Therefore α can be considered a function of the density ρ allowing for the following definition:

$$\nu_\rho^N = \bar{\nu}_{\alpha(\rho)}^N, \tag{30}$$

being an invariant measure corresponding to an average number of particles ρ at each site. Another set of measures one could consider are those that require the number of particles to be constant. Particle density is after all a conserved quantity (in time) for this system. However if we define

$$\nu_{N,K}(.) \equiv \nu_\alpha^N \left(. \left| \sum_{x \in \mathbb{T}_N^d} \eta(x) = K \right. \right) \tag{31}$$

then the dependence on α vanishes. These measures converge to the measures $\nu_{\alpha(K/N^d)}^N$ in the following sense. For all parameters $B > 0$ and all bounded functions $f : \mathbb{N}_0^{\mathbb{T}_N^d} \to \mathbb{R}$

$$\lim_{N \to \infty} \sup_{0 \leq K \leq BN^d} \left| \mathbb{E}_{\nu_{N,K}}(f) - \mathbb{E}_{\nu_{\alpha(K/N^d)}}(f) \right| = 0. \tag{32}$$

The parameter B can be interpreted as a bound on the density and so we see that the two measures behave similarly as long as they correspond to the same average density. The measures where the particle numbers can change form the so called grand canonical ensemble, while the measures that keep the total number of particles constant form the canonical ensemble. The relation between the two is often called the equivalence of ensembles. This is, in a more general form, an important theorem in statistical physics.

25

3.2 Hydrodynamic limits

In the previous paragraph the zero range process was defined on the torus and some of its invariant measures were discussed. However the invariant measures give trivial dynamics no matter how much time is accelerated, so, in addition to those, more general measures should be discussed. If we were given a particle density function, then at least we would want to be able to define a sequence of measures indexed by N, such that for each N the corresponding measure reflects the macroscopic density profile. Clearly these measures are not, in general, invariant, but they lead to far more interesting dynamics. We would expect that proper rescaling of time, for such a system, results in a change in the density profile over time. First we make precise what it means for a single measure to be associated to a density profile. Then sequences of measures, indexed by N will be considered and finally the system will be made time dependent. Given a density profile $\rho_0 : \mathbb{T}^d \to \mathbb{R}_+$ we would require, for a measure to be associated to ρ_0, that the expected number of particles of a position in the microscopic system is equal to the value of ρ_0 in the corresponding position of the macroscopic system. To this end we define, for a single value of N, a product measure associated to a profile ρ_0 as follows [11].

Definition 3.1. *Given a continuous profile $\rho_0 : \mathbb{T}^d \to \mathbb{R}_+$, we denote by $\nu^N_{\rho_0(.)}$ the product measure on $\mathbb{N}_0^{\mathbb{T}_N^d}$ with marginals given by*

$$\nu^N_{\rho_0(.)}\{\eta : \eta(x) = k\} = \nu^N_{\rho_0(x/N)}\{\eta : \eta(0) = k\} \tag{33}$$

for all $x \in \mathbb{T}_N^d$, $k = 0, 1$. This measure is called the product measure with slowly varying parameter associated to $\rho_0(.)$

Here the right hand side uses the definition for the product measure from the previous with density parameter $\rho_0(x/N)$. The factor $1/N$ corresponds to the way the discrete torus is embedded in the smooth torus. Note that a generalisation for other particle systems is quite natural. The strength of this notion becomes more apparent if we look at the behaviour of sequences of measures $\nu^N_{\rho_0(.)}$ associated with a density profile ρ_0. First we define two functions, let $\Psi : \mathbb{N}_0^{\mathbb{Z}^d} \to \mathbb{R}$ be a cylinder function and associated to it let $\tilde{\Psi} : \mathbb{R}_+ \to \mathbb{R}$ be given by

$$\tilde{\Psi}(\alpha) \equiv E_{\nu_\alpha}(\Psi) = \int \Psi(\eta)\nu_\alpha(d\eta). \tag{34}$$

Let τ_x denote translation by x, meaning $\tau_x\eta(y) = \eta(x + y)$. Naturally we can let this operator act on functions via $(\tau_x\Psi)(\eta) = \Psi(\tau_{-x}\eta)$. If we are now given a density profile $\rho_0 : \mathbb{T}^d \to \mathbb{R}_+$ and a sequence $\nu^N_{\rho_0(.)}$ of measures with slowly varying parameter, each associated to ρ_0, then for all continuous functions

$G : \mathbb{T}^d \to \mathbb{R}$, all bounded cylinder functions Ψ and $\delta > 0$

$$\lim_{N \to \infty} \nu_{\rho_0(.)}^N \left(\left| \frac{1}{N^d} \sum_{x \in \mathbb{T}_N^d} G(x/N)(\tau_x \Psi)(\eta) - \int_{\mathbb{T}^d} G(u)\tilde{\Psi}(\rho_0(u))du \right| > \delta \right) = 0.$$
(35)

The proof of this equation relies on the dominated convergence theorem and Chebyshev's inequality and can be found in [12]. Any sequence of measures satisfying the above equation is said to be a weak local equilibrium of the density ρ_0. This is in fact a stronger requirement than what is needed for a sequence of probability measures to be associated to a profile ρ_0. This latter notion is the one on which we base hydrodynamic limits. A central role in its definition is played by the cylinder function $\Psi(\eta) = \eta(0)$. Motivated by the fact that the product measures $\nu_{\rho_0(.)}^N$ form a weak local equilibrium of ρ_0 the following requirement seems a realistic goal, if not too weak.

Definition 3.2. *A sequence $(\mu^N)_{N \geq 1}$ of probability measures on $\mathbb{N}_0^{\mathbb{T}_N^d}$ is associated to a density profile $\rho_0 : \mathbb{T}^d \to \mathbb{R}_+$ if, for every continuous function $G : \mathbb{T}^d \to \mathbb{R}$, and for every $\delta > 0$, we have*

$$\lim_{N \to \infty} \mu^N \left(\left| N^{-d} \sum_{x \in \mathbb{T}_N^d} G(x/N)\eta(x) - \int_{\mathbb{T}^d} G(u)\rho_0(u)du \right| > \delta \right) = 0 \qquad (36)$$

Since the product measures $\nu_{\rho_0(.)}^N$ are a weak local equilibrium of ρ_0 they are also associated to ρ_0. But it can also be shown directly, because the variance of $N^{-d} \sum_{x \in \mathbb{T}_N^d} \eta(x)$ vanishes for $N \to \infty$. And since the expectation of $\tau_x \eta(0)$ is equal to $\rho(x/N)$ this can be recognized as a Riemann sum. Using Chebyshev's inequality and combining it with an estimate of the expectation of $N^{-d} \sum_{x \in \mathbb{T}_N^d} \eta(x)$ by the corresponding integral, the limit is obtained. So, as was mentioned before, discrete and smooth systems can be compared by looking at how they integrate functions. It is in the above sense that G is integrated in the same way by the discrete system as by $\rho_0 du$. To make this integration on the torus more specific, consider the following set of measures on the torus:

$$\pi^N(\eta, du) = N^{-d} \sum_{x \in \mathbb{T}_N^d} \eta(x)\delta_{x/N}(du) \qquad (37)$$

These are called empirical measures, the advantage of considering this set of measures is that we can work in the space of probability measures on the smooth torus $\mathcal{M}_+(\mathbb{T}^d)$, instead of having a different set of measures for each N. The space $\mathcal{M}_+(\mathbb{T}^d)$ is a topological space equipped with the weak topology. The integral of a continuous function G with respect to the empirical measure is now given by

$$< \pi^N, G >= N^{-d} \sum_{x \in \mathbb{T}_N^d} G(x/N)\eta(x).$$

This concludes the framework for the convergence from a discrete to a smooth system. The final piece of the puzzle is adding time dependence. Doing this naively might very well result in a situation where all dynamics disappears. Rescaling time with an appropriate N-dependent factor θ_N can then restore nontrivial dynamic behaviour. So, starting out with a sequence of measures associated to an initial distribution ρ_0, the goal will be finding a time dependent density function such that for all continuous $G : \mathbb{T}^d \to \mathbb{R}$, the following holds:

$$\lim_{N \to \infty} \mu^N \left(\left| N^{-d} \sum_{x \in \mathbb{T}^d_N} G(x/N) \eta_{t\theta_N}(x) - \int_{\mathbb{T}^d} G(u)\rho(t,u)du \right| > \delta \right) = 0, \quad (38)$$

where ρ is a solution to some partial differential equation with initial condition that $\rho(0,u) = \rho_0(u)$. Often ρ is, a priori, only a solution to a weak differential equation. Therefore we will now give a short introduction to weak solutions. In physics one often finds solutions to equations that are only piecewise differentiable. Therefore it is not surprising that, given a set differential equations, one might allow solutions that do not satisfy these equations in the normal sense. Such a broader set of solutions can be obtained by allowing weak solutions. The idea is based on the notion of partial integration. Consider, for example, the equation $\frac{\partial^2 f}{\partial x^2}(x) = g(x)$ where f and g are real functions. Suppose we would have a solution, then, for all smooth functions ϕ on $\mathbb{R}$ with compact support, we expect the following to hold:

$$\int_{\mathbb{R}} g(x)\phi(x)dx = \int_{\mathbb{R}} \frac{\partial^2 f}{\partial x^2}(x)\phi(x)dx = \int_{\mathbb{R}} f(x)\frac{\partial^2 \phi(x)}{\partial x^2}dx. \quad (39)$$

The boundary terms disappear because the test function, ϕ, has compact support. This is a motivation for defining a weak solution of the above equation as any function f on $\mathbb{R}$ that satisfies

$$\int_{\mathbb{R}} g(x)\phi(x)dx = \int_{\mathbb{R}} f(x)\frac{\partial^2 \phi(x)}{\partial x^2}dx. \quad (40)$$

Now we will discuss the precise definition of a weak solution for the following class of differential equations:

$$\begin{cases} Lu = f & \text{in } U \\ u = 0 & \text{on } \partial U \end{cases} \quad (41)$$

Where L is of the form

$$Lu = -\sum_{i,j=1}^{n} a^{ij}(x)u_{x_i x_j} + \sum_{i=1}^{n} b^i(x)u_{x_i} + c(x)u \quad (42)$$

and U is an open bounded subset of $\mathbb{R}^n$ and $f \in L^2(U)$. The restriction that the solution vanishes on the boundary is not a very strong one in that a transformation can be used to relate a solution for non vanishing boundaries to a solution

of the problem above. This paragraph is largely based on [19] and further details can be found there. Before we continue a little background information about the spaces at hand is needed. In particular we must choose a space in which we look for solutions, this will be a so called Sobolev space. This is the definition of a Sobolev space as it is given in [19]

Definition 3.3. *The Sobolev space $W^{k,p}(U)$ consists of all locally summable functions $u : U \to \mathbb{R}$ such that for each multi-index α with $|\alpha| \leq k$, $D^\alpha u$ exists in the weak sense and belongs to $L^p(U)$*

So for this definition to be useful we need to understand what it means to be a weak derivative, we say that the αth-weak partial derivative of u is given by v if

$$\int_U u D^\alpha \phi dx = (-1)^{|\alpha|} \int_U v\phi dx \qquad (43)$$

for all functions $\phi \in C_c^\infty(U)$. The Sobolev spaces can be equipped with a norm that turns them into a Banach space. Using this topology the boundary conditions can be incorporated naturally by defining $W_0^{k,p}$ as the closure of $C_c^\infty(U)$ in $W^{k,p}$. Finally let $H_0^k(U) \equiv W_0^{k,2}(U)$. The definition of a weak solution to our boundary value problem is now given by

Definition 3.4. *Let the bilinear form associated to the operator L be given by*

$$B(u,v) = \int_U \sum_{i,j=1}^n a^{ij}(\partial_{x_i} u)(\partial_{x_j} v) + \sum_{i=1}^n b^i \partial_{x_i} v + cuvdx \qquad (44)$$

for $u, v \in H_0^1(U)$. We then say that $u \in H_0^1(U)$ is a weak solution of the boundary-value problem (41) if

$$B(u,v) = (f,v) \qquad (45)$$

for all $v \in H_0^1(U)$, where $(.,.)$ denotes the inner product in $L^2(U)$.

Existence and uniqueness theorems for such solutions can be found in [19]. However a final note on the boundary conditions is needed. A boundary value problem of the form

$$\begin{cases} Lu = f & \text{in } U \\ u = g & \text{on } \partial U \end{cases} \qquad (46)$$

where ∂U is a C^1 boundary, can be transformed to a new boundary value problem:

$$\begin{cases} L\hat{u} = \hat{f} & \text{in } U \\ \hat{u} = 0 & \text{on } \partial U \end{cases} \qquad (47)$$

where $\hat{u}$ is related to u via $\hat{u} = u - \omega$ and $\hat{f} = f - L\omega$, with ω a function in H^1 with trace equal to g.

3.3 Entropy and Dirichlet forms

Now, given a particle system, demonstrating the existence of a set of differential equations for ρ such that equation (38) is indeed satisfied by its solution is the main goal. An important method to this end is called the entropy method. It is a method that, for a variety of particle systems, can be used to solve the above problem. Originally it was introduced by Guo, Papanicolaou and Varadhan [18]. In applying the entropy method to the zero range process we will closely follow a set of lecture notes by Sethuraman[14]. Further information about the entropy method can also be found in [11]. In this paragraph some background is provided that is needed to treat the entropy method. First relative entropy is introduced, a quantity from which the entropy method derives its name. Second the Dirichlet form is defined and third we will introduce a particle system, the zero range process, that will be analysed using this method. The definitions and properties of entropy and the Dirichlet form are included in the text as a matter of reference. Their use is largely restricted to a set of (essential) lemma's for which proofs the reader will be directed at other texts. Let us start with the definition of entropy, denote by $H(\mu|\nu)$ the relative entropy of μ with respect to ν, both probability measures on a countable space Ω. This quantity is defined as

$$H(\mu|\nu) = \sup_f \left(E_\mu f - \log(E_\nu(e^f)) \right), \tag{48}$$

where the supremum is over all bounded, measurable functions $f : \Omega \to \mathbb{R}$. If this entropy is finite then it is equal to

$$H(\mu|\nu) = \sum_{x \in \Omega} \mu(x) \log \left(\frac{\mu(x)}{\nu(x)} \right). \tag{49}$$

This quantity has an interesting property in the context of Markov chains. If ν is an invariant measure of a Markov chain on a countable space E and if $\{S(t)\}$ is the semigroup associated to this chain then the following holds.

Theorem 3.1. *For every probability measure μ, we have*

$$H(\mu S(t)|\nu) \leq H(\mu|\nu) \tag{50}$$

If both sides are equal and finite then, for an indecomposable chain, the measures are equal.

This shows that, in a Markov chain, the relative entropy of the measure describing the state of the system with respect to an invariant measure will not increase. This becomes very useful if a relevant quantity can be estimated by relative entropy with respect to an invariant measure, since a bound on the initial relative entropy gives a bound on the relative entropy at all times. This is a feature used in the entropy method.

The second quantity, the Dirichlet form, is one that is relevant at a technical level and will therefore not be as important in this text. However it does occur and its definition will be given as a matter of reference. The Dirichlet form on

$L^2(\pi)$, with π an invariant measure of a continuous time, irreducible Markov process with a finite state space S, is given by

$$D(f) = - < f, LF >_\pi \qquad (51)$$

for all $f \in L^2(\pi)$, where L is the generator of that Markov process. Several results, relevant to the entropy method, about Dirichlet forms and entropy can be found in [12] and [14].

3.4 Guessing the differential equation

With the tools and definitions from the previous paragraphs an attempt can be made to guess the appropriate differential equations for ρ. From this point onward the system will be one dimensional. Let the initial distribution be given by $\nu_{\rho_0(.)}$, such that given N and a site $x \in \mathbb{T}_N$ that site has a number of particles distributed as $\kappa_{\alpha(\rho(x/N))}$. Once again we use the time dependent empirical measure

$$\pi_t^N = \frac{1}{N} \sum_{x \in \mathbb{T}_N} \eta_t(x) \delta_{x/N}. \qquad (52)$$

The strategy that will be used is as follows. First a new martingale, consisting of several terms, will be defined. Next it will be shown that the martingale vanishes for $N \to \infty$. Therefore the terms of the martingale must either vanish or cancel out. Then we will attempt to write these terms as functions of the empirical measure. The result will then consist of three integrals with $\rho(t, u)$ dependent integrands, thereby forming a weak formulation of the corresponding differential equation. Some background for martingales can be found in most textbooks on stochastic processes. If X_t is a Markov process on a countable state space and f is a function on the state space then the following object is a Martingale

$$M_t = f(X_t) - f(X_0) - \int_0^t Lf(x_s)ds. \qquad (53)$$

Therefore, if we accelerate the zero range process by a factor θ_N, and make a specific choice for f, we also get a martingale:

$$M_{\theta_N t}^G = < G, \pi_{\theta_N t}^N > - < G, \pi_0^N > -\theta_N \int_0^t L < G, \pi_{\theta_N s}^N > ds, \qquad (54)$$

where G is, again, a continuous function on the torus. Now an estimate will be made in order to show that the martingale vanishes. Given a martingale M_t its quadratic variation is defined as

$$< M_t > = \lim_{||P|| \to 0} \sum_{k=1}^n (M_{t_k} - M_{t_{k-1}})^2, \qquad (55)$$

where P is a partition of the interval $[0, t]$ and $||P|| = \max\{t_n - t_{n-1}, t_{n-1} - t_{n-2}, ..., t_1 - t_0\}$, the length of the longest interval in the partition. So, in the

limit, the length of all intervals of the partition goes to 0. Now the quadratic variation can be used to put an estimate on the martingale, using the Burkholder-Davis-Gundy inequality. Computing the quadratic variation of the above martingale, using rules derived from the above definition, results in

$$
\begin{aligned}
< M_{\theta_N t}^G > &= \theta_N \int_0^t L(< G, \pi_{\theta_N s}^N >)^2 - 2 < G, \pi_{\theta_N t}^N > L < G, \pi_{\theta_n s}^N > ds \\
&= \frac{\theta_N}{N^4} \sum_{x,y} (\nabla_{x,x+y}^N G)^2 g(\eta(x)) p(y),
\end{aligned}
\tag{56}
$$

where $\nabla_{x,x+y}^N$ is a discrete derivative, defined as

$$
\nabla_{a,b}^N f = N(f(a/N) - f(b/N)).
\tag{57}
$$

The above expression for the quadratic variation is of the order $O(\theta_N N^{-3})$. At this point we will assume that $\theta_N = N^2$, then it follows that the martingale vanishes in the limit $N \to \infty$. The justification of the choice for θ_N follows when the process turns out to have nontrivial dynamics. Now consider the term containing the generator. Using explicit computation one arrives at

$$
\theta_N L < G, \pi_{\theta_N t}^N > = \frac{\theta_N}{N^2} \sum_{x,y} g(\eta(x)) p(y) N \left(G\left(\frac{x+y}{N}\right) - G\left(\frac{x}{N}\right) \right).
\tag{58}
$$

Using a technique called summation by parts, a discrete analogue to integration by parts, this can be reduced to

$$
\frac{\theta_N}{2N^3} \sum_{x,y} g(\eta(x)) p(y) N^2 \left(G\left(\frac{x+y}{N}\right) - 2G\left(\frac{x}{N}\right) + G\left(\frac{x-y}{N}\right) \right).
\tag{59}
$$

The dependence on the state of the process is now contained in the factor $g(\eta(x))$ in each term. Unfortunately this cannot be written as a closed function of the empirical measure. Instead we can try to write it as the sum of a closed expression depending on the empirical measure and a vanishing remainder term. In order to guess the equations governing the hydrodynamic behaviour these estimates will, for now, be made in a non-rigorous way. Let $\eta_{N^2 t}^{N\epsilon}$ be the number of particles in a small cube as a function of the position of the center of the cube:

$$
\eta_{N^2 t}^{N\epsilon}(x) = \frac{1}{2N\epsilon + 1} \sum_{|y-x| \le N\epsilon} \eta_{N^2 t}(x).
\tag{60}
$$

The notation $|y - x| \le N\epsilon$ meaning that the summation is performed over all elements of the torus within in a cube with sides of length $2N\epsilon + 1$. We would expect that $g(\eta(x))$ should be well approximated by the the expected value of $g(\eta(0))$ if we were to take the probability measure ν_ρ^N with $\rho = \eta_{N^2 t}^{n\epsilon}$. This would then mean that, for large N

$$
\frac{1}{2N} \sum_{x,y} g(\eta(x)) p(y) \Delta_{x,x+y}^N G \approx \frac{1}{N} \sum_{x,y} p(y) \Psi(\eta_{N^2 t}^{N\epsilon}(x)) \Delta_{x,x+y}^N G,
\tag{61}
$$

with $\Psi(\rho) = \mathbb{E}_{\nu_\rho^N}(g(\eta(0)))$ and $\Delta_{x,x+y}$ another discrete derivative given by

$$\Delta_{x,y}^N f = N^2 \left(f\left(\frac{x+y}{N}\right) - 2f\left(\frac{x}{N}\right) + f\left(\frac{x-y}{N}\right) \right). \tag{62}$$

Since we can express $\eta_{N^2 t}^{N\epsilon}$, approximately in terms of the empirical measure via

$$\eta_{N^2 t}^{N\epsilon}(x) = <(2\epsilon)^{-1}\mathbb{1}_{[\frac{x}{N}-\epsilon, \frac{x}{N}+\epsilon]}, \pi_{N^2 t}^N> \tag{63}$$

and subsequently, under the assumption that $\pi_{N^2 t}^N$ converges weakly to $\pi(t,u)du$:

$$\eta_{N^2 t}^{N\epsilon} \approx \frac{1}{2\epsilon}\int \mathbb{1}_{[-\epsilon,\epsilon]}(u - x/N)\pi(t,u)du. \tag{64}$$

Combining these results leads to

$$\int_{\mathbb{T}} G(u)\pi(t,u)du = \int_{\mathbb{T}} G(u)\rho_0(u)du + \frac{1}{2}\int_0^t \int_{\mathbb{T}} \Delta G(u)\Psi(\pi(s,u))duds. \tag{65}$$

This is the weak formulation of the following differential equation:

$$\begin{aligned} \partial_t \rho(t,u) &= \Delta\Psi(\rho(t,u)) \\ \rho(0,u) &= \rho_0(u) \end{aligned} \tag{66}$$

3.5   Entropy method for the zero range process

At this point there is a well defined system, namely the zero range model, to work with and a rigorous notion of how a continuum limit of such a system should be taken. We also have an idea about what the differential equation, governing the continuous system, should be. At this point the entropy method can finally be applied to show the correctness of the expected limiting behaviour. In this paragraph this method will be used, however not in its full rigour. The proof of several technical notions will not be discussed, a full treatment can be found in [11] and [12]. Throughout this paragraph it will be made clear whenever such a proof is omitted. Summarizing, the goal is to show the validity of the following theorem:

Theorem 3.2. *We have for all $G \in C^2(\mathbb{T})$ and $\delta > 0$ that*

$$\lim_{N\to\infty} \mathbb{P}_{\nu_{\rho_0}(.)}\left(\left| <G, \pi_{N^2 t}^N> - \int_{\mathbb{T}} G(u)\rho(t,u)du \right| > \delta \right) = 0 \tag{67}$$

where ρ is the unique weak solution of the hydrodynamic equation

$$\partial_t \rho = \frac{1}{2}\Delta\rho \ \text{and} \ \rho(0,u) = \rho_0(u) \tag{68}$$

First a sketch of the relevant steps. The process at hand defines, for each N, a set of probability measures Q^N on $D([0, T], M_+(\mathbb{T}))$. The first step of the proof is showing that the set of measures $\{Q^N\}$ is tight. Tightness in this context means that for all $\epsilon > 0$ there exists a compact subset K of the probability space such that $\inf_N Q^N(K) > 1 - \epsilon$. Here tightness implies relative compactness. This means that each sequence of measures $\{Q^N\}$ has a weakly convergent subsequence. The proof of the tightness of $\{Q^N\}$ is the first technicality that will be omitted. Instead relative compactness will be treated as a given. Then there is the second step concerning properties of the probability densities of the limit points of $\{Q_N\}$. This result will be stated below without proof. Then we will need another lemma, in order to make the treatment of $\eta_{N^2 t}^{N\epsilon}$ in the previous paragraph more precise. The proof of this lemma will also be omitted, however some comments will be made since in this part certain important differences will arise for the asymmetric exclusion process. Then finally based on these three lemma's the proof of the hydrodynamic behaviour can be given. So at this point tightness is assumed to be a given. Then the following lemma gives us an important property of the limit points of Q^N:

Lemma 3.1. *All limit points of $\{Q^N\}$ are supported on trajectories with L^2 densities $\pi(t, u)du$ such that there is a constant $C = C(\bar{\rho})$, where $\bar{\rho} = \sup_{u \in \mathbb{T}} \rho_0(u)$, and*

$$\mathbb{E}_Q \int_0^T \int \pi(t, u)^2 du dt < C \tag{69}$$

In fact the trajectories are absolutely continuous for each time $t \in [0, T]$. This has an interesting interpretation in that it implies that particles cannot pile up in such a way they behave, macroscopically, as a point mass. Finally a replacement theorem is needed, for $\Psi(\rho) = \mathbb{E}_{\nu_\rho}(g(\eta(0)))$ and τ_x the shift operator we have

Lemma 3.2.

$$\limsup_{\epsilon \downarrow 0} \limsup_{N \uparrow \infty} \mathbb{E}_{\nu_{\rho_0}(.)} \left(\int_0^T \frac{1}{N} \sum_{x \in \mathbb{T}_N} \tau_x V_{N\epsilon}(\eta_{N^2 s}) ds \right) = 0 \tag{70}$$

where

$$V_l(\eta) = \left| \frac{1}{(2l + 1)} \sum_{|y| \leq l} g(\eta(y)) - \Psi(\eta^l(0)) \right| \tag{71}$$

The above theorem is used to estimate the martingale equation of the previous paragraph by a closed expression in terms of the empirical measure. Some aspects of this lemma will be discussed after the proof of the theorem.

Proof of Theorem 3.2. Let $G \in C^2([0, T] \times \mathbb{T})$, note that we allow G to depent on time.

$$M_{\theta_N t}^G = < G(t, .), \pi_{\theta_N t}^N > - < G(0, .), \pi_0^N > -\theta_N \int_0^t L < G(s, .), \pi_{\theta_N s}^N > ds \tag{72}$$

34

Using partial summation again the following equality is obtained.

$$M_{N^2 t}^G = < G(t,.), \pi_{\theta_N t}^N > - < G(0,.), \pi_0^N > +$$
$$- \int_0^t \frac{1}{N} \sum_{x \in \mathbb{T}_N} \left(\frac{\partial}{\partial s} G(s, x/N) \eta_{N^2 s}(x) + \frac{1}{2} \Delta G(s, x/N) g(\eta_{N^2 s}(x)) \right) ds + o(1) \tag{73}$$

Due to the assumptions on g and the conservation of particle number the quadratic variation of the martingale vanishes. Using Doob's inequality we can estimate the supremum over $t \in [0,T]$ of the expectations of the martingales squared by the expectation at time T:

$$\sup_{t \in [0,T]} \mathbb{E}_{\nu_{\rho_0(.)}} < M_{N^2 t}^G >^2 \le \mathbb{E}_{\nu_{\rho_0(.)}} < M_{N^2 T}^G >^2 = O(N^{-1}). \tag{74}$$

The problem is now to close the expression in terms of the empirical measure, using the replacement lemma we have the following estimate for the integral term in the martingale

$$\lim_{\epsilon \downarrow 0} \lim_{N \uparrow \infty} \mathbb{E}_{\nu_{\rho_0(.)}} \int_0^T \frac{1}{N} \sum_{x \in \mathbb{T}_N} \Delta G(s, x/N) \left| g(\eta_{N^2 s})(x) - \Psi(\eta^{N\epsilon}(x)) \right| ds = 0, \tag{75}$$

where $\eta^{N\epsilon}(x)$ again denotes the particle density in a small box around x, with the corresponding subscript omitted. Combining this replacement with the vanishing of the martingale we obtain

$$\lim_{\epsilon \downarrow 0} \lim_{N \uparrow \infty} \mathbb{E}_{\nu_{\rho_0(.)}} \big| < G, \pi_{N^2 T} > - < G, \pi_0 > +$$
$$- \int_0^T < \frac{\partial}{\partial s} G, \pi_{N^2 s} > ds - \frac{1}{2} \int_0^T \int_{\mathbb{T}} \Delta G(s, u) \Psi(\pi_{N^2 s}^{(\epsilon)}(u)) du ds \big| = 0, \tag{76}$$

where the following notation was used $\pi_s^{(\epsilon)}(u) = < (2\epsilon)^{-d} \mathbb{1}_{[-\epsilon, \epsilon]}(-u), \pi_s >$. The quantity in between the absolute value brackets is, in the Skorokhod topology, a continuous function. Now Fatou's lemma gives an upper bound to the expectation under a limit point Q of $\{Q^N\}$. Note that the sequence $\{Q^N\}$ takes care of the acceleration of the process and that therefore the factor N^2 can be dropped:

$$\lim_{\epsilon \downarrow 0} \mathbb{E}_Q \big| < G, \pi_T > - < G, \pi_0 > +$$
$$- \int_0^T < \frac{\partial}{\partial s} G, \pi_s > ds - \frac{1}{2} \int_0^T \int_{\mathbb{T}} \Delta G(s, u) \Psi(\pi_s^{(\epsilon)}(u)) du ds \big| = 0. \tag{77}$$

Using lemma (3.1) and the fact that Ψ is bounded and continuous dominated convergence can be used to estimate the above quantity by a closed function of π. The following equality makes the replacement explicit:

$$\lim_{\epsilon \downarrow 0} \mathbb{E}_Q \int_0^T \int_{\mathbb{T}} |\Psi(\pi_s^{(\epsilon)})(u) - \Psi(\pi(u, s))| du ds = 0. \tag{78}$$

35

The conclusion is that Q is supported on trajectories such that

$$
\begin{aligned}
< G, \pi_T >= \\
< G, \pi_0 > +\frac{1}{2} \int_0^T \int_{\mathbb{T}} < G_s, \pi_s > ds + \frac{1}{2} \int_0^T \int_{\mathbb{T}} \Delta G(s, u) \Psi(\pi(s, u)) du ds.
\end{aligned}
\tag{79}
$$

Form the literature it is known that the weak solution with L^2 integrable density, satisfying the known conditions, is unique. So the limit points of $\{Q^N\}$ are supported on trajectories with deterministic densities $\pi(t, u)$ satisfying the above weak differential equation. Now let h_t be a projection operator that sends a trajectory $< \gamma_t >$ in $D([0, T], M_+(\mathbb{T}))$ to $\gamma_t \in M_+(\mathbb{T})$. This projection is not continuous, however it is continuous restricted to the support of limit points Q of $\{Q^N\}$, since these points are supported on continuous trajectories. It now follows that $Q^N \circ h_t^{-1} \to Q \circ h_t^{-1}$ and thus that $\pi_{N^2 t}^N$ converges, in probability, to $\rho(t, u) du$ where ρ is the unique solution to the above differential equation $\quad\square$

3.6 Comments on lemma's of the entropy method

In the previous paragraph the outline of the proof of the hydrodynamic behaviour was discussed, however the concept entropy was not used. In fact it was hidden in the lemma's and the assumptions made for the particular system at hand. It can be shown that, in the zero range process in d dimensions $H(\nu_{\rho_0(.)}^N|\nu_\rho^N) = O(N^d)$. Because ν_ρ^N is an invariant measure the above estimate on the original entropy implies that

$$
H(\nu_{\rho_0(.)}^N S_N(t)|\nu_\rho^N) \leq H(\nu_{\rho_0(.)}^N|\nu_\rho^N) = O(N^d),
\tag{80}
$$

where S_N is the semigroup related to the process accelerated by a factor of N^2. In the proof on the tightness of $\{Q_N\}$ the first inequality can be used. Lemma (3.1) can be broken down in a part that puts estimates on the trajectories in $D([0, T], \mathcal{M}_+)$ at a given time t, assuming an entropy inequality at that time. Then the remainder of the lemma follows from the original entropy inequality and its evolution in time. In order to prove the replacement lemma two rather technical lemma's, called the one block estimate and the two block estimate respectively, have to be proven. It is also in this proof that estimates on relative entropy are used in addition to some estimates on Dirichlet forms. In the adjusted zero range process, to be discussed later, several changes must be made to keep the proof of the replacement lemma valid. It is beyond the scope of this text to give more then rough indications about what changes. For more details the reader could go to [11] for the adjustments and [12] for the background material.

4 The driven tracer particle in the exclusion process

In this chapter we will discuss the behaviour of a tagged particle in the exclusion process when it experiences a small force. This topic is treated in [13], where the validity of the Einstein relation for this system is proven. The content of this chapter is largely derived from that article. First the results, concerning the long term behaviour of the tagged particle, will be introduced and discussed. Then the Einstein relation, as stated in this article, will be treated and it will be shown that this indeed corresponds to the notions introduced in earlier chapters. The underlying proofs, as given in [13], will be postponed to the next chapter, but, as was done for the entropy method, they will not be treated in their entirety. An interesting aspect of the approach that will be used is the introduction of an invertible transformation, changing the exclusion process to a variety of the zero range process. Because of this approach several quantities in this chapter may look unnatural at first, however, in the context of this transformation they make a lot more sense.

4.1 Defining the exclusion process

As was done for the zero range process in the chapter on hydrodynamic limits we should settle on some definitions and notations for the exclusion process. Specifically we should have a notation that makes the existence of a tagged particle experiencing a force apparent. In the exclusion process the particles are distributed on $\mathbb{Z}$, for convenience let us call the direction with increasing values "right" and the other direction "left". Without loss of generality we can now assume that the force acting on the tagged particle is directed to the right. This force can be modelled by saying that an attempt to jump to the right happens with probability $\frac{1}{2} < p \leq 1$, while a jump to the left happens with probability $q = 1 - p$. This can be justified on physical grounds and more specifically by the Arrhenius equation. The argument goes as follows, if we apply a force to a particle that can jump to neighboring sites then this force corresponds to a change in the potential of the system. This then implies a change in jump rate between sites i and j scaling as $e^{-\frac{\Delta E_{ij}}{k_B T}}$, with ΔE_{ij} the change in potential between the sites. In a first order approximation this leads to a change in jump rate with a factor of $1 + \frac{F}{k_B T}$ with F the force in that direction. So approximately the jump rates become $p = \frac{1}{2}(1 + \frac{F}{k_B T})$ and $q = \frac{1}{2}(1 - \frac{F}{k_B T})$. We now need a way to describe the states of the system, define the state space S as a collection of pairs (X, ξ). Here $X \in \mathbb{Z}$ denotes the position of the tagged particle and $\xi \in \{0, 1\}^{\mathbb{Z}}$ will indicate for each site of $\mathbb{Z}$ whether it is occupied by a particle that is not the tracer particle. If it is, then ξ will assign that site a 1 otherwise a 0, note that since ξ does not consider the position of the tagged particle occupied, automatically $\xi(X) = 0$. The probability space is then the set of càdlàg functions from $\mathbb{R}$ to S. We can now define this Markov process via its infinitesimal generator. Let $\xi^{z,z+1}$ denote the state one obtains by interchanging

the values of $\xi(z)$ and $\xi(z+1)$ in ξ, then define the exclusion process via its generator as follows

Definition 4.1. *The exclusion process with right jump probability p is a Markov process with state space $\{0,1\}^{\mathbb{Z}}$ and the following generator*

$$\mathcal{L}f(X,\xi) = \tfrac{1}{2}\sum_{z\neq X-1,X}(f(X,\xi^{z,z+1}) - f(X,\xi))+$$
$$p(1-\xi(X+1))(f(X+1,\xi) - f(X,\xi)) \qquad (81)$$
$$q(1-\xi(X-1))(f(X-1,\xi) - f(X,\xi))$$

with $(X,\xi) \in \{0,1\}^{\mathbb{Z}}$ as described above and $f : \mathbb{Z} \times \{0,1\}^{\mathbb{Z}}$ a local function

The generator in this definition consists of three terms. The first part, $\tfrac{1}{2}\sum_{z\neq X-1,X}(f(X,\xi^{z,z+1}) - f(X,\xi))$, describes the contribution due to jumps of the particles that do not experience any force. The factor $\tfrac{1}{2}$ is the rate at witch particles jump in a both directions. The summation is over all possible jumps except the ones of the tagged particle, while the summand reflects the change in the value of a function due to the interchange of $\xi(z)$ and $\xi(z+1)$. This sum contains a lot of terms where nothing happens, but naturally these vanish since f does not change for such a jump. The second term accounts for jumps of the tagged particle to the right, happening at rate $p(1 - \xi(X+1))$. In this term the rate p is reduced to 0 if the target site is already occupied. Likewise the third term describes jumps to the left.

4.2 The behaviour of the tagged particle

The results discussed in this section are based on taking hydrodynamic limits of the systems at hand. As mentioned before the derivation uses the hydrodynamic limit of a zero range process, this will make the results for the tagged particle look relatively inelegant. In the previous paragraph the exclusion process was defined, we will continue by defining the quantities needed to formulate the main theorems of this chapter. First some relevant measures will be introduced, then two functions will be defined that are actually remnants of the transformation between the two processes. These functions describe how to transform densities between the exclusion process and the zero range process. As such they will appear in the sets of differential equations governing the hydrodynamics of the zero range process and also in the expression for the position of the tagged particle. The differential equations used in this chapter actually describe the particle density in the zero range process rather then the densities of the exclusion process. So let us start with the relevant measures, what we need are candidates for the initial distribution of the exclusion process. Define $\mathbb{Z}_* \equiv \mathbb{Z}-0$ we can then define an initial measure for the non tagged particles on $\{0,1\}^{\mathbb{Z}_*}$. Let $x \in \mathbb{Z}$ and $0 \leq \alpha \leq 1$, then

$$\mu_\alpha(\xi : \xi(x) = 1) = \alpha. \qquad (82)$$

This would be a system where each site has a probability α to be occupied. We can generalise this approach by considering product measures with slowly

varying parameter associated to a density profile $\kappa_0 : \mathbb{R} \to [0, 1]$. For this system this is done by defining

$$\mu_{\kappa_0}^N(\{\xi : \xi(x) = 1\}) = \kappa_0(x/N). \tag{83}$$

Now the measure $\mu_{\kappa_0}^N$ assigns a probability $\kappa_0(x)$ to the site xN being occupied. This of course is similar to how initial conditions were chosen by for the zero range process. Sequences of associated measures will indeed give proper hydrodynamic behaviour initially. The probability measures corresponding to this process will be denoted by $\mathbb{P}_{\mu_{\kappa_0}^N}$ defined on D_S. Aside from the above probability measures we need the functions $H : \mathbb{R} \to \mathbb{R}$ and $F : \mathbb{R} \to \mathbb{R}$ defined as

$$H(x) \equiv \int_0^A \kappa_0(u)du; \tag{84}$$

$$F(y) \equiv \frac{1}{\kappa_0(H^{-1}(y))} - 1. \tag{85}$$

The inverse of H is well-defined since κ_0 is positive on its domain. The last definition we will need for the main results consists of two sets of differential equations. In the first case we will use the following definition

Definition 4.2. *Let* $\rho : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ *be the unique weak solution of*

$$\begin{cases} \partial_t \rho = \frac{1}{2}\Delta\Phi(\rho) \\ \rho(t, 0) = 0 \\ \rho(0, .) = F_+(.) \end{cases} \tag{86}$$

Here F_+ *denotes the restriction of* F *to* $\mathbb{R}_+$ *and* $\Phi(\rho) \equiv \frac{\rho}{1+\rho}$.

The precise meaning of a weak solution to this equation will be discussed in the proof. For now it is sufficient to know that ρ corresponds to the particle density in the zero range process. The function F appears here because, given a density κ_0 of the exclusion process then F is the corresponding density of the zero range process. The restriction to the positive real numbers is made because this equation describes a special case where the tagged particle always moves to the right, that is $p = 1$. In the second case, for $p < 1$ we use another definition:

Definition 4.3. *Let* $\rho : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ *be the unique weak solution of*

$$\begin{cases} \partial_t \rho = \frac{1}{2}\Delta\Phi(\rho) \\ p\Phi(\rho(t, 0+)) = q\Phi(\rho(t, 0-)) \\ \partial_u \Phi(\rho(t, 0+)) = \partial_u \Phi(\rho(t, 0-)) \\ \rho(0, .) = F(.) \end{cases} \tag{87}$$

This set of equations covers the (more relevant) cases where $p < 1$. We can now state the two main theorems as given in [13].

Theorem 4.1. *Assume $p = 1$. Fix a profile $\kappa_0 : \mathbb{R}_+ \to [0,1]$ such that $\sigma \leq \kappa_0 \leq 1 - \sigma$ for some $\sigma > 0$. Then, for every $\delta > 0$,*

$$\lim_{N \to \infty} P_{\mu_{\kappa_0}^N} \left[\left| \frac{X_{tN^2}}{N} - v_t \right| > \delta \right] = 0 \tag{88}$$

where

$$v_t = \int_0^\infty \{F(x) - \rho(t,x)\} \, dx \tag{89}$$

with ρ the solution of equation (86).

This theorem describes the behaviour of the tagged particle on large time scales. As before time is rescaled to preserve dynamics. This case does not correspond to the situation where a small force acts on the particle, however it does provide insight in the behaviour of the exclusion process. One complication in the theorem is the convergence of the integral in (89). For it to make sense it must be defined defined in a slightly unexpected manner. Let the set of real functions $\{H_n : n \geq 1\}$ be defined as

$$H_n(x) = \left(1 - \frac{x}{n} \right)^+. \tag{90}$$

Then the limit for $n \to \infty$ of $\int_0^\infty H_n(x) \{F(x) - \rho(t,x)\} \, dx$ converges and this is what is meant by the right hand side of equation (89). For $p < 1$ a similar results can be derived and is stated below.

Theorem 4.2. *Assume $p < 1$. For $\alpha < 1$ define $\psi_\alpha(x) = \alpha \mathbb{1}\{x < 0\} + (q\alpha/p)\mathbb{1}\{x > 0\}$. Let $\kappa_0 : \mathbb{R} \to [0,1]$ such that $\psi_\alpha \leq \kappa_0 \leq 1 - \sigma$ for some $\sigma > 0$, $0 < \alpha < 1$. Then, for every $\delta > 0$, (88) holds if v_t is given by (89) where ρ is now the solution of differential equation (87) instead of equation (86), that was used in the previous theorem.*

So once again we get a result for $\frac{x_{tN^2}}{N}$ as long as the initial conditions are sufficiently homogeneous. From this we can derive the Einstein relation for this system.

4.3   The Einstein relation

In [13] the Einstein relation is formulated as the third theorem in the following form

Theorem 4.3. *If the initial state, for the case $p < 1$ is given by μ_α, then*

$$\lim_{p-q \to 0} \frac{v_t}{p - q} = \frac{1 - \alpha}{\alpha} \sqrt{\frac{2t}{\pi}} \tag{91}$$

Note that in this case the system is required to start out homogeneously, whereas in the previous theorems the systems were allowed to deviate from homogeneity to a certain extent. Let us look at why this indeed corresponds to

the Einstein relation as it was formulated earlier. The Einstein relation uses the diffusion coefficient of a system when no force is present. For this system the coefficient can be found in the literature. Let μ_α be the distribution for a set of particles in the exclusion process without a force. From [10] we know that, if we choose a tracer particle in this system, then $\frac{X_t}{t^{1/4}}$ has a distribution that converges to a Gaussian distribution with variance $\frac{1-\alpha}{\alpha}\sqrt{2/\pi}$. This expression leads, in our one-dimensional system, to

$$D = \frac{<X_t^2>}{2t} = \frac{1-\alpha}{\alpha}\frac{1}{\sqrt{2\pi t}}. \tag{92}$$

Now the other quantity we need is mobility, let us denote it by μ. It is defined as

$$\mu = \lim_{F\downarrow 0}\frac{V(F,t)}{F}$$

Where V is the velocity of the particle as a function of the force F and time t. In our lattice model we should interpret velocity as a derivative to time of the expected position of the tracer particle: $<V(F,t)> = \frac{\partial <X_t>}{\partial t}$. Now the theorems of the previous paragraph come in handy. From these it can be derived, by solving equation (87), that for a homogeneous initial state $\lim_{t\to\infty}\frac{X_t}{\sqrt{t}} = v$, with v a constant. From this limit we see that for $t \gg 0$ we get the following approximation for V:

$$V = \frac{\partial <X_t>}{\partial t} = \lim_{s\to 0}\frac{<X_{t+s}> - <X_t>}{s} = \lim_{s\to 0}\frac{\sqrt{t+s}-\sqrt{t}}{s}v = \frac{1}{2\sqrt{t}}v. \tag{93}$$

And so, using the relation between p and F, namely that up to first order in F the probability to jump to the right is given by $p = \frac{1}{2} + \frac{F}{2k_B T}$, we obtain

$$\lim_{p-q\to 0}\frac{v_t}{p-q} = \lim_{F\to 0}k_B T\frac{v_t}{F} = 2\mu k_B T\sqrt{t} \tag{94}$$

Using these results in theorem 4.3, we can rewrite that equality as

$$D(t) = \mu(t)k_B T. \tag{95}$$

This approximation is good for large times and small forces and thus the Einstein relation holds in this system. What remains to be shown is the validity of the theorems describing the position of the tagged particle and to solve the differential equations for the situation where a small force is applied.

5 The structure of the proof for the $p = 1$ case

In the previous chapter some strong statements were made about the behaviour of the tagged particle, in this chapter the focus will lie on proving their validity. First the proof of Theorem 4.1 will be discussed, then a few comments will be made about the proof of theorem 4.2 and how it is different. As mentioned several times a transformation will be defined, changing the the exclusion process to a variety of the zero range process. Then, for this system, we can prove the existence of a hydrodynamic limit and this result can then be transformed back to prove the theorems.

5.1 Transforming to the zero range process

The transformation can best be described by an invertible map

$$\mathcal{T} : \mathbb{Z} \times \{0, 1\}^{\mathbb{Z}} \to \mathbb{N}_0^{\mathbb{Z}}. \tag{96}$$

This map sends states of the exclusion process to states of the zero range process. Let (X, ξ) be a state of the exclusion process, we can then attach numbers to each particle in the following way. The tagged particle is given a 0, the xth particle to its right is labelled by x and the yth particle to its left is labelled by $-y$. We can now define $\eta(x)$ to be the number of empty sites between particle x and particle $x + 1$. So given a state in the exclusion process we can assign to each integer x a natural number corresponding to the number of holes between particles x and $x + 1$ in the exclusion process. The dynamics of this process follow from the dynamics of the exclusion process. We will prove that we do obtain the zero range process using the notation of [13].
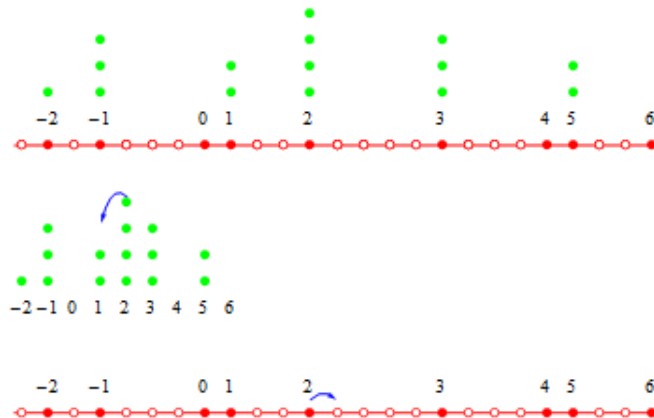

Figure 5: An illustration of how the exclusion process transforms into the zero range process. The first figure shows how to assign particles to each number, the second then illustrates the connection in the dynamics of the two systems.

Let δ_x denote the state with no particles except for one at position x. Furthermore let subtraction on states be defined by $(\eta_1 - \eta_2)(x) = \eta_1(x) - \eta_2(x)$ for any two states η_1, η_2 for which this is nonnegative for all $x \in \mathbb{N}_0$. Addition can be defined similarly, the only difference is that the nonnegativity constraint becomes trivial. Then let $\sigma^{x,y}\eta = \eta - \delta_x + \delta_y$ be the operator that takes one particle from position x to position y. Denote by $L_{x,y}$ the following operator on bounded functions on the state space:

$$L_{x,y}h = \frac{1}{2}g(\eta(x))(h(\sigma^{x,y}\eta) - h(\eta)). \tag{97}$$

The following lemma gives the desired property of the transformation

Lemma 5.1. *The transformation $\mathcal{T}$ defined above transforms the exclusion process with generator (81) into a zero range process with generator (103) given below.*

Proof. It is trivial that the transformation maps a state (X, ξ) in the state space of the exclusion process into a state η of the zero range process. In order to show that the dynamics of this new process is, up to a slight adjustment, the dynamics of the zero range process, as we encountered it before, we will look at the generators. First note that for any $z \neq -1, 0$ we have that, if either z or $z + 1$ is occupied, the state $\mathcal{T}(X, \xi^{z,z+1})$ corresponds to

$$g(\eta(x))\sigma^{x,x+1}\eta + g(\eta(x+1))\sigma^{x+1,x}\eta \tag{98}$$

where x and $x + 1$ are the labels of the gaps to the left and right of the jumping particle respectively. If neither z nor $z + 1$ is occupied there is no contribution in the generator so no corresponding value of x has to be found. If both z and $z+1$ are occupied then the contribution to the original generator vanishes, while the terms $g(\eta(x))\sigma^{x,x+1}\eta + g(\eta(x+1))\sigma^{x+1,x}\eta$ cancel each other. As such the summation

$$\frac{1}{2}\sum_{z \neq X-1, X}(f(X, \xi^{z,z+1}) - f(X, \xi)) \tag{99}$$

in the generator of the exclusion process results under the transformation, in the following contribution to the generator of the zero range process:

$$\sum_{z \neq -1}(L_{z,z+1} + L_{z+1,z}). \tag{100}$$

Note that only $z = -1$ is excluded because both $\sigma^{-1,0}\eta$ and $\sigma^{0,-1}\eta$ would only occur in that term. The behaviour of the tagged particle, given by

$$p(1-\xi(X+1))(f(X+1, \xi)-f(X,\xi))+q(1-\xi(X-1))(f(X-1,\xi)-f(X,\xi)) \tag{101}$$

similarly translates into

$$2pL_{0,-1} + 2qL_{-1,0}. \tag{102}$$

This then results in the full generator being given by

$$L = 2pL_{0,-1} + 2qL_{-1,0} + \sum_{z \neq -1} (L_{z,z+1} + L_{z+1,z}). \tag{103}$$

This, with $g = \mathbb{1}_{[1,\infty)}$, is indeed a generator of the zero range process as we encountered it before, except for slightly different behaviour in the origin. $\square$

The anomaly in the origin is the main difficulty in determining the hydro-dynamic limit of this process. However we do have a nice feature that helps in recovering the limiting behaviour of the tagged particle. As it turns out it is possible to give an expression for X_t, the position of the tagged particle, in terms of the states η of the zero range process at times 0 and t:

$$X_t = \sum_{x \geq 0} (\eta_0(x) - \eta_t(x)) \tag{104}$$

This equality will turn out to be a crucial step in translating the results for the zero range process in results for the exclusion process. Other important quantities of the two process that can be related to each other are the sets of distributions of these processes. For the exclusion process consider a product Bernoulli measure μ_α, with parameter α. The probability for the length of the gap between particle x and $x + 1$ to be k is given by $p_k = \alpha(1 - \alpha)^k$. The expected length of the gap is then given by $\beta = \frac{1}{\alpha} - 1$, since β corresponds to the particle density in the zero range process it seems to be a natural parameter to use for probability measures in this process. The product measure for the zero range process corresponding to the homogeneous distribution in the exclusion process has marginals given by

$$\nu_\beta^+(\{\eta : \eta(x) = k\}) = \frac{1}{1 + \beta} \left(\frac{\beta}{1 + \beta} \right)^k. \tag{105}$$

The expectation of $\eta(x)$ indeed equals β, also this measure is reversible with respect to $L_{x,x+1} + L_{x+1,x}$. The latter condition was of course expected, because in the absence of a force on the tagged particle, the homogeneous distribution is invariant for the exclusion process. The main theorems that we wish to prove have weaker restrictions on their initial measures then for them to be homogeneous. In the case of theorem 4.1 the constraint is only that the initial distribution consists of a sequence of Bernoulli product measures associated to a density profile $\sigma \leq k_0 \leq 1 - \sigma$, for some $\sigma > 0$. Ideally this translates into a sequence of measures in the corresponding zero range processes that have a nice limiting behaviour. For this theorem we have $p = 1$ and thus we can reduce the complexity of the problem by disregarding what happens to the left of the tagged particle. This is justified by the fact that, since the tagged particle never attempts to jump to the left the only interaction, namely blocking the path of another particle, is absent for the tagged particle and hence for all other particles to its right. In the zero range process this feature can be implemented

by only looking at the dynamics on the nonnegative numbers and to consider -1 as an absorption point. In this case the generator reduces to

$$L = L_b + \sum_{z \neq -1} (L_{z,z+1} + L_{z+1,z}) \tag{106}$$

where

$$L_b h(\eta) = g(\eta(0))(h(\eta - \delta_0) - h(\eta)). \tag{107}$$

For this simplified system, let $\{\lambda_{\rho_0(.)}^N\}$ be a sequence of distributions in the zero range process corresponding to $\{\mu_{\kappa_0}^N\}$. We expect the process to have proper hydrodynamic behaviour. So there should exist a bounded function $\rho_0 : \mathbb{R}_+ \to \mathbb{R}_+$ such that for each continuous $G : \mathbb{R}_+ \to \mathbb{R}$ with compact support and each $\delta > 0$,

$$\lim_{N \to \infty} \mathbb{P}_{\lambda_{\rho_0}^N}^N \left(\left| \frac{1}{N} \sum_x G(x/N)\eta(x) - \int G(u)\rho_0(u)du \right| > \delta \right) = 0. \tag{108}$$

Furthermore we expect $\rho_0(u) = F(u)$ for almost all $u \in \mathbb{R}_+$. Later on we will prove both these statements and the corresponding restrictions for the case $p < 1$.

5.2 Weak solutions for the transformed system

Previously two differential equations (86) and (87) were introduced and in addition to that weak solutions to differential equations were discussed. However, since we are dealing with an infinite system do not have a precise definition of what counts as a weak solution to these equations. We will discuss this now for equation (86). Let $\rho_0 : \mathbb{R}_+ \to \mathbb{R}$ be an initial density profile.

Definition 5.1. *A weak solution of equation (86), in the layer $[0,T] \times \mathbb{R}_+$, with initial condition ρ_0, is a bounded function $\rho : [0,T] \times \mathbb{R}_+ \to \mathbb{R}$ satisfying the following conditions*

1. *$\Phi(\rho(t,u))$ is absolutely continuous in the space variable and*

$$\int_0^T ds \int_{\mathbb{R}_+} du e^{-u} \left(\partial_u \Phi(\rho(s,u)) \right)^2 < \infty$$

2. *$\rho(t,0) = 0$ for almost every $0 \leq t \leq T$*

3. *For every smooth function with compact support $G : \mathbb{R}_+ \to \mathbb{R}$ vanishing at the origin and every $0 \leq t \leq T$*

$$\int_{\mathbb{R}_+} du \rho(t,u)G(u) - \int_{\mathbb{R}_+} du \rho_0(u)G(u) = -\frac{1}{2} \int_0^t ds \int_{\mathbb{R}_+} du \frac{dG}{du}(u) \partial_u \Phi(\rho(s,u))$$

With $\Phi(\rho) = \frac{\rho}{1+\rho}$.

In the usual definitions for weak differential equations the solution is defined on a bounded set, now only time is bounded. Space however is not and therefore there is the first requirement, namely a convergence restriction. The second requirement corresponds to $\rho(t,0) = 0$, enforced for almost every $0 \le t \le T$. The last requirement is where we find back both the equation $\partial_t \rho = \frac{1}{2}\Delta\Phi(\rho)$ and our initial condition. The last requirement is what we expect from the definition of a weak solution as we saw it before. We can see that this is indeed a reasonable requirement by looking at the following quantity

$$\int_{\mathbb{R}_+} du \frac{\partial \rho(t,u)}{\partial t} G(u). \tag{109}$$

The left hand side of the third requirement follows directly from applying the fundamental theorem of calculus to (109) and filling in the boundaries. The right hand side is obtained by using the differential equation to substitute away the time derivative. The uniqueness of weak solutions to this system can be proven, arguments for this can be found in an appendix to [13].

5.3 The hydrodynamic limit and its implications

In this section it will be explained how the proof of theorem 4.1 depends on the hydrodynamic behaviour of the zero range process from the previous section. Some preliminary notions will be needed, the first being a partial ordering on probability measures on the state space. This can be done if the probability space itself is equipped with a preorder. A relation $\le$ on a set S is a preorder if it is reflexive and transitive, so for $a, b, c \in S$

1. $a \le a$;

2. $a \le b \wedge b \le c \Rightarrow a \le c$

Note that in the case of $\mathbb{N}_0^{\mathbb{N}_0}$ a preorder can be defined as follows, if $a = (a_0, a_1, ...), b = (b_0, b_1, ...) \in S$ then $a \le b$ iff $a_i \le b_i$ for all $i \in \mathbb{N}_0$. This preorder allows to distinguish upward closed sets. Let $U \subset S$ be a subset, such that if $x \in U$ and $y \in S$ with $x \le y$ then $y \in U$, then U is called upward closed. Let μ and ν be two measures on the same probability space, such that $\mu(U) \le \nu(U)$ for all upward closed sets $U \subset S$, then we say that $\mu \le \nu$. This ordering on measures on $\mathbb{N}_0^{\mathbb{N}_0}$ can now be used to introduce lower and upper bounds on a sequence of measures, a precondition needed to prove the hydrodynamic limit. We now come back to the set of product measures ν_α^+ with marginals given by

$$\nu_\alpha^+\{\eta : \eta(x) = k\} = \frac{1}{1+\alpha}\left(\frac{\alpha}{1+\alpha}\right)^k. \tag{110}$$

Given a measure with constant parameter in the exclusion process these are precisely the corresponding measures in the zero range process. To show the existence of a hydrodynamic limit, for a time interval $[0, T]$, the sequence of initial measures μ^N musts satisfy the following two conditions.

I There exists $0 < \lambda < \alpha$ such that $\nu_\lambda^+ \leq \mu^N \leq \nu_\alpha^+$ for all $N \geq 1$

II There exists a bounded function $\rho_0 : \mathbb{R}_+ \to \mathbb{R}_+$ such that for each continuous function $G : \mathbb{R}_+ \to \mathbb{R}$ with compact support and each $\delta > 0$,

$$\lim_{N \to \infty} \mu^N \left(\left| \frac{1}{N} \sum_x G(x/N)\eta(x) - \int du\, G(u)\rho_0(u) \right| \geq \delta \right) = 0 \qquad (111)$$

In words the second requirement is that at time $t = 0$ the system has a proper hydrodynamic limit. At this point we have to consider the Skorokhod spaces to account for time dependence. For each probability measure μ on $\mathbb{N}_0^{\mathbb{N}_0}$ define a probability measure $\mathbb{P}_\mu^N$ on the space $D(\mathbb{R}_+, \mathbb{N}_0^{\mathbb{N}_0})$ induced by the Markov process with generator (106) accelerated by N^2 and initial measure μ. Expectations with respect to this measure are denoted by $\mathbb{E}_\mu^N$. Now the theorem that actually describes the dynamics of the system is the following

Theorem 5.1. *Given a sequence $\{\mu^N, N \geq 1\}$ satisfying the requirements I and II above, a continuous function $G : \mathbb{R}_+ \to \mathbb{R}$ with compact support and a constant $\delta > 0$ we have*

$$\lim_{N \to \infty} \mathbb{P}_{\mu^N}^N \left(\left| \frac{1}{N} \sum_x G(x/N)\eta_{tN^2}(x) - \int G(u)\rho(t, u)du \right| > \delta \right) = 0 \qquad (112)$$

where ρ is the unique solution of (86).

Applying the entropy method to prove this theorem involves showing the convergence of a sequence Q^N on the path space of this Markov process. In the next paragraph this will be discussed and it will be shown how 5.1 follows from it. Comparing this theorem with the hydrodynamics we previously encountered we see that the only way in which this expression deviates is in it satisfying a different differential equation. More importantly the expression in Theorem 4.1 follows from it. The crucial equality to relate this result to the exclusion process is equation (104). This equation expresses the displacement of the tagged particle in terms of states of the zero range process. If we multiply time with a factor N^2 and divide the expression by N it reads as follows:

$$\frac{X_{tN^2}}{N} = \frac{1}{N} \sum_{x \geq 0} \eta_0(x) - \eta_{tN^2}(x). \qquad (113)$$

From this it follows that, if the transformed measures from theorem 4.1 satisfy the requirements put on the measures for theorem 5.1, that

$$\lim_{N \to \infty} \mathbb{P}_{\mu^{\mathbb{N}}}^{\mathbb{N}} \left(\left| \frac{X_{tN^2}}{N} - \frac{1}{N} \sum_{x \geq 0} \eta_0(x) + \int \rho(t, u)du \right| > \delta \right) = 0. \qquad (114)$$

To show that this plausibly reduces to theorem 4.1 a non-rigorous argument will be given. Later in this chapter we will come back to this problem. Assuming

47

hydrodynamic behaviour of the transformed measures at $t = 0$, we can rewrite the summation in the above expression as

$$\frac{1}{N} \sum_{x \geq 0} \eta_0(x) = \int \rho(0, u) du. \tag{115}$$

The right hand side of this expression refers to the particle density in the continuum limit of the zero range process. Given a particle density κ_0 in the exclusion process then it seems plausible that, for $N \to \infty$, the average width of a gap is given by the function

$$\sum_{n \geq 1} n P(n) = \sum_{n \geq 1} n(1 - \kappa_0)^n \kappa_0 = \frac{1}{\kappa_0} - 1. \tag{116}$$

Furthermore a position in the exclusion process translates to a position in the zero range process by counting the number of particles from the origin and determining the proper sign. In the continuum limit the position x in the exclusion process therefore seems to correspond to a position

$$u = \int_0^x \kappa_0(z) dz \tag{117}$$

in the zero range process. Combining these equation leads to the expression in theorem 4.1

$$\lim_{N \to \infty} \mathbb{P}^{\mathbb{N}}_{\mu^{\mathbb{N}}} \left(\left| \frac{X_{tN^2}}{N} - \int \mathcal{F}(u) - \rho(t, u) du \right| > \delta \right) = 0 \tag{118}$$

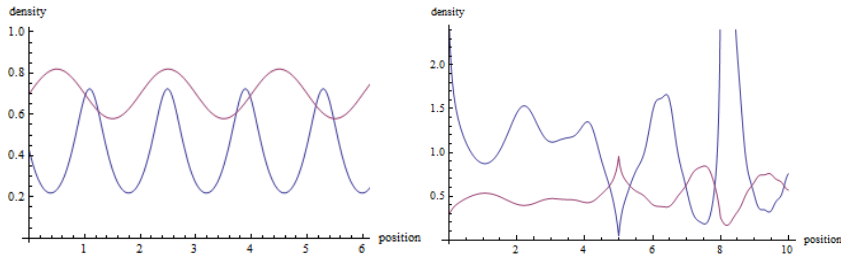with $F(u) = \frac{1}{\kappa_0(H^{-1}(u))} - 1$ and $H(x) = \int_0^x \kappa_0(z) dz$.

Figure 6: Two illustrations of the behaviour of F, the purple graph, given a profile κ_0, the blue graph.

5.4 Entropy method

The entropy method was previously discussed for the normal zero range process, this time an adjustment is needed for the different behaviour at the origin. Also we do not work on the torus but on $\mathbb{R}$, here a sketch will be given on how these

problems can be tackled. However, as is the case for the entropy method itself we will not give a full proof. To start off, denote by Q_μ^N the probability measure on the path space $D([0,T], \mathcal{M}_+)$ corresponding to the accelerated process. Here $\mathcal{M}_+$ denotes the space of positive Radon measures on $\mathbb{R}_+$. Like before the statement regarding the hydrodynamic limit is reduced to the following statement.

Theorem 5.2. *The sequence $\mathbb{Q}_{\mu^N}^N$ converges to the probability measure Q concentrated on the absolutely continuous path $\pi(t, du) = \rho(t, u)du$ whose density is the solution of (86)*

The argument to reproduce the equality in theorem 5.1 is also the same as before. One can take a projection operator h_t sending a path $< \pi_{N^2t}^N : t \in [0,T] >$ to $\pi_{N^2t}^N$. h_t is continuous on the support of Q and thus we have the following convergence $Q^N \circ h_t^{-1} \to Q \circ h_t^{-1}$. Therefore $\pi_{N^2t}^N$ converges in law to $\rho(t, u)du$ and thus the convergence in theorem 5.1 follows. So up to this point the difference with the entropy method as it was discussed before is rather limited. Now however we encounter the first problem, caused by the different behaviour in the origin. The problem being that we only have one invariant measure at hand, δ_0. The entropy of any relevant measure on a torus relative to this measure would be infinite. However the problem is much worse since we try to work on $\mathbb{N}_0$, meaning that we would get infinite entropy, even if we had a nice invariant measure. To circumvent the lack of a nice invariant measure we can look at a measure that is, in some sense, close to the invariant measure. First if we assume that $H(\mu^N | \nu_\beta^+) \leq C_0 N$. Then for $x \geq 0$ Let $\gamma_x = \beta(1+x)/N$ for $0 \leq x \leq N-1$ and $\gamma_x = \beta$ for $x \geq N$. Let the product measure $\nu_{\gamma(.)}^N$ be defined via its marginals as

$$\nu_{\gamma(.)}^N \{\eta : \eta(x) = k\} = (1 - \gamma_x)\gamma_x^k. \tag{119}$$

The relative entropies of the measures μ^N with respect to these measures is, under our assumptions bounded, one can compute that $H(\mu^N | \nu_{\gamma(.)}^N) \leq C_1 N$ where the constant C_1 only depends on α, β and the constant C_0. However the assumption that $H(\mu^N | \nu_\beta^+)$ is bounded is false. In [11] a coupling argument is given that shows that the behaviour of measures restricted to finite subsets (scaling with N) of $\mathbb{N}_0$ is similar enough to that of the actual measures. Thereby allowing usage of the entropy method as long as the relative entropy of the measures on these sets is bounded. These measures have a relative entropy scaling with N and are finite as long as N is finite. Now based on this construction one can work towards proposition 5.1. If $D_\gamma(f)$ is the Dirichlet form of the probability density f with respect to $\nu^N \gamma(.)$, given by

$$D_\gamma(f) = D_{\gamma,b}(f) + \sum_{x \geq 0} D_{x,x+1}(f) \tag{120}$$

where

$$D_{\gamma,b}(f) = \tfrac{1}{2} \int g(\eta(0)) \left(\sqrt{f(\eta - \delta_0)} - \sqrt{f(\eta)} \right)^2 d\nu^N_{\gamma(.)}$$

$$D_{x,x+1}(f) = \tfrac{1}{2} \int g(\eta(x)) \left(\sqrt{f(\eta + \delta_{x+1} - \delta_x)} - \sqrt{f(\eta)} \right)^2 d\nu^N_{\gamma(.)} \qquad (121)$$

Then finally in the next proposition, the proof of which can be found in [11], the last needed estimates on the entropy are given.

Proposition 5.1. *Let S_t^N be the semigroup associated to the generator L corresponding to the accelerated process. Denote by $f_t = f_t^N$ the Radon-Nikodym derivative of $\mu^N S_t^N$ with respect to $\nu^N_{\gamma(.)}$. There exists a finite constant $C = C(\beta)$ such that*

$$\partial_t H(\mu^N S_t^N | \nu^N_{\gamma(.)}) \leq -N^2 D_\gamma(f_t) + CN \qquad (122)$$

From this point onward the differences with the entropy method for the normal zero range process largely boil down to adjustments in the (proof of the) replacement lemma which is beyond the scope of this text. So what we have seen is how one can transform the exclusion process to a zero range process and what approach can be used to determine its hydrodynamic limit in the special case $p = 1$. For this case we still have to show that the transformation behaves as expected in the continuum limit, this will be done in the next chapter after we have compared the $p < 1$ case with the $p = 1$ case.

6 The $p < 1$ case and calculating the Einstein relation

The situation $\frac{1}{2} < p < 1$ is quite similar to the $p = 1$ case so our treatment here will be somewhat shorter. However for both cases we have postponed the proof that in the continuum limit the transformation (96) behaves as expected. So after the discussion of the situation with $p < 1$ we will finish that proof. Finally we then need to obtain the validity of the Einstein relation from this theorem, that will be the last point of discussion in this chapter.

6.1 The approach for the $p < 1$ case

Recall that the state space is equal to the state space of the normal zero range process, being $\mathbb{Z}^{\mathbb{N}_0}$, and the generator is given by

$$L = \sum_{x \neq -1} \left(L_{x,x+1} + L_{x+1,x} \right) + 2pL_{0,-1} + 2qL_{-1,0}. \tag{123}$$

A fortunate difference with the case that was discussed earlier is the existence of an invariant measure, namely the product measure $\bar{\nu}_\phi^i$ with marginals given by

$$\bar{\nu}_\phi^i \{ \eta : \eta(x_0 = k \} = \frac{1}{Z(\phi_x)} \frac{\phi_x^k}{g(k)!}, \tag{124}$$

where $\phi_x = p\phi$ if $x \leq -1$ and $\phi_x = q\phi$ if $x \geq 0$ and Z is the normalisation factor. The hydrodynamic behaviour is assumed to be reflected by equation (87) so we have to be precise about what it means to for a function to be a weak solution to this equation.

Definition 6.1. *A weak solution of the partial differential equation (87) is a function $\rho : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ such that*

1. *$\Phi(\rho(t, u))$ is absolutely continuous in the space variable and for every $t > 0$*

$$\int_0^t ds \int_\mathbb{R} du e^{-|u|} \left(\partial_u \Phi(\rho(s, u)) \right)^2 < \infty$$

2. *$p\Phi(\rho(t, 0+)) = q\Phi(\rho(t, 0-))$ for almost every $t \geq 0$*

3. *For every smooth function with compact support $G : \mathbb{R} \to \mathbb{R}$ vanishing at the origin and every $0 \leq t \leq T$*

$$\int_\mathbb{R} du \rho(t, u) G(u) - \int_\mathbb{R} du \rho_0(u) G(u) = -\frac{1}{2} \int_0^t ds \int_\mathbb{R} du \frac{dG}{du}(u) \partial_u \Phi(\rho(s, u))$$

Here we have a problem because ρ is supposed to be a solution to a weak differential equation. This implies that we know it to be measurable, but not

that these limits are well defined. Therefore the second requirement must be interpreted as

$$\lim_{\epsilon \to 0} \int_0^t h(s) \left(p\Phi\left(\frac{1}{\epsilon} \int_0^\epsilon \rho(s,u)dy\right) - q\Phi\left(\frac{1}{\epsilon} \int_{-\epsilon}^0 \rho(s,u)du\right) \right) ds = 0 \quad (125)$$

for all continuous functions h. For uniqueness we refer again to [11]. For initial states the following requirements are imposed

I' There exist $0 < \lambda < \alpha$ such that $\bar{\nu}_\lambda^+ \leq \mu^N \leq \bar{\nu}_\alpha^+$ for all $N \geq 1$

II' There exists a bounded function $\rho_0 : \mathbb{R} \to \mathbb{R}_+$ such that for each continuous function $G : \mathbb{R} \to \mathbb{R}_+$ with compact support and each $\delta > 0$,

$$\lim_{N \to \infty} \mu^N \left(\left| \frac{1}{N} \sum_x G(x/N)\eta(x) - \int duG(u)\rho_0(u) \right| \geq \delta \right) = 0 \quad (126)$$

The hydrodynamic limit is then given by the usual equation

Theorem 6.1. *The hydrodynamic limit of the zero range process with generator (123) on $\mathbb{Z}^{\mathbb{N}_0}$ with a sequence of initial measures satisfying I' and II' above is given by*

$$\lim_{N \to \infty} \mathbb{P}_{\mu^N}^N \left(\left| \frac{1}{N} \sum_x G(x/N)\eta_t(x) - \int G(u)\rho(t,u)du \right| \geq \delta \right) = 0 \quad (127)$$

where ρ is a weak solution to (87).

This can then be proven via an adjustment to equation (5.2) to this situation. Let $Q_{\mu^N}^N$ be the probability measure on the path space $D(\mathbb{R}_+, \mathcal{M}_+)$ induced by the accelerated process.

Theorem 6.2. *The sequence $\mathbb{Q}_{\mu^N}^N$ converges to the probability measure concentrated on the absolutely continuous path $\pi(t,du) = \rho(t,u)du$ whose density ρ is the solution of 87*

The argument to obtain theorem 6.1 from theorem 6.2 is the same as for the $p = 1$ case. As in the $p = 1$ case one then has to make an assumption about a bound on the entropy, namely $H(\mu^N | \bar{\nu}_\beta^i) \leq C_0 N$, for some β and N. A similar coupling argument, given in [13], allows to remove this assumption afterwards. From there on one has to obtain a version of the replacement lemma for the entropy method to work, details can also be found in [13].

6.2 The properties of the transformation $\mathcal{T}$

As for the entropy method we merely provided some of the background needed to apply it to the case of the zero range process with an anomaly in the origin. Basing the validity of the hydrodynamic limit of zero range process on the

results of [13] would still not be sufficient to support the conclusion that the hydrodynamic behaviour of the zero range process implies the validity of the theorems 4.1 and 4.2. The following statement is what we wish to prove in both cases, where the only difference lies in the definitions of ρ and F:

$$\lim_{N \to \infty} \mathbb{P}_{\mu^\mathbb{N}}^\mathbb{N} \left(\left| \frac{X_{tN^2}}{N} - \int F(u) - \rho(t, u) du \right| > \delta \right) = 0. \qquad (128)$$

Proving this is the purpose of this paragraph. The key is showing the validity of the following proposition.

Proposition 6.1. *In the case $p = 1$, consider a sequence of probability measures μ^N satisfying the corresponding assumptions for the initial measures, I' and II'. Then, for every $t \geq 0$ and $\delta > 0$,*

$$\lim_{N \to \infty} \mathbb{P}_{\mu^N}^N \left(\left| N^{-1} \sum_{x \geq 0} (\eta_t(x) - \eta_0(x)) - \int_0^\infty \rho(t, u) - \rho_0(u) du \right| > \delta \right) = 0 \quad (129)$$

where ρ is the solution of (86). In the case $p < 1$, consider a sequence of probability measures μ^N satisfying the corresponding assumptions, I and II, for the initial measures. Then, for every $t \geq 0$ and $\delta > 0$ the same expression holds, where ρ is now the solution to (87).

This proposition follows from the hydrodynamic limit in combination with the constraints on the probability measures μ^N. This proposition then implies the validity of theorems 4.1 and 4.2 if the measures introduced in in those theorems indeed transform to measures satisfying the constraints needed for proposition 6.1. This is made precise in the following proposition

Proposition 6.2. *Fix a sequence of initial states $\mu_{\rho_0(.)}^N$ satisfying the assumptions of Theorem 4.1 or 4.2. The sequence $\mathcal{T}\mu_{\rho_0(.)}^N$ satisfies the corresponding assumptions, I and II or I' and II' respectively, for the initial measures.*

Proof. The proof will be given for $p = 1$ the case $p < 1$ is analogous. First we will check the existence of the bounds on the measures. We are given an initial profile $\rho_0 : \mathbb{R}_+ \to [0, 1]$ for which there exists a $\sigma > 0$ such that $\sigma \leq \rho_0 \leq 1 - \sigma$. We also have a product measure $\mu_{\rho_0(.)}^N$ associated to the density profile. Now let $\nu_{\rho_0(.)}^N$ be a product measure for the zero range model given by $\mathcal{T}\mu_{\rho_0(.)}^N$. If it can be shown that, given a product measure on $\{0, 1\}^\mathbb{N}$ that is bounded above by μ_ρ^+ for some $0 < \rho < 1$, then $\mathcal{T}\mu$ is bounded below by $\nu_{(1-\rho)/\rho}^+$. Likewise for a lower bound μ_ρ^- the image under the transformation has an upper bound $\nu_{(1-\rho)/\rho}^-$. Assume now that $\mu \leq \mu_\rho^+$, for $x \geq 1$ denote by γ_x the probability of finding a particle at x under the measure μ. Then $\gamma_x \leq \rho$. For $j \geq 1$ we denote by N_j the random variable corresponding to the position of the jth particle. Now a coupling argument will be used, if we consider the joint process on $\{0, 1\}^{\mathbb{N}_0} \times \{0, 1\}^{\mathbb{N}_0}$ with distribution $\mu \times \rho$ then $N_1^\mu \geq N_1^\rho$ and $N_{j+1}^\mu - N_j^\mu \geq$

53

$N_{j+1}^{\rho} - N_j^{\rho}$ for $j \geq 1$. Where the superscripts indicate that the position of the jth particle is taken under the corresponding distribution. Now applying $\mathcal{T}$ to this coupling measure yields a measure on $\mathbb{N}_0^{\mathbb{N}_0} \times \mathbb{N}_0^{\mathbb{N}_0}$ with first marginal given by $\mathcal{T}\mu$ and second marginal $\mathcal{T}\mu_{\rho}^+ = \nu_{(1-\rho)/\rho}^+$. The coupling measure is concentrated on states (η^1, η^2) below the diagonal, thus $\mathcal{T}\mu \geq \nu_{(1-\rho)/\rho}^+$. This then shows $\nu_{\sigma/(1-\sigma)}^+ \leq \nu_{\rho_0(.)}^N \leq \nu_{(1-\sigma)/\sigma}^+$, concluding the first part of the proof. It remains to show the validity of an initial hydrodynamic limit. The validity of this requirement already follows if we know that for all $\delta > 0$:

$$\lim_{N \to \infty} \nu_{\rho_0(.)}^N \left(\left| \frac{1}{N} \sum_{x=0}^{\lfloor BN \rfloor} \eta(x) - \int_0^B F(u)du \right| > \delta \right) = 0. \tag{130}$$

In order to show this note that by definition of F we have that, for all $B > 0$:

$$\int_0^B F(u)du = H^{-1}(B) - B. \tag{131}$$

The total number of sites between the origin and $n \in \mathbb{N}$ equals the number of particles plus the sum of the length of all the holes. This leads to the estimate

$$\sum_{x=0}^{n} \xi(x) + \sum_{y=0}^{-1+\sum_{x=0}^{n} \xi(x)} \eta(y) \leq n+1 \leq \sum_{x=0}^{n} \xi(x) + \sum_{y=0}^{\sum_{x=0}^{n} \xi(x)} \eta(y), \tag{132}$$

with ξ a state in the exclusion process corresponding to the state η in the zero range process. The last piece of information needed is the convergence of $\frac{1}{N} \sum_{x=0}^{\lfloor nN \rfloor} \xi(x)$ to $\int_0^n \kappa_0(u)du$. Using the convergence of the sum over $\xi(x)/N$ and substituting $n = \lfloor NH^{-1}(B) \rfloor$ in the inequality then yields an estimate from which the desired result quickly follows. $\qquad \square$

6.3 Solving the differential equation

We now have rather precise results about the behaviour of the tagged particle in the exclusion process. Finally we should look how the third theorem, the Einstein relation, follows from these results. We could try to do this by solving the differential equation governing the system for a constant density profile. However it is possible to start out with a more general initial density:

$$\rho_0(u) = \rho_+ \mathbb{1}_{\{x \geq 1\}}(u) + \rho_- \mathbb{1}_{\{x < 0\}}, \tag{133}$$

with ρ_+ and ρ_- such that the requirements of theorem 4.2 are met. In [11] this problem is then solved by reducing the set of equation to a new set of equations describing a solution to the self-scaling Stefan problem. From the solution to this problem the validity of the Einstein relation follows. In this paragraph their solution to the problem will be presented, a few comments are omitted in this

text. For convenience, recall that the differential equation we wish to solve is given by:

$$\begin{cases} \partial_t \rho = \frac{1}{2}\Delta\Phi(\rho) \\ p\Phi(\rho(t,0+)) = q\Phi(\rho(t,0-)) \\ \partial_u \Phi(\rho(t,0+)) = \partial_u \Phi(\rho(t,0-)) \\ \rho(0,.) = F(.) \end{cases} \tag{134}$$

The first step is recognizing that the equation is solved by densities of the form

$$\rho(t,u) = \phi(u/\sqrt{t}), \tag{135}$$

where $\phi : \mathbb{R} \to \mathbb{R}$ is the solution of

$$\begin{cases} -z\phi'(z) = \partial_z^2(\phi(z)) \\ \frac{\phi'(0+)}{1+\phi(0+))^2} = \frac{\phi'(0-)}{(1+\phi(0-))^2} \\ \phi(\pm\infty) = \rho_\pm \\ p\Phi(\phi(0+)) = q\Phi(\phi(0-)) \end{cases} \tag{136}$$

with $\Phi(\rho) = \frac{\rho}{1-\rho}$. This can be checked by explicit computation. The resulting function v_t is in this case given by $v\sqrt{t}$ where the parameter v is given by

$$v = \int_0^\infty (\rho_+ - \phi(y))\, dy. \tag{137}$$

This can be seen by using the fact that on $\mathbb{R}_+$ the value of F is given by ρ_+ and subsequently performing a coordinate transformation $y = \frac{u}{\sqrt{t}}$. Note that the convergence of the integral must still be understood in the way it was described in the previous chapter. The integrand of v can be written differently by using $\phi(\infty) = \rho_+$:

$$\rho_+ - \phi(y) = \int_y^\infty \partial_z \phi(z)dz. \tag{138}$$

Using the differential equation and then integration by parts leads to the derivative of $\Phi(\phi(u))$:

$$v = \frac{\phi'(0+)}{1+\phi(0+))^2}. \tag{139}$$

The following transformation corresponds to the inverse of the transformation $\mathcal{T}$

$$\begin{aligned} x(z) &= \int_0^z (1+\phi(y))dy \\ m(x) &= \frac{1}{1+\phi(z(x))} \end{aligned} \tag{140}$$

The function m introduced this way solves the following equation

$$\begin{cases} m''(x) = -(x+v)m'(x) \\ -v = \frac{m'(0+)}{m(0+)} = \frac{m'(0-)}{m(0-)} \\ p(1-m(0+)) = q(1-m(0-)) \\ m(\pm\infty) = \alpha_\pm = \frac{1}{1+\rho_\pm} \end{cases} \tag{141}$$

55

This is a self-scaling Stephan problem, a solution to this problem can be found in [11], the text there is based on [20]. It is solved by

$$m(x) = \begin{cases} A_+ + B_+ \int_0^x xe^{-\frac{1}{2}y^2 - vy} dy \\ A_+ + B_- \int_0^x e^{-\frac{1}{2}y^2 - vy} dy \end{cases} \tag{142}$$

with the following relations on the parameters:

$$\begin{aligned} p(1 - A_+) &= q(1 - A_-) \\ -v &= \frac{B_+}{A_+} = \frac{B_-}{A_-} \\ \alpha_\pm &= A_\pm J(\pm v) \end{aligned} \tag{143}$$

Here J is defined by $J(v) = 1 - v \int_0^\infty e^{-1/2y^2 - vy} dy$. Then the following equality can be derived:

$$p\left(1 - \frac{\alpha_+}{J(v)}\right) = q\left(1 - \frac{\alpha_-}{J(-v)}\right). \tag{144}$$

This finally allows for a first order estimate in the case of a constant density profile ($\alpha_+ = \alpha_-$). From the previous equality it follows that the following holds:

$$(p - q)\frac{1 - \alpha}{\alpha} = \frac{pJ(v - v) - qJ(v) - (p - q)J(v)J(-v)}{J(v)J(-v)}. \tag{145}$$

Then, when $p - q$ is small, this leads to the first order estimate

$$v = (p - q)\sqrt{\frac{2}{\pi}\frac{\alpha}{1 - \alpha}} + o(p - q). \tag{146}$$

Thus concluding the validity of the Einstein relation.

7 Summary

We set out to discuss the Einstein relation in anomalous diffusive systems and in particular for the one dimensional exclusion process. It was shown that this relation indeed holds for this particle system. In order to obtain this result some of the theory of continuous time Markov process was discussed so that we could rigorously describe the systems at hand. Then this theory was used to introduce the concept of hydrodynamic limits of discrete particle systems. These limits were defined via weak convergence of empirical measures to a measure given by $\rho(t, u)du$, where ρ solves a weak partial differential equation and du is the Lebesque measure. Some background information was then given about how the existence of such a hydrodynamic limit can be shown by the entropy method. Afterwards the article by Landim, Olla and Volchan was discussed and their approach, transforming the exclusion process to the zero range process, was explained. There we saw how the hydrodynamic behaviour of this transformed process can be translated back to the behaviour of a tagged particle. We concluded by solving the differential equations describing that system. Having a solution for this system raises the question under what conditions anomalously diffusive system satisfy the Einstein relation. For now we only have results for individual systems like the exclusion process.

References

[1] Einstein, A. *On the movement of small particles suspended in stationary liquids required by the molecular-kinetic theory of heat* (in German) Ann. d. Phys., 17 (1905), 549-560.

[2] Philibert, J. *One and a Half Century of Diffusion: Fick, Einstein, before and beyond*, Diffusion Fundamentals, Vol. 4, (2011)

[3] Snieder, R. *Retrieving the Greens function of the diffusion equation from the response to a random forcing.* Phys Rev E Stat Nonlin Soft Matter Phys. (2006)

[4] Kubo, R. *The fluctuation-dissipation theorem* Rep. Prog. Phys. 29 255 (1966)

[5] Panja, D. *Anomalous polymer dynamics is non-Markovian: memory effects and the generalized Langevin equation formulation.* J. Stat. Mech. (2010)

[6] Richardson, L.F. *Atmospheric Diffusion Shown on a Distance-Neighbour Graph.* Proceedings of the Royal Society of London, Vol. 110, No. 756 (Apr. 1, 1926), pp. 709-737

[7] Dubkov, A.A., Spagnolo, B., Uchaikin, V.V., *Lévy Flight Superdiffusion: An Introduction.* Intern. Journ. of Bifurcation and Chaos, Vol. 18, No. 9, (2008)

[8] Gradenigo, G., Sarracino, A., Villamaina, D. Vulpiani, A, *Einstein relation in superdiffusive systems.* J. Stat. Mech. (2012)

[9] Beijeren, H. Kehr, K.W., Kutner, R., *Diffusion in concentrated lattice gases. III. Tracer diffusion on a one dimensional lattice.* Physical Review B, Vol 28, Number 10, (1983)

[10] Arratia, R. *The motion of a tagged particle in the simple symmetric exclusion system in $\mathbb{Z}$.* Ann. Prob. 11, 362-373 (1983)

[11] Landim, C. *Hydrodynamic Limit of Interacting Particle Systems.* Lecture notes, (2002)

[12] Kipnis, C., Landim, C., *Scaling limits of interacting particle systems.* Springer-Verlag, Berlin, (1999)

[13] Landim, C., Olla, S., Volchan, S.B., *Driven Tracer Particle in One Dimensional Symmetric Simple Exclusion.* Comm. Math. Phys. 192, 287 - 307 (1998)

[14] Sethuraman, S., lecture notes on hydrodynamic limits available at http://math.arizona.edu/ sethuram/588

[15] Landim, C. *Hydrodynamic Limits of Interacting Particle Systems* . ICTP Lecture Notes 17, School and Conference on Probability Theory, 57-100 (2004)

[16] Timo Seppäläinen, *Translation Invariant Exclusion Processes.* (book in progress), (2003)

[17] Ligett, T.M. *Interacting Particle Systems.* Springer-Verlag, New York, (1985)

[18] Guo, M.Z., Papanicolaou, G.C. and Varadhan, S.R.S. *Nonlinear diffusion limit for a system with nearest neighbor interactions.* Comm. Math. Phys. 118, 3159 (1988)

[19] Evans, L.C. *Partial Differential Equation.* American Mathematical Society, (1998)

[20] Burlatsky, S.F., De Coninck, J., Moreau, M., Oshanin, G.S., *Dynamics of the shock gront propagation in a one-dimensional hard core lattice gas.* Preprint (1997)