An analysis of the Rudvalis Shuffle

David Tadeu Lopes Moreira
david.tadeu.98@ist.utl.pt

Instituto Superior Técnico, Lisboa, Portugal

December, 2021

Abstract

We will study a specific way to shuffle a finite number of cards. The shuffle consists of drawing the first card and flipping a coin: if it lands heads, we place the card at the last position of the deck, while if tails come out, the card is placed at the second-to-last position of the deck. This shuffle was introduced by Arunas Rudvalis (cf. [2], page 90), so we call it the Rudvalis shuffle. First, we define basic properties of Markov chains and then, we study the Rudvalis shuffle, using Markov chains and the mixing time. The first goal is to study the problem solved by Wilson in [11], i.e., to estimate a lower bound (as a function of the number of cards in the deck) for the minimum number of times we have to shuffle the cards if we want the deck to be well shuffled. The second goal is to apply the Rudvalis shuffle and look at the configuration of the cards as a particle system (identifying the red cards as empty sites and the black cards as occupied sites), making use of continuous time Markov chains. The main result we obtain is the Hydrodynamic Limit, which characterizes the evolution of the particle density.
Keywords: Markov chains, partial differential equations, hydrodynamic limit, particle system, mixing time, transport equation.

1. Introduction

A Markov chain is a stochastic process, i.e., a collection of random variables indexed on time, which has the property (Markov property) that the future of the process depends only on its present and it is conditionally independent of its past. There are many real life occurrences of Markov chains but our focus in this paper will be a Markov chain generated by a card shuffle, more precisely, a random walk followed by the permutations of the cards when the deck is shuffled. Here the state space is the symmetric group S_n (whose elements are called permutations) and the transitions are given by the specific chosen way to shuffle the cards.

We are interested on the number of shuffles it takes for a deck to be well shuffled. But what does it mean for the deck to be well shuffled? Well, it means that after a certain number of shuffles, all the possible $n!$ permutations are equally likely. We are going to consider a shuffle, which was introduced by Arunas Rudvalis [2], and it is very simple to understand. We take the first card of the deck and we insert it in the last position with probability $1/2$ or in the second-to-last position with probability $1/2$.

We shall look for the number of shuffles needed for the deck to be well shuffled after successively applying the previous shuffle. But how can we do this? We will start by defining some basic concepts of discrete time Markov chains, namely, the mixing time, which gives the time it takes for the chain to be arbitrarily close to its stationary distribution. If the deck reaches the stationary state, which is the uniform distribution, this means that at that moment any permutation is equally likely. In other words, when the state of the deck reaches stationarity, the deck is well shuffled. Thus, finding the number of shuffles it takes for the deck to be well shuffled boils down to estimating the mixing time of the Markov chain described by the Rudvalis shuffle. Having realized that, we will use some properties of eigenvectors to estimate the mixing time of the Rudvalis shuffle. The upper bound of the mixing time Rudvalis shuffle is $O(n^3 \log(n))$ [6] (after that number of shuffles the deck is well shuffled). Here we explain the method used by Wilson in [11], which shows that the lower bound for the shuffle is also $O(n^3 \log(n))$ (we need at least that number of shuffles for the deck to be well shuffled). Therefore we can conclude that the mixing time for the Rudvalis shuffle is $O(n^3 \log(n))$.

For the second part of this paper, we will interpret the Rudvalis shuffle as an interacting particle system on n sites, associating black cards with particles and red cards with vacant sites. We want instead to find what happens to the system when n goes to infinity, i.e., we want to know how the particles behave when the number of sites goes to

infinity.

This scenario is motivated by the interaction between particles (microscopic) in a real system, such as a gas or a fluid (macroscopic). In this kind of systems, the number of particles is very large and thus it is impossible to study the behaviour of each particle individually in order to understand the collective behaviour of the system, i.e., it is not easy to study the macroscopic space (for example the gas) using the microscopic space (looking at the motion of all individual particles). To overcome this problem, what it is done instead is to consider that the particles move randomly according to some probability law (this is shown to model well real macroscopic systems) and use the results obtained for the microscopic case to draw conclusions for the macroscopic space.

The last strategy is done by considering an initial density of particles at the macroscopic level and a probabilistic law for the microscopic particles (this will be determined by the Rudvalis shuffle). Then we associate the density of the macroscopic system with an initial distribution of particles in the microscopic system, to which we associate a continuous Markov chain, modelling the microscopic system. This system conserves one or more quantities (in our case, the number of red and black cards is the preserved quantity) but the distribution of this quantity changes through time. The question we ask is how can we approximate the distribution of particles through time, and the hydrodynamic limit tells us that the density of particles is described by a solution of a partial differential equation (PDE). In other words, the hydrodynamic limit allows finding a macroscopic law (in general a solution of a PDE) through a microscopic underlying random dynamic.

The hydrodynamic limit is really interesting in the sense that it allows us to relate concepts of probability theory with partial differential equations. This is because the density of particles in the macroscopic space is usually described by a weak solution of a PDE. So if we take a random process (the card shuffle), this random process gives rise, in the limit, to a deterministic process. Furthermore, if we prove the uniqueness of the weak solution of that PDE, the hydrodynamic limit allows concluding the existence of a weak solution to a PDE which a priori we did not know it existed. In other words, we use a random process to prove the existence of a (weak) solution of a PDE. This is exactly what we aim to do in the second part of this paper: explore the Rudvalis shuffle (which is a random process) and use it to find the (weak) solution of a PDE, in our case, the transport equation.

2. Discrete time Rudvalis shuffle

2.1. Discrete time Markov chains

In this section we start with an overview of discrete time Markov chains and some of its properties.

2.1.1 Basic properties

Let $(\tilde{\Omega}, \mathcal{F}, \mathbb{P})$ be a probability space. Throughout the rest of this section, we shall denote by $\{X_t\}_{t\in\mathbb{N}_0}$ an irreducible and aperiodic Markov chain with finite state space Ω, transition matrix P and initial measure $\mu_0(\cdot) = P(X_0 \in \cdot)$. Let $\mathcal{F}_t := \sigma(X_s \mid s \leq t)$ be the σ-algebra generated by $X_1, X_2, \ldots, X_t$ and $\{\mathcal{F}_t\}_{t\in\mathbb{N}_0}$ the filtration associated with $\{X_t\}_{t\in\mathbb{N}_0}$. Moreover, denote by $\mathbb{E}_{\mu_0}$ the expectation starting from μ_0. It is well known that every Markov chain with finite state space has at least one stationary distribution and that if the Markov chain is irreducible, the stationary distribution is unique [9]. Let π be the stationary distribution of $\{X_t\}_{t\in\mathbb{N}}$. Furthermore, the distribution of the chain converges (in total variation distance) to the unique stationary distribution π (Section 4.3 of [9]).

But how much time t does it take for the distribution of the chain to get arbitrarily close to the stationary distribution? With that question in mind, we define a function that keeps track of the distance between the distribution of the chain at time t and the stationary distribution:

$$d(t) = \max_{x\in\Omega} \left\| P^t(x,\cdot) - \pi \right\|_{TV},$$

where

$$P^t(x,y) := \mathbb{P}(X_t = y | X_0 = x)$$

is the probability of reaching state y in t time steps, given that $X_0 = x$, that is, given that the chain started at state x.

Since $d(t)$ is a decreasing function (Exercise 4.2 of [9]), i.e., the more the chain evolves, the closer it is to stationarity, we define the *mixing time* of a Markov chain as follows:

Definition 2.1 (Mixing time)**.** For $\varepsilon \in (0,1)$,

$$t_{mix}(\varepsilon) = \min\{t : d(t) \leq \varepsilon\}.$$

2.1.2 Eigenfunctions

It will be useful to introduce the concept of *eigenfunctions* and *eigenvalues* of the transition matrix of a Markov chain. Moreover, there are properties relating eigenfunctions and the stationary distribution of a Markov chain.

Definition 2.2 (Eigenfunction)**.** Let P be a transition matrix of a Markov chain with finite state space Ω. A function f on Ω is an eigenfunction (or eigenvector) of P associated with the eigenvalue λ if $Pf = \lambda f$.

The fact that the transition matrix P is right stochastic implies that all the eigenvalues have absolute value bounded from above by 1.

Lemma 2.1. *Let $\{X_t\}_{t\in\mathbb{N}_0}$ be a discrete Markov chain with finite state space Ω, transition matrix P and initial distribution μ_0. Let f be an eigenfunction of P with eigenvalue λ. Then*

1. $\mathbb{E}_{\mu_0}[f(X_t)] = \lambda^t f(X_0)$.
2. $|\lambda| \leq 1$ *and there exists always one eigenvalue equal to* 1.

Given a function $\Psi : \Omega \to \mathbb{C}$ and a probability measure μ on a finite space Ω, we denote by $E_\mu[\Psi]$ the expectation of Ψ under μ.

Lemma 2.2. *Let P be a transition matrix of a Markov chain with finite state space Ω and stationary distribution π. If f is an eigenfunction with eigenvalue $\lambda \neq 1$, then $E_\pi[f] = 0$.*

2.2. Rudvalis shuffle

We will now take a look at a very specific way to shuffle a deck of cards. Since a card shuffle consists of a random walk on the permutations of the deck, this shuffle can be seen as a Markov chain.

Let us consider a deck with n cards numbered from 1 to n (we label the first card with 1, the second with 2, and so on). At each time $t \in \mathbb{N}$, we remove the card on top of the deck and insert that card at the second position from the bottom with probability p and at the bottom of the deck with probability $1 - p$, where $p \in (0, 1)$.

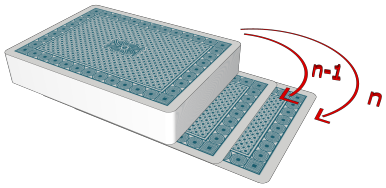


Figure 1: Inserting the top card at positions $n - 1$ or n with probability p or $1 - p$ respectively.

In order to see the dynamics of the shuffle, consider the following example of two consecutive shuffles for a deck with $n = 19$ cards.

Example 2.3. *Suppose the deck has $n = 19$ cards. We can use a (not necessarily fair) coin to decide where we place the first card. If it lands "heads", we place the first card at position $n = 19$, if it lands "tails", we place it at position $n - 1 = 18$. The following image shows the four different possibilities after two shuffles. The top card and bottom cards of the initial configuration of the deck are respectively the* 3 *of diamonds and* 8 *of clubs (we can imagine the deck in a circle where the last and first cards are "connected"). After the first shuffle, the* 3 *of diamonds either goes to position $n - 1$ or position n. We see that in both cases, all the cards but the* 3 *of diamonds (and possibly the* 8 *of clubs) are pushed up one position, i.e., almost all the cards rotate counterclockwise. If the deck keeps being shuffled, the cards keep "rotating" and as we will see later on, this is what makes the Rudvalis shuffle so unique.*

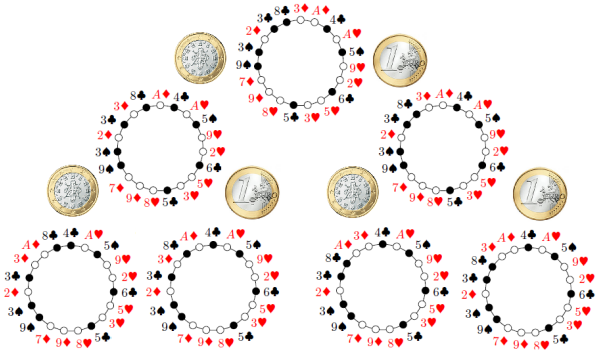


Figure 2: Example of the Rudvalis shuffle with $n = 19$.

The discrete-time Markov chain described by Figure 1 was introduced by Arunas Rudvalis (in fact it is a slight generalization of the original shuffle proposed by Rudvalis, which considers $p = \frac{1}{2}$), and for this reason, it is known as the Rudvalis shuffle. Its state space is the symmetric group S_n of all possible permutations of an n-element set. For $\sigma \in S_n$, we write $\sigma(x) = y$ if card y is in position x of the deck. The transition matrix P is given by

$$P(\sigma, \xi) = \begin{cases} p, & \text{if } \xi = \sigma^{n-1} \\ 1 - p, & \text{if } \xi = \sigma^n \\ 0, & \text{otherwise} \end{cases}$$

where σ^x is the permutation obtained from σ after we move the top card to position x.

We denote the Rudvalis Markov chain by $\{\sigma_t\}_{t\in\mathbb{N}_0}$. Moreover, we denote by $P^t(\sigma, \cdot)$ the distribution of the deck after t shuffles, given that the initial permutation of the deck is σ. For example, $P^t(\sigma, \xi)$ is the probability of having permutation ξ after t shuffles, given that the starting permutation is σ.

We denote by $\mathbb{E}_\sigma$ the expectation with respect to the law of the process when the initial distribution $\mu_0(\cdot) = \mathbb{P}(\sigma_0 \in \cdot)$ is concentrated on σ, that is, the initial permutation of the deck is σ.

If the Rudvalis shuffle was not an irreducible chain, it would not make much sense to call it a shuffle (we would not be able to achieve some specific configurations). But as it can be easily verified, the Rudvalis chain is irreducible.

Since the chain is irreducible and has a finite state space, it has a unique stationary distribution U_n,

which is the uniform distribution on S_n (cf. Proposition 2.16 of [9]). In fact, it is simple to verify that $U_nP = U_n$. For any permutation σ,

$$(U_nP)(\sigma) = \sum_{\xi\in S_n} U_n(\xi)P(\xi,\sigma) = \sum_{\xi\in S_n} \tfrac{1}{n!}P(\xi,\sigma)$$
$$= \tfrac{1}{n!}(P(\xi_1,\sigma) + P(\xi_2,\sigma)) = \tfrac{1}{n!} = U_n(\sigma),$$

where ξ_1 and ξ_2 are the permutations such that $\sigma = \xi_1^{n-1} = \xi_2^n$.

If we keep shuffling this way, how long must we shuffle the deck until it is well shuffled? But what does it mean for the deck to be well shuffled? It means that the arrangement of the deck is close to random. Thus, another way to ask this question is: How many shuffles does it take for the distribution of the deck to become close to uniform? As we have seen before, this can be done using the mixing time of the Markov chain. While an upper bound of $O(n^3\log(n))$ for the mixing time was already found (cf. [6]), in this paper we focus on finding a lower bound for the mixing time of this chain. The bound for the mixing time translates into a bound on the number of shuffles which tells us that if we shuffle the deck any number of times below that bound, we know for sure that the deck is still not well shuffled. In order to find this bound, we first find a lower bound for the total variation distance between any two probability measures. Then, we consider the case in which one of the measures is the uniform measure on S_n and the other is the distribution of the state of the deck after t shuffles. This will allow finding a lower bound for $d(t)$. Finally, we will use this result to obtain a lower bound for $t_{mix}(\varepsilon)$.

Thus, we need to find $g(n,\varepsilon)$ such that $t_{mix}(\varepsilon) \geq g(n,\varepsilon)$. If we can find such a function, we have just shown that for any initial configuration of the deck, we need to shuffle the deck at least $g(n,\varepsilon)$ times if we want the deck to become well shuffled. However, shuffling $g(n,\varepsilon)$ times does not guarantee that the deck will be well shuffled.

2.2.1 Lower bound for the mixing time of the Rudvalis shuffle

We will use the following proposition to find a lower bound for the distance between the distribution of the chain at time t and the stationary distribution, i.e., a lower bound for $d(t)$. If we can bound $d(t)$ by a "large" number, i.e., if we show that for small $\delta > 0$, there is t^* (possibly depending on ε) such that $d(t^*) \geq 1-\delta$ then we can conclude, since d is decreasing, that $t_{mix}(\varepsilon) > t^*$.

Proposition 2.4. *Let μ and ν be two probability distributions on a finite state space Ω and let $f : \Omega \to \mathbb{C}$ be a complex-valued function on Ω. Then,*

$$\|\mu-\nu\|_{TV} \geq 1 - 8\,\frac{\max\{Var_\mu(f), Var_\nu(f)\}}{|\,|E_\mu[f]| - |E_\nu[f]|\,|^2}. \quad (1)$$

This result is particularly useful if we consider the Rudvalis Markov chain $\{\sigma_t\}_{t\in\mathbb{N}_0}$ and take $\mu := P^t(\sigma,\cdot)$ (μ is time dependent), the distribution of the state of the deck after t shuffles given that the initial state is $\sigma \in S_n$, and $\nu := U_n$, the uniform measure on S_n (ν does not depend of t). Indeed, by Proposition 2.4 we have a lower bound for $\|P^t(\sigma,\cdot) - U_n\|_{TV}$, which gives a lower bound for $d(t)$. We would just need to compute $E_\mu[f(\sigma_t)] := \mathbb{E}_\sigma[f(\sigma_t)], E_\nu[f(\sigma_t)], Var_\mu(f(\sigma_t))$ and $Var_\nu(f(\sigma_t))$.

When we look at (1), we realize that in order to get $d(t)$ close to 1 (which means that the chain has not mixed by time t), we would like the variances to be small and the difference of expectations to be large. But now the question is, how can we find a function $f : S_n \to \mathbb{C}$ which provides these properties? The main idea, as we will see, is to use the eigenfunctions of the transition matrix of a Markov chain.

In the rest of this section, we follow the approach in [11]. The idea is to find an eigenvector Ψ (with eigenvalue λ) not for the original chain, but for another chain (X_t, Y_t), where Y_t is obtained by taking into account more information about the shuffle. Wilson [11] refers to the chain $\{(X_t,Y_t)\}_{t\in\mathbb{N}_0}$ as the *lifted chain*. We want the eigenvector of the lifted chain to have the property that for all x, y_1, y_2, the eigenvector verifies $|\Psi(x,y_1)| = |\Psi(x,y_2)|$, so that $|\Psi(X_t,Y_t)|$ is a function of X_t alone. Now let us see how to define Y_t and then obtain an eigenfunction Ψ with the desired properties.

Recall that the state σ_t of the Rudvalis chain is the permutation giving the order of the cards at time t. It will be convenient to look at the position of a particular card. Hence denote the position of card k (with $k \in \{1,\cdots,n\}$) at time t by $X_t(k) := \sigma_t^{-1}(k)$. Thus $\{X_t = (X_t(1),\ldots,X_t(n))\}_{t\in\mathbb{N}_0}$ is a Markov chain with state space S_n. But we have to consider more information about the Rudvalis chain. We observe that when we place the top card in position n, every card is cyclically shifted counterclockwise and when we place the top card in position $n-1$, every card except the last is cyclically shifted (counterclockwise). Thus, at each time t, besides keeping track of the position of card k, $X_t(k)$, we will also count, with Y_t, the number of times the deck is shifted counterclockwise (modulo n). We shall consider the lifted chain $\{(X_t,Y_t)\}_{t\in\mathbb{N}_0}$ with state space $(S_n,\mathbb{Z}_n)$, where

$$\begin{cases} X_t(k) = \sigma_t^{-1}(k) \\ Y_t = t \pmod n. \end{cases}$$

With that in mind, given a card k, define

$$\Psi_k(X_t,Y_t) = v(X_t(k))e^{\frac{2\pi i}{n}Z_t(k)} \quad (2)$$

where $Z_t(k) = (X_t(k) - X_0(k) + Y_t) \pmod n$ and $v : \{1,\ldots,n\} \to \mathbb{C}$ is a function, to be determined

later, which will turn, for all $k \in \{1, \cdots, n\}$, Ψ_k into an eigenfunction (for the lifted chain) with the same eigenvalue λ. Furthermore, define

$$\Psi(X_t, Y_t) = \sum_{k=1}^{n} \Psi_k(X_t, Y_t)$$

which is an eigenfunction with eigenvalue λ, since $\{\Psi_k\}_{k=1}^n$ are eigenfunctions with the same eigenvalue λ. Furthermore, note that $|\Psi_k(X_t, Y_t)| = |v(X_t(k))| \left| e^{\frac{2\pi i}{n} Z_t(k)} \right| = |v(X_t(k))|$, which means that for fixed x, and for y_1, y_2, the eigenvector satisfies $|\Psi_k(x, y_1)| = |\Psi_k(x, y_2)| = |v(x)|$, hence $|\Psi(X_t, Y_t)|$ is a function of X_t only, as we wanted. Finally, the function f to consider in Proposition 2.4 is

$$\begin{aligned} f : S_n &\to \mathbb{C} \\ \sigma_t &\mapsto \Psi(X_t, Y_t). \end{aligned}$$

The next proposition determines the eigenvalue for the previously defined eigenvector.

Proposition 2.5. *The random walk performed by a card k under the lifted Rudvalis shuffle has an eigenvector of the form*

$$\Psi_k(x, y) = v(x) e^{2\pi i z/n},$$

where $v(x)$ is the x-th number in the list $\lambda^{n-2}, \ldots, \lambda^2, \lambda, 1, \chi$, the eigenvalue is

$$\lambda = 1 - \frac{p}{1-p} \frac{4\pi^2}{n^3} + O(\tfrac{1}{n^5}),$$

$\chi = 1 + \frac{p}{1-p} \frac{2\pi i}{n} + O(\frac{1}{n^2})$, *and* $z = x - k + y$.

In order to apply Proposition 2.4, we need to compute the expected value and variance of $\Psi(X_t, Y_t)$. The expectation can be easily computed, but the variance demands more effort, which is why we compute instead $\mathbb{E}[|\Psi(X_{t+1}) - \Psi(X_t)|^2]$ and use the following Lemma.

Lemma 2.6. *Let $\{X_t\}_{t \in \mathbb{N}_0}$ be a Markov chain with finite state space Ω, with eigenfunction Ψ and eigenvalue λ, and let μ be a probability distribution on Ω. If $\Re(\lambda) \geq \frac{1}{2}$ and $|\lambda| < 1$, then*

$$Var_\mu[\Psi(X_t)] \leq \frac{\mathbb{E}_\mu[|\Psi(X_{t+1}) - \Psi(X_t)|^2]}{2(1 - \Re(\lambda))}.$$

We need to compute the eigenvector at time $t = 0$ in order to get its expected value.

Lemma 2.7. *Let $\{(X_t, Y_t)\}_{t \geq 0}$ be the lifted Markov chain described above and let Ψ be the eigenvector with eigenvalue λ. Then,*

$$\Psi(X_0, Y_0) = n + O(\tfrac{1}{n}).$$

In particular, this value is independent from the initial configuration of the deck.

Now we can compute the expected value and the variance of $\Psi(X_t, Y_t)$.

Lemma 2.8. *Let (X_t, Y_t) be the lifted Markov chain described above, $\Psi = \sum_{k=1}^{n} \Psi_k$ the eigenvector (with eigenvalue λ) obtained from Proposition 2.5 and $\mu_t^\sigma := P^t(\sigma, \cdot)$. Then, for any $t \in \mathbb{N}_0$, the expectation and the variance of the complex random variable $\Psi(X_t, Y_t)$ under the probability distribution μ_t^σ are given, respectively, by*

$$E_{\mu_t^\sigma}[\Psi(X_t, Y_t)] = \lambda^t \Psi(X_0, Y_0) = \lambda^t n + O(\tfrac{1}{n})$$

and

$$Var_{\mu_t^\sigma}[\Psi(X_t, Y_t)] = O(n).$$

We can finally use (1) to find a lower bound for the mixing time of the Rudvalis shuffle.

Theorem 2.9 (Lower bound on the mixing time of the Rudvalis shuffle)**.** *For the Rudvalis shuffle with n cards,*

$$d(t) \geq 1 - 8e^{-2t \log|\lambda|} \frac{O(1)}{n + O(\frac{1}{n})},$$

where

$$\lambda = 1 - \frac{p}{1-p} \frac{4\pi^2}{n^3} + O(\tfrac{1}{n^5}).$$

Furthermore, for $\varepsilon \in (0, 1)$,

$$t_{mix}(\varepsilon) \geq \frac{1-p}{p} \frac{1}{8\pi^2} n^3 \log(n) + O(n^3).$$

Remark. *A generalization of the Rudvalis shuffle is the top to bottom-k shuffle which places the first card uniformly at random in one of the bottom k_n positions (for $k_n = 2$, we have the Rudvalis shuffle). In [3] we can see that if k_n is a constant, we would still have a lower bound (and upper bound) of order $n^3 \log(n)$.*

3. Hydrodynamic limit for the continuous time Rudvalis shuffle

Now that we have studied the discrete time Rudvalis Markov chain, it is time to extend it to continuous time. In this section, we focus on the case $p = \frac{1}{2}$ (in order to simplify the computations and notation). After defining the process, we will convert it into a particle system and study the hydrodynamic limit of this system. Moreover, we shall prove the existence and uniqueness of a (weak) solution of a PDE, the transport equation, and this is the main goal of this section.

As we will see, the hydrodynamic limit can be seen as the convergence (in probability) of a random measure to an absolutely continuous measure (with respect to the Lebesgue measure) whose density is the unique weak solution of a PDE. If we prove

this convergence together with the uniqueness of the solution, we will have proved that there exists a unique weak solution of that PDE [4].

To prove this convergence, the entropy method is employed. First, we prove that the sequence of probability measures (associated with the random measures of the density) is relatively compact, which tells us that every subsequence has a weakly convergent subsequence. Then, we characterize their limit points by showing that they are unique. Since all the (weakly) convergent subsequences converge to the same limit point, the whole sequence converges to that unique limit point.

3.1. Rudvalis shuffle

Let us define a continuous-time version of the Rudvalis Markov chain. Let us consider a Poisson process $\{\mathcal{T}_k : k \in \mathbb{N}\}$ of rate 1, that is, $\mathcal{T}_0 = 0$ and $\{(\mathcal{T}_k - \mathcal{T}_{k-1}) : k \in \mathbb{N}\}$ is a sequence of i.i.d. exponential random variables with mean 1 which is responsible for giving our process the Markov property. The process $\{\mathcal{T}_k : k \in \mathbb{N}\}$ is known as the *Harris process* or the *clock process* (see [5] for instance). It is important to observe that in Harris' construction, the probability that two clocks ring simultaneously is equal to 0. At each time a clock rings, we toss a fair coin (since we fixed $p = \frac{1}{2}$). If the coin lands heads up, then we remove the top card from the deck (the card at position 1) and insert it at position n. If the coin lands tails up, then we remove the top card from the deck and insert it at position $n-1$. In particular, every card performs a continuous-time random walk on the deck.

The continuous-time Rudvalis shuffle is the Markov process $\{\sigma_t : t \geq 0\}$ with state space S_n and *infinitesimal generator* (cf. Chapter 2 of [10]) given on $f : S_n \to \mathbb{R}$, by

$$\mathcal{L}_n^R f(\sigma) = \frac{1}{2} \sum_{x=n-1}^{n} \Theta_x f(\sigma),$$

where $\Theta_x f(\sigma) = f(\sigma^x) - f(\sigma)$.

Denote by $\mathbb{T} = \mathbb{R}/\mathbb{Z} = [0,1)$ the one-dimensional continuous torus (macroscopic space), by $\mathbb{T}_n = \mathbb{Z}/n\mathbb{Z} = \{0, 1, \ldots, n-1\}$ the one-dimensional discrete torus with n points (microscopic space) and by $\Omega_n = \{0,1\}^{\mathbb{T}_n}$ the space of functions from $\mathbb{T}_n$ to $\{0,1\}$.

Recall that $\sigma(x) = y$ if, and only if, card y is at position x in permutation σ. Now we will color the cards, so, in order to identify the color of the card at a position x, we define $\eta(x) = \mathbb{1}_{\{\sigma(x) \text{ is black}\}}$. Thus, the projection $\sigma \mapsto P(\sigma) = \eta$ induces a Markov process with state space Ω_n, which we shall call the *space of configurations*, and with infinitesimal generator given on $f : \Omega_n \to \mathbb{R}$, by

$$\mathcal{L}_n f(\eta) = \frac{1}{2} \sum_{x=n-1}^{n} \Theta_x f(\eta),$$

where $\Theta_x f(\eta) = f(\eta^x) - f(\eta)$ and $\eta^x = P(\sigma^x)$.

Now we can think of a simple, yet remarkable, mapping. Taking each position of the deck as a site, the black cards as particles and the red cards as empty sites, a particle system emerges!

Remark. *For $x \in \mathbb{T}_n$ and $\eta \in \Omega_n$, $\eta(x)$ is the number of particles at site x:*

$$\eta(x) = \begin{cases} 1, & \text{if card } x \text{ is black (site } x \text{ is occupied)} \\ 0, & \text{if card } x \text{ is red (site } x \text{ is empty)} \end{cases}.$$

Example 3.1. *Consider a simple example of a deck of $n = 5$ cards.*

Figure 3: The configuration of the particle system corresponding to a deck with 5 cards.

The top card of the deck, the 3 of hearts, is red, so we do not assign a particle to site 1, the king of clubs is black, so we assign a particle to site 2. We keep doing this procedure until there are no more cards. Moreover, note that in Figure 2 we can see a particle system in motion.

Our goal is to study the evolution in space and time of this system. In order to restrict the system to $\mathbb{T}$, we shall consider the process evolving in $\mathbb{T}_n/n$ and then we will take $n \to \infty$. We can identify the microscopic space with $\mathbb{T}_n$ (discrete) and the macroscopic space with $\mathbb{T}$ (continuous). A site x in the microscopic space $\mathbb{T}_n$, can be identified with $\frac{x}{n}$ in the macroscopic space $\mathbb{T}$. This means that the particles move between the sites $0, \frac{1}{n}, \frac{2}{n}, ..., 1$. If we let n go to ∞, then we can think of an initial density of particles, denoted by $\gamma : \mathbb{T} \to [0,1]$. Moreover, if we let the system evolve through time, we can wonder what function describes the density of particles at each time, given that the initial density is γ. As we will see, this function will be given by a weak solution (proved to be unique) of some PDE. This kind of behaviour of the density is the so-called hydrodynamic limit.

Observe that applying the generator $\mathcal{L}_n^R$ to functions that only depend on the colors of the cards is equivalent to applying the generator $\mathcal{L}_n$. If we want to see a non trivial evolution of the system

when $n \to \infty$, speeding up the process is absolutely necessary. For this reason, from now on we will speed up the process in the hyperbolic time scale tn.

We will still denote by $\{\eta_t\}_{t\geq 0}$ the Markov process with generator $n\mathcal{L}_n$ and denote the initial measure by μ_n. Owing to the fact that it was derived from the Rudvalis card shuffle, we call the process $\{\eta_t\}_{t\geq 0}$ the Rudvalis process. In order to study this process, we start by defining the space where it lives.

Definition 3.1 (Space $D_X[0,T]$)**.** Let $T > 0$ and X be a metric space. Denote by $D_X[0,T]$ the space of functions $x : [0,T] \to X$ that are right-continuous and have left-hand limits:

(i) for $0 \leq t < T$, $x(t^+) = \lim_{s\downarrow t} x(s)$ exists and $x(t^+) = x(t)$;

(ii) for $0 < t \leq T$, $x(t^-) = \lim_{s\uparrow t} x(s)$ exists;

and which is endowed with the Skorohod topology (cf. Chapter 3 of [1]).

Remark. *We call the elements of $D_X[0,T]$ trajectories and denote them by $x_\cdot$, i.e., for each $t \in [0,T]$, we denote $x(t)$ by x_t.*

The trajectories of the Rudvalis process $\{\eta_t\}_{t\geq 0}$ belong to $D_{\Omega_n}[0,T]$, i.e., for each time t, η_t is an element of Ω_n. Moreover, the function $t \mapsto \eta_t$ is càdlàg (it has jumps).

Let $\mathbb{P}_{\mu_n}$ be the probability measure in $D_{\Omega_n}[0,T]$ induced by the Rudvalis process with initial measure μ_n. Denote by $\mathbb{E}_{\mu_n}$ the expectation with respect to $\mathbb{P}_{\mu_n}$.

3.2. Initial measures

We have an initial profile, that we assume to be measurable, denoted by $\gamma : \mathbb{T} \to [0,1]$, i.e., γ is a function which assigns mass to points on the torus.

Definition 3.2 (Sequence of measures associated with a profile)**.** A sequence of probability measures $\{\mu_n\}_{n\in\mathbb{N}}$ on Ω_n is said to be associated with a measurable profile $\gamma : \mathbb{T} \to [0,1]$ if for every $\delta > 0$ and every continuous function $G : \mathbb{T} \to \mathbb{R}$

$$\lim_{n\to\infty} \mu_n\Big(\eta \in \Omega_n : \Big|\frac{1}{n}\sum_{x\in\mathbb{T}_n} G\left(\tfrac{x}{n}\right)\eta(x) - \int_{\mathbb{T}} G(u)\gamma(u)\,du\Big| > \delta\Big) = 0.$$

Example 3.2. *Let $\gamma : \mathbb{T} \to [0,1]$ be continuous. An example of a sequence of measures satisfying the assumption above is $\{\nu^n_{\gamma(\cdot)}\}_{n\in\mathbb{N}}$, where $\nu^n_{\gamma(\cdot)}$ is the Bernoulli product measure associated with γ:*

$$\nu^n_{\gamma(\cdot)}(\eta) = \prod_{x\in\mathbb{T}_n} [\eta(x)\gamma(\tfrac{x}{n}) + (1-\eta(x))(1-\gamma(\tfrac{x}{n}))]$$

which means that this measure assigns a particle independently to each site x with probability $\gamma(\frac{x}{n})$.

Since the Bernoulli product measures are quite simple, we can take them as initial distributions when one wants to obtain scaling limits.

3.3. Empirical measure

We have already seen how to obtain a particle system from the Rudvalis process. Now it is time to study the evolution of the Rudvalis process $\{\eta_t\}_{t\geq 0}$ with infinitesimal generator $n\mathcal{L}_n$ and state space Ω_n.

Let $\mathcal{M}$ be the set of positive measures on $\mathbb{T}$ with mass bounded by 1 endowed with the weak topology. If $\{\pi^n\}_{n\in\mathbb{N}}$, $\pi \in \mathcal{M}$, we say that π^n converges weakly to π, which we denote by $\pi^n \xrightarrow[n\uparrow\infty]{w} \pi$, if for all $G \in C(\mathbb{T})$

$$\int_{\mathbb{T}} G\ \pi^n(du) \xrightarrow[n\uparrow\infty]{} \int_{\mathbb{T}} G\ \pi(du).$$

Given a function $G : \mathbb{T} \to \mathbb{R}$, we denote the integral of G with respect to a measure μ (on $\mathbb{T}$) by $\langle \mu, G\rangle := \int G(u)\ \mu(du)$.

In order to study a scaling limit of the Rudvalis process, we introduce a measure which gives weight $\frac{1}{n}$ to each occupied site of the configuration η.

Definition 3.3 (Empirical Measure)**.** For each configuration $\eta \in \Omega_n$, we define the empirical measure $\pi^n(\eta, du)$ in $[0,1]$ by $\pi^n(\eta, du) = \frac{1}{n}\sum_{x\in\mathbb{T}_n} \eta(x)\delta_{\frac{x}{n}}(du)$, where δ_y is the Dirac measure concentrated on $y \in \mathbb{T}$.

Since we are interested in studying the behaviour of this measure with respect to the Rudvalis process $\{\eta_t\}_{t\geq 0}$, define

$$\pi^n_t(du) := \pi^n(\eta_t, du) = \frac{1}{n}\sum_{x\in\mathbb{T}_n} \eta_t(x)\delta_{\frac{x}{n}}(du) \in \mathcal{M}.$$

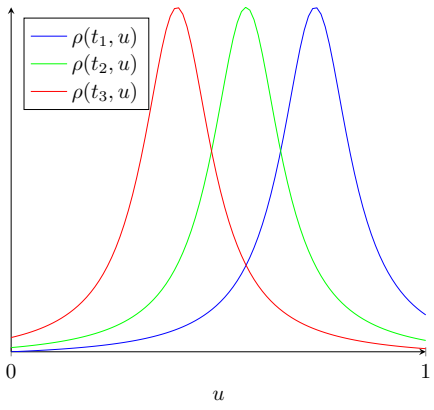
3.4. Transport equation

From now on, fix $T > 0$. When we think about the Rudvalis shuffle, the behaviour of the cards is quite similar to a wave: we take the top card and put it in the bottom (or almost at the bottom), pushing all the cards one position up. We shuffle the deck over and over again, that is, the cards are being pushed up continuously (Figure 2). With that picture in mind, we analyze the following PDE: the *transport equation* (sometimes called the one-way wave equation), given formally by

$$\begin{cases} \partial_t\rho(t,u) = \nabla\rho(t,u), & \text{for } t \in [0,T], u \in \mathbb{T}, \\ \rho(0,u) = \gamma(u), & \text{for } u \in \mathbb{T}, \end{cases} \tag{3}$$

where ∇ is the space derivative and $\gamma : \mathbb{T} \to [0,1]$ is a measurable function.

Observe that the solutions of this equation are functions that behave like "waves" moving to the left, i.e., if γ is C^1 in space, then the function $\rho(t,u) = \gamma(u+t)$ is a (strong) solution of (3). For example, we would have the following graph for a solution of (3) captured at three distinct times $0 < t_1 < t_2 < t_3 < T$.



3.5. Motivation for the weak solution

Define $C^{2,1}([0,T]\times\mathbb{T})$ as the space of functions $f:[0,T]\times\mathbb{T}\to\mathbb{R}$ with first derivative in time and second derivative in space, both being continuous and let $f \in C^{2,1}([0,T]\times\mathbb{T})$. Multiplying both sides of the equation $\partial_s\rho(s,u) = \partial_u\rho(s,u)$ by $f(s,u)$ and integrating it on time $[0,t]$, for $t<T$, and on space $\mathbb{T}$, and, applying Fubini's theorem and using the initial condition $\rho(0,u) = \gamma(u)$, we can define the concept of *weak solution* for the transport equation (3).

Definition 3.4 (Weak solution of the transport equation on $\mathbb{T}$)**.** Let $\gamma:\mathbb{T}\to[0,1]$ be a measurable function and $T>0$. A measurable function $\rho:[0,T]\times\mathbb{T}\to[0,1]$ is said to be a weak solution of equation (3) if for all $f\in C^{2,1}([0,T]\times\mathbb{T})$ and all $t\in[0,T]$, ρ satisfies

$$\int_{\mathbb{T}}\rho(t,u)f(t,u)-\gamma(u)f(0,u)\,du-\int_0^t\int_{\mathbb{T}}\rho(s,u)\partial_s f(s,u)\,du\,ds + \int_0^t\int_{\mathbb{T}}\rho(s,u)\partial_u f(s,u)\;du\;ds = 0.$$

3.6. Heuristic argument for the weak solution

As stated above, our goal is to prove that the density of particles is ruled by the weak solution of a PDE, in our case, the transport equation. In order to do that, we work with the microscopic space $\mathbb{T}_n$ and then take the limit on the number of sites, n. For that reason, it is useful to define the notion of derivative in the discrete case.

Definition 3.5 (Discrete left derivative)**.** Given $G:\mathbb{T}\to\mathbb{R}$, $n\in\mathbb{N}$ and $x\in\mathbb{T}_n$, the discrete left derivative at $\frac{x}{n}$ is defined by

$$\nabla_n^- G\left(\tfrac{x}{n}\right) = n\left(G\left(\tfrac{x}{n}\right) - G\left(\tfrac{x-1}{n}\right)\right).$$

Computing the generator on $\langle\pi_t^n, G\rangle$ is one of the first steps when one wants to find how the empirical measure evolves in space/time. The following lemma is useful in that regard.

Lemma 3.3. *Let $\{\eta_t\}_{t\geq0}$ be the Markov process with generator $n\mathcal{L}_n$ (Rudvalis process). Then, for any $t\geq0$ and $G:\mathbb{T}\to\mathbb{R}$*

$$n\mathcal{L}_n\langle\pi_t^n,G\rangle = -\langle\pi_t^n,\nabla_n^- G\rangle - \frac{1}{2n}(\eta_t(1)-\eta_t(0))\nabla_n^- G(0).$$

Now fix $f\in C^{2,1}([0,T]\times\mathbb{T})$ and apply Dynkin's formula (Appendix 1.5 of [8]) with $F(t,\eta_t) = \langle\pi_t^n, f_t\rangle$. Define

$$\begin{aligned} M_t^n(f) &:= \langle\pi_t^n,f_t\rangle - \langle\pi_0^n,f_0\rangle - \int_0^t(\partial_s + n\mathcal{L}_n)\langle\pi_s^n,f_s\rangle\;ds,\\ N_t^n(f) &:= (M_t^n(f))^2 - \int_0^t\Gamma^n(\langle\pi_s^n,f_s\rangle)\;ds, \end{aligned} \tag{4}$$

where $\Gamma^n(\langle\pi_s^n,f_s\rangle) := n\mathcal{L}_n\langle\pi_s^n,f_s\rangle^2 - 2\langle\pi_s^n,f_s\rangle n\mathcal{L}_n\langle\pi_s^n,f_s\rangle$ is the *carré du champ* operator. Then, $\{M_t^n\}_{t\geq0}$ and $\{N_t^n\}_{t\geq0}$ are martingales with respect to the natural filtration $\mathcal{F}_t = \sigma(\eta_s : s\leq t)$.

Using Lemma 3.3, we can rewrite the first martingale as

$$\begin{aligned} M_t^n(f) = \langle\pi_t^n,f_t\rangle - \langle\pi_0^n,f_0\rangle - \int_0^t\partial_s\langle\pi_s^n,f_s\rangle\;ds \\ + \int_0^t\langle\pi_s^n,\nabla_n^- f_s\rangle\;ds + \frac{1}{2n}\int_0^t(\eta_s(1)-\eta_s(0))\nabla_n^- f_s(0)\;ds. \end{aligned} \tag{5}$$

Since the expected value of a martingale remains constant and $M_0^n(f)=0$, we have for all t that $\mathbb{E}_{\mu_n}[M_t^n(f)] = 0$. Consider the function $\rho_t^n(x) := \mathbb{E}_{\mu_n}[\eta_t(x)]$. The martingale gives us an idea (in the discrete case) of the condition satisfied by the solution of the PDE which we are expecting. Taking the expectation with respect to $\mathbb{P}_{\mu_n}$ in (5), we get the following equation for ρ_t^n

$$\begin{aligned} 0 = \frac{1}{n}\sum_{x\in\mathbb{T}_n} f_t\left(\tfrac{x}{n}\right)\rho_t^n(x) - \frac{1}{n}\sum_{x\in\mathbb{T}_n} f_0\left(\tfrac{x}{n}\right)\rho_0^n(x) \\ - \int_0^t\frac{1}{n}\sum_{x\in\mathbb{T}_n}\partial_s f_s\left(\tfrac{x}{n}\right)\rho_s^n(x)\;ds \\ + \int_0^t\frac{1}{n}\sum_{x\in\mathbb{T}_n}\nabla_n^- f_s\left(\tfrac{x}{n}\right)\rho_s^n(x)\;ds + O(\tfrac{1}{n}). \end{aligned}$$

If we look carefully and assume that $|\rho_t^n(x) - \rho_t(\frac{x}{n})| \xrightarrow[n\uparrow\infty]{} 0$, these are Riemann sums for the integrals in Definition 3.4.

3.7. Hydrodynamic limit

Now that we have seen an heuristic argument to obtain a weak solution for equation (3), let us see how to prove rigorously the hydrodynamic limit for

the dynamics of the Rudvalis process. Recall Definition 3.1 and the space $D_{\Omega_n}[0,T]$, where $\mathbb{P}_{\mu_n}$ is defined. The process $\{\pi_t^n\}_{t\geq 0}$, induced by the Rudvalis process, can be seen as a Markov process, but this time in $D_{\mathcal{M}}[0,T]$ endowed with the Skorohod topology (cf. Chapter 3 of [1]).

Let $\{\mathbb{Q}^n\}_{n\in\mathbb{N}}$ be the sequence of probability measures in $D_{\mathcal{M}}[0,T]$ induced by the process $\{\pi_t^n\}_{t\geq 0}$ and by the probability measure $\mathbb{P}_{\mu_n}$. Equivalently, $\mathbb{Q}^n$ is induced by the application

$$\begin{aligned}\pi^n : (D_{\Omega_n}[0,T], \mathbb{P}_{\mu_n}) &\to (D_{\mathcal{M}}[0,T], \mathbb{Q}^n)\\ \eta_\cdot &\to \pi^n(\eta_\cdot, du).\end{aligned}$$

As we will see, the sequence $\{\mathbb{Q}^n\}_{n\in\mathbb{N}}$ of probability measures converges weakly to a limit point.

Recall the concept of a sequence of measures associated with a profile, given in Definition 3.2. We now present the main result of this section.

Theorem 3.4 (Hydrodynamic Limit)**.** *Consider the Rudvalis process $\{\eta_t\}_{t\geq 0}$. Fix a measurable profile $\gamma : \mathbb{T} \to [0,1]$ and let $\{\mu_n\}_{n\in\mathbb{N}}$ be a sequence of probability measures associated with γ. Then, for any $t \in [0,T]$, for any $\delta > 0$ and any $G \in C(\mathbb{T})$, it holds*

$$\lim_{n\to\infty} \mathbb{P}_{\mu_n}\Big(\eta_\cdot \in D_{\Omega_n}[0,T] : \Big|\frac{1}{n}\sum_{x\in\mathbb{T}_n} G\left(\frac{x}{n}\right)\eta_t(x) - \int_{\mathbb{T}} G(u)\rho(t,u)\, du\Big| > \delta\Big) = 0, \tag{6}$$

where $\rho(t,\cdot)$ is the unique weak solution of (3) with initial condition $\rho(0,\cdot) := \gamma(\cdot)$.

Remark. *We call equation (3) the hydrodynamic equation of the Rudvalis process.*

With respect to $\mathbb{P}_{\mu_n}$, the measure induced by the Rudvalis process with infinitesimal generator $n\mathcal{L}_n$ and initial measure μ_n, verifies $\pi_t^n \xrightarrow{w} \pi_t$ so another way to express the limit in (6) is to say that π_t^n converges in probability to π_t (because the convergence is only verified under $\mathbb{P}_{\mu_n}$). The hydrodynamic limit is a consequence of the following proposition.

Proposition 3.5. *Let $\gamma : \mathbb{T} \to [0,1]$ be measurable and consider a sequence of measures $\{\mu_n\}_{n\in\mathbb{N}}$ on Ω_n associated with γ. Let $\mathbb{Q}^*$ be the probability measure concentrated on a trajectory $\pi_\cdot$ of $D_{\mathcal{M}}[0,T]$ consisting of absolutely continuous measures with respect to the Lebesgue measure, i.e., $\pi_t(du) = \rho_t(u)du$ where the density $\rho_\cdot$ is the unique weak solution of (3). Then*

$$\mathbb{Q}^n \xrightarrow{w} \mathbb{Q}^*.$$

Since $\mathbb{Q}^n$ is the measure induced by $\mathbb{P}_{\mu_n}$ and the application π^n, the weak convergence of $\mathbb{Q}^n$ to $\mathbb{Q}^*$ is the same as the convergence in distribution of π_t^n to π_t, for any $t \in [0,T]$. Since $\pi_t(du) = \rho_t(u)du$ is a deterministic measure, the previous convergence is also valid in probability (with respect to $\mathbb{P}_{\mu_n}$) and whence, by the previous observations, Proposition 3.5 implies Theorem 3.4. For this reason, when we want to prove the hydrodynamic limit, we often follow the sequence of steps given below:

1. The sequence $\{\mathbb{Q}^n\}_{n\in\mathbb{N}}$ is relatively compact in Skorohod's topology(Proposition 1.2 of [8]).
2. The limit points of the subsequences of $\{\mathbb{Q}^n\}_{n\in\mathbb{N}}$ are concentrated on trajectories of measures which are absolutely continuous with respect to the Lebesgue measure and whose densities $\rho(t,u)$ are weak solutions of the hydrodynamic equation.
3. The hydrodynamic equation has a unique weak solution.

3.7.1 Relative compactness

The concepts of tightness and relatively compactness are related in a well know theorem by Prohorov (Theorem 5.1 and Theorem 5.2 of [1]).

When proving the tightness of $\{\mathbb{Q}^n\}_{n\in\mathbb{N}}$, we will use Dynkin's formula as in Section 3.6. Moreover, the following lemma will be useful. But before, let us just introduce the next notation (similar to the big O). For two sequences $\{f_k\}_{k\in\mathbb{N}}$ and $\{g_k\}_{k\in\mathbb{N}}$, we say that $f_n \lesssim g_n$ if there exists a positive constant C, independent of n, such that $f_n \leq C\ g_n$. Moreover, given a space X and $G : X \to \mathbb{R}$, denote the $L^\infty(X)$-norm by $\|G\|_\infty = \sup_{x\in X}|G(x)|$.

Lemma 3.6. *Recall the operator Γ^n defined in (4). Let $G \in C^1(\mathbb{T})$. Then,*

$$\Gamma^n(\langle \pi_s^n, G\rangle) \lesssim \frac{\|\nabla G\|_\infty^2}{n}.$$

Proposition 3.7. *The sequence $\{\mathbb{Q}^n\}_{n\in\mathbb{N}}$ is relatively compact.*

The next step is to characterize all possible limit points, i.e., to find their properties.

3.7.2 Characterization of limit points

Having proven that $\{\mathbb{Q}^n\}_{n\in\mathbb{N}}$ is relatively compact, it is time to characterize the limit points of the subsequences. We divide this proof into two parts. First, we show that any limit point $\mathbb{Q}^*$ is concentrated on trajectories (in $D_{\mathcal{M}}[0,T]$) which are absolutely continuous with respect to the Lebesgue measure. Then, we prove that the density ρ is a weak solution of (3).

The following lemma gives a sufficient condition for a measure to be absolutely continuous with respect to the Lebesgue measure.

Lemma 3.8. *Let μ be a measure which satisfies $|\langle \mu, G\rangle| \leq \int_0^1 |G(u)|\ \lambda(du)$ for all $G \in C([0,1])$. Then, $\mu \ll \lambda$, where λ is the Lebesgue measure.*

We prove that the sufficient condition of the lemma holds for trajectories on which $\mathbb{Q}^*$ is concentrated and we can conclude the following.

Proposition 3.9. *Let $\mathbb{Q}^*$ be a limit point of a subsequence of the sequence $\{\mathbb{Q}^n\}_{n\geq 1}$. Then, $\mathbb{Q}^*$ is concentrated on trajectories of measures absolutely continuous with respect to the Lebesgue measure.*

The previous result means that, for each t, $\pi_t(du) = \rho(t,u)du$ where ρ is the density of π.

For the second part, we prove rigorously the ideas explored before in the heuristic argument (Section 3.6).

Proposition 3.10. *Let $\mathbb{Q}^*$ be a limit point of a subsequence of $\{\mathbb{Q}^n\}_{n\geq 1}$ and assume without loss of generality that $\mathbb{Q}^n \xrightarrow{w} \mathbb{Q}^*$. Let $f \in C^{2,1}([0,T]\times\mathbb{T})$ and define*

$$A = \{\pi_\cdot \in D_{\mathcal{M}}[0,T] : \sup_{t\in[0,T]} \langle \rho_t, f_t\rangle - \langle \gamma, f_0\rangle - \int_0^t \langle \rho_s, \partial_s f_s\rangle\ ds + \int_0^t \langle \rho_s, \nabla f_s\rangle\ ds = 0\}. \tag{7}$$

Then, $\mathbb{Q}^*$ is concentrated on A, that is, $\mathbb{Q}^*(A) = 1$.

We still might have different subsequences of $\{\mathbb{Q}^n\}_{n\in\mathbb{N}}$ converging to limit points concentrated on trajectories whose densities might not be the same. It is thus left to show the uniqueness of the weak solution of (3).

3.7.3 Uniqueness of weak solutions of the transport equation

Recall Definition 3.4. Assume that $\rho^l : [0,T]\times\mathbb{T} \to [0,1]$, $l = 1,2$, are two weak solutions of (3) starting from the same initial condition. Define $\hat{\rho} = \rho^1 - \rho^2$ and observe that $\hat{\rho}(0,u) = 0$ for any $u \in \mathbb{T}$. For $t \in [0,T]$ and $u \in \mathbb{T}$, let $\phi_m(t) = \frac{1}{\sqrt{T}} e^{2\pi i m t/T}$, $\xi_k(u) = e^{2\pi i k u}$ and $\psi_{m,k}(t,u) = \phi_m(t)\xi_k(u) = \frac{1}{\sqrt{T}} e^{2\pi i (mt/T + ku)}$. Then, $\{\psi_{m,k} : m,k \in \mathbb{N}_0\}$ is an orthonormal basis of $L^2([0,T]\times\mathbb{T})$ (cf. Chapter 7 of [7]) for the inner product $\langle f, g\rangle = \int_0^T \int_{\mathbb{T}} f(t,u)\overline{g(t,u)}\ du\ dt$, where $\overline{g}$ is the conjugate of g. Moreover, for any $m,k \in \mathbb{N}_0$ we have $\partial_s \psi_{m,k}(s,u) = 2\pi i \frac{m}{T} \psi_{m,k}(s,u)$ and $\nabla \psi_{m,k}(s,u) = \frac{2\pi i k}{\sqrt{T}} \psi_{m,k}(s,u)$. Therefore, replacing f_s by $\psi_{m,k}(s,\cdot)$ in Definition 3.4 for fixed m and k, we obtain

$$\int_{\mathbb{T}} \hat{\rho}(t,u)\psi_{m,k}(t,u)du = 2\pi i \left(\tfrac{m}{T} - k\right) \int_0^t \int_{\mathbb{T}} \hat{\rho}(s,u)\psi_{m,k}(s,u)\, du\, ds.$$

Taking a time derivative on both sides of the above expression, we conclude that $x_{m,k}(t) = \int_{\mathbb{T}} \hat{\rho}(t,u)\psi_{m,k}(t,u)\, du$ is a solution of the first order linear ordinary differential equation $y'(t) - 2\pi i \left(\frac{m}{T} - k\right) y(t) = 0$, that is, $x_{m,k}(t) = x_{m,k}(0)\, e^{2\pi i (m/T - k)t} = 0$. Recall that we can write $\hat{\rho}(t,u) = \sum_{m=0}^{\infty}\sum_{k=0}^{\infty} c_{m,k}\, \psi_{m,k}(t,u)$, where $c_{m,k} = \langle \hat{\rho}, \psi_{m,k}\rangle = 0$. Therefore we conclude that $\hat{\rho}(t,u) = 0$ a.s. (almost surely), i.e., $\rho^1 = \rho^2$ a.s., which proves the uniqueness of weak solution of (3).

References

[1] P. Billingsley. *Convergence of Probability Measures*. Wiley Interscience, 1999.

[2] P. Diaconis. Group representations in probability and statistics. *Lecture notes-monograph series*, 1988.

[3] S. Goel. Analysis of top to bottom-k shuffles. *The Annals of Applied Probability*, 16:30–55, 2006.

[4] M. Guo, G. Papanicolaou, and S. Varadhan. Nonlinear diffusion limit for a system with nearest neighbor interactions. *Communications in Mathematical Physics*, 118(1):31–59, 1988.

[5] T. Harris. Nearest-neighbor markov interaction processes on multidimensional lattices. *Advances in Mathematics*, 9:66 – 89, 1972.

[6] M. V. Hildebrand. *Rates of convergence of some random processes on finite groups*. Harvard University, 1990.

[7] J. K. Hunter and B. Nachtergaele. *Applied analysis*. World Scientific Publishing Company, 2001.

[8] C. Kipnis and C. Landim. *Scaling limits of interacting particle systems*. Springer-Verlag, Berlin, 1999.

[9] D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov chains and mixing times*. American Mathematical Soc., 2017.

[10] T. M. Liggett. *Continuous time Markov processes: an introduction*. American Mathematical Soc., 2010.

[11] D. Wilson. Mixing time of the rudvalis shuffle. *Electronic Communications in Probability*, 8:77–85, 2003.