The hydrodynamic limit of a crystal surface jump diffusion with Metropolis-type rates

by

Anya Katsevich

Senior Honors Thesis
Department of Mathematics
University of North Carolina at Chapel Hill

March 30, 2017

Approved: [signature]

Dr. Jeremy Marzuola, Thesis Advisor

Dr. Amarjit Budhiraja, Reader

Dr. Katie Newhall, Reader

1 Introduction

Crystal surface diffusion refers to the way in which atoms on the surface of a crystal redistribute to eventually settle into a configuration with minimal surface energy. Along with epitaxy, or crystal growth, crystal surface diffusion is important to study due to the role it plays in the production of thin films, which have wide-ranging applications in microelectronics. For example, the deformation of a crystal surface to an equilibrium state plays a central role in fuel cells that rely on thin crystal films, as the conversion efficiency of chemical energy to electricity depends on the surface configuration of the film. As is characteristic of large microscopic systems, we can gain more insight into the nature of the dynamics of surface diffusion by studying it at the macroscopic level than at the level of individual atoms. Since the physical process is microscopic, however, a faithful mathematical model of the diffusion should describe it with microscopic dynamics. Given a model of the microscopic dynamics, then, we are presented with the challenge of deriving macroscopic dynamics in the limit as the number of particles approaches infinity. This is known as a hydrodynamic, or scaling, limit; it is particularly appealing from the modeling perspective because the input is the true, microscopic dynamics, while the output is a much easier to analyze continuum equation. The main goal of this work is to derive such a scaling limit for a specific dynamics governing the microscopic process of crystal surface diffusion.

1.1 Acknowledgments

I would like to thank Professor Jeremy Marzuola for his guidance and dedication throughout the one and a half years he has advised me. This thesis would not have come about if it weren't for the many engaging and helpful conversations we had about concepts ranging in scope from broad ideas to the nitty gritty details of local limit theorem proofs. I owe my enthusiasm to this topic and to the research process as a whole to Professor Marzuola. I would also like to thank Professor Budhiraja for discussions which helped me understand the problem from a more probabilistic point of view.

2 Background

One way to model a system of particles is to track the position and momentum of each one and use equations of motion to predict the variables' change over time. However, it is often more convenient to take the statistical mechanics viewpoint, introducing randomness into the system. According to this viewpoint, one divides the space into microscopically-sized cells and represents the system through field variables, which measure the average number and momentum of particles in each cell. These variables are taken to be random, to account for fluctuations resulting from particle movement between cells. Thus, the state of a microscopic system is best viewed as a probability measure: the set of configurations and their associated probabilities. For example, global equilibrium corresponds to the maximum entropy probability distribution. This is the state of greatest disorder, which we should expect the distribution of particles of the crystal surface to converge to with time (in accordance with the second law of thermodynamics).

As the number of cells approaches infinity, the microscopic system becomes increasingly well-represented by a continuous, macroscopic system, and the random dynamics by a deterministic one (an evolution equation). The central principle that hydrodynamic limit proofs rely on to "smooth out" random fluctuations in the limit is that microscopic systems quickly (on a macroscopic time scale) reach a local equilibrium state (i.e. local equilibrium measure). In this state, the particles in small but macroscopic regions are distributed according to a homogeneous equilibrium-like distribution that smoothly varies across space (see e.g. the introduction of [2] for an illustration of this idea). Because the behavior is homogeneous in each region, we may represent the local state by a single number without loss of information. This is the local analogue of assigning single thermodynamic quantities like temperature or pressure to an entire gas in equilibrium. However, determining this local equilibrium measure and rigorously justifying that the microscopic process takes on this distribution is challenging. The measure is tied to a certain local surface tension function σ, which is particularly sensitive to the way the microscopic model is defined. This function also appears in the PDE governing the macroscopic dynamics of surface diffusion processes, making it the link between microscopic and macroscopic dynamics.

A macroscopic crystal surface will be represented by a height profile, i.e. a function $h(x), x \in \mathbb{T}$, where $\mathbb{T} = [0,1]$ with endpoints identified. A microscopic surface will be represented by a function $h_N(j), j \in \mathbb{Z}/N\mathbb{Z}$, i.e. as N regularly spaced columns of atoms stacked above or below a substrate. The heights may either be restricted to the integers (to model columns of integer numbers of atoms), or allowed to be continuous. In [4], Marzuola restricts the heights to the integers, and models the diffusion as a Markov jump process with transitions occuring when atoms jump to neighboring lattice sites. In [5], Nishikawa considers continuous heights and models the microscopic diffusion with a stochastic differential equation. Both obtain

$$\partial_t h(t,x) = -\Delta \mathrm{div}\left[\nabla\sigma(\nabla h(t,x))\right] \tag{1}$$

as the PDE governing the evolution of the macroscopic profile h, albeit with differing definitions of σ, which will be discussed in this thesis.

This equation can be seen as a gradient descent in the free energy of the surface $h(t,\cdot)$ in $H^{-1}(\mathbb{T})$, which turns out to be the natural space to consider such a diffusion process in. The inner product of the space is given by $(f,g)_{H^{-1}} = ((-\Delta)^{-1}f, g)_{L^2}$. Letting $\frac{\delta\Sigma(h)}{\delta h}$ represent the functional derivative of h in H^{-1} and ϕ be a test function, we have

$$\begin{aligned}\int_{\mathbb{T}} (-\Delta)^{-1}\frac{\delta\Sigma(h)}{\delta h}\phi(x)dx &= \left(\frac{\delta\Sigma(h)}{\delta h}, \phi\right)_{H^{-1}} \\ &= \lim_{\epsilon\to 0}\frac{\Sigma(h+\epsilon\phi)-\Sigma(h)}{\epsilon} \\ &= \lim_{\epsilon\to 0}\int_{\mathbb{T}}\frac{\sigma(\nabla(h(x)+\epsilon\phi(x)))-\sigma(\nabla(h(x)))}{\epsilon}dx \\ &= \int_{\mathbb{T}}\nabla\sigma(\nabla h(x))\cdot\nabla\phi(x)dx = -\int_{\mathbb{T}}\mathrm{div}\left[\nabla\sigma(\nabla h(x))\right]\phi(x)dx,\end{aligned} \tag{2}$$

from which we see that $\frac{\delta\Sigma(h)}{\delta h} = \Delta\mathrm{div}\left[\nabla\sigma(\nabla h(x))\right]$.

In this thesis, we consider the setup in [4] (discrete heights) but a different set of transition rates, and derive the hydrodynamic limit. The transition rates we consider belong to a class of rates that depend only on the energy difference from before to after the jump; such rates, known as Metropolis-type, are important to study because they are commonly used by computational chemists in numerical simulations. The basic framework of the PDE derivation follows that of Marzuola in [4]. Both this work and that of Marzuola are based on the paper of Krug [3], who used physical arguments to derive PDEs such as (1) based on both adatom and Metropolis rates.

The rest of the thesis is organized as follows. In Section 3, we describe the microscopic space and microscopic dynamics. In Section 4, we explain the particular scaling limit we consider and derive the PDE governing the continuum evolution, postponing some calculations to Section 6. In Section 5, we present our numerical results and compare them with the theory.

3 The microscopic model

We index the microscopic spaces by N; the Nth microscopic configuration space of crystal surfaces is represented by height profiles of integer numbers of particles above and below the periodic lattice $\mathbb{Z}/N\mathbb{Z}$. We will denote a generic profile by $\mathbf{h} = (h_1, ..., h_N)$. Since the total number of particles is preserved by a diffusion process, we restrict the configuration space to

$$\mathbb{Z}_N^m := \{\mathbf{h} : h_i \in \mathbb{Z}, i = 1, ..., N, \sum_{i=1}^N h_i = m(N)\},$$

for some fixed $m(N)$. The height gradients corresponding to a profile $\mathbf{h}$ are denoted by $z_i = h_{i+1} - h_i, i = 1, ..., N$ (with circular indexing). The gradient values, rather than the heights themselves, play the key role in driving the diffusion process to equilibrium, via the profile's Hamiltonian,

$$H(\mathbf{h}) = \sum_{i=1}^N z_i^2.$$

We assume the system is in contact with a heat bath at constant temperature T; thus, the ensemble on $\mathbb{Z}_N^m$ representing thermal equilibrium is the canonical ensemble. It is given by

$$p_0(\mathbf{h}) = \frac{e^{-KH(\mathbf{h})}}{\mathcal{Z}},$$

where $K = \frac{1}{k_\beta T}$, k_β is the Boltzmann constant, and $\mathcal{Z}$ is the normalization constant.

The diffusion on $\mathbb{Z}_N^m$ is a Markov jump process; we denote it by $\mathbf{h}_N^t = (h_1(t), ..., h_N(t))$. Transitions between states occur when the top particle at a lattice site jumps to a neighboring lattice site with a certain instantaneous transition probability, or rate. Such a transition can be represented via the operator J_i^k, defined by

$$(J_i^k\mathbf{h})(j) = \begin{cases} h(j) - 1 & j = i \\ h(j) + 1 & j = k \\ h(j) & j \neq i, k \end{cases} \tag{3}$$

With this notation, the permissible transitions are

$$\mathbf{h} \mapsto J_i^k\mathbf{h}, |i - k| = 1.$$

Note that these transitions preserve the total number of particles m. The transition rates are defined by

$$r_N^{\pm}(i, \mathbf{h}) = \lim_{t\to 0} \frac{1}{t} P(\mathbf{h}_N^t = J_i^{i\pm 1}\mathbf{h} | \mathbf{h}_N^0 = \mathbf{h}).$$

To ensure that the process is reversible and invariant with respect to the equilibrium measure p_0, the rates must satisfy detailed balance, i.e. we must have

$$\begin{aligned} r_N^+(i, \mathbf{h})p_0(\mathbf{h}) &= r_N^-(i + 1, J_i^{i+1}\mathbf{h})p_0(J_i^{i+1}\mathbf{h}), \\ r_N^-(i, \mathbf{h})p_0(\mathbf{h}) &= r_N^+(i - 1, J_i^{i-1}\mathbf{h})p_0(J_i^{i-1}\mathbf{h}). \end{aligned} \tag{4}$$

There are many rates that satisfy detailed balance; in [4], for example, Marzuola considers the so-called adatom rates $r^+(i) = r^-(i) = e^{-2Kn(i)}$, where $n(i)$ is a so-called coordination number that quantifies the energy it takes for the topmost atom at site i to break the bonds with its nearest neighbors. While the adatom rates are physically motivated, the rates we consider here are motivated by computational convenience. They are defined by

$$r_N^{\pm}(i, \mathbf{h}) = \exp\left[-\frac{K}{2}(H(J_i^{i\pm 1}\mathbf{h}) - H(\mathbf{h}))\right].$$

The generator $\mathcal{A}_N$ of the Markov process $\mathbf{h}_N^t$ quantifies the instantaneous change in the average value of an observable f of the process. Namely,

$$(\mathcal{A}_N f)(\mathbf{h}) := \lim_{t\to 0} \frac{\mathbb{E}[f(\mathbf{h}_N^t)|\mathbf{h}_N^0 = \mathbf{h}] - f(\mathbf{h})}{t} = \sum_{i=1}^{N} (f(J_i^{i+1}\mathbf{h}) - f(\mathbf{h}))r^+(i, \mathbf{h}) + (f(J_i^{i-1}\mathbf{h}) - f(\mathbf{h}))r^-(i, \mathbf{h}). \tag{5}$$

One can also show that

$$f(\mathbf{h}_N^t) = f(\mathbf{h}_N^0) + \int_0^t (\mathcal{A}_N f)(\mathbf{h}_N^s)ds + M_f^t,$$

where M_f^t is a zero-mean martingale.

4 Scaling limit

Establishing a hydrodynamic limit is typically achieved in two steps. First, assuming a limiting macroscopic profile $h(t, \cdot) : \mathbb{T} \to \mathbb{R}$ exists, one derives the PDE it satisfies. One then proves the solution to this PDE is unique, and shows that the microscopic profile at each time t (appropriately rescaled to some $\bar{h}_N(t, \cdot)$) does indeed converge to $h(t,\cdot)$ as $N \to \infty$. The appropriate convergence regime is in $H^{-1}(\mathbb{T})$, i.e. one should show that $\mathbb{E}||\bar{h}_N^t - h(t,\cdot)||_{H^{-1}(\mathbb{T})} \to 0$ as $N \to \infty$ for each t, where the expectation is taken with respect to the measure of the process at time t given a particular initial profile h_0. Here, we will focus only on the first step, that is, deriving the PDE under the assumption that a sufficiently smooth limit h exists.

Recall that $\mathbf{h}_N^t = (h_1(t), ..., h_N(t))$ and consider the rescaled process $\bar{h}_N(t,\cdot) : \mathbb{T} \to \mathbb{R}$, defined by

$$\bar{h}_N(t,x) = \frac{1}{N} h_{\lfloor Nx \rceil}(N^4 t),$$

where $\lfloor y \rceil$ is defined to mean the nearest integer to y. We have scaled down the distance between lattice sites to live in the unit interval, and scaled the heights by $\frac{1}{N}$ so that the total mass, $\frac{m(N)}{N}$, does not blow up. The time scale was chosen to yield a meaningful limit. In particular, it reflects that microscopic events occur at a very fast rate on the macroscopic time scale. The general structure of the derivation presented in the following sections follows that of Marzuola in [4].

4.1 Window average framework

As described in Section 2, the standard technique that hydrodynamic limit proofs rely on is to show that after a short amount of time on the macroscopic time scale, the microscopic system should reach a local equilibrium state (we will provide numerical confirmation of this fact but do not provide proof). In order to take advantage of this property, we can approximate $\bar{h}_N(t,x)$ by sliding window averages of the function over a small but macroscopic window centered at $\frac{\lfloor Nx \rceil}{N}$. Specifically, define the sets

$$S_{k,\delta} = (\frac{k}{N} - \frac{\delta}{2}, \frac{k}{N} + \frac{\delta}{2})$$

(where addition is modulo 1) and window averages

$$\phi_{k,\delta}(t) = \frac{1}{\delta} \int_{S_k} \bar{h}_N(t,x)dx = \frac{1}{N\delta} \frac{1}{N} \sum_{j:\frac{j}{N} \in S_k} h_j(N^4 t).$$

Letting $k(x) = \lfloor Nx \rceil$, we have that if $\frac{1}{N} << \delta << 1$, then

$$h(t,x) \approx \bar{h}_N(t,x) \approx \phi_{k(x),\delta}(t).$$

We therefore have

$$h(t,x) - h(0,x) \approx \frac{1}{N\delta} \frac{1}{N} \sum_{j:\frac{j}{N} \in S_{k(x)}} h_j(N^4 t) - h_j(0) = \frac{1}{N\delta} \frac{1}{N} \sum_{j:\frac{j}{N} \in S_{k(x)}} \int_0^{N^4 t} (\mathcal{A}_N \pi_j)(h_N(s))ds + M_j(N^4 t), \tag{6}$$

where π_j is the projection operator onto h_j. Using (5) with $f = \pi_j$, we see that

$$(\mathcal{A}_N \pi_j)(\mathbf{h}) = \left(r_N^+(j-1,\mathbf{h}) - r_N^+(j,\mathbf{h})\right) + \left(r_N^-(j+1,\mathbf{h}) - r_N^-(j,\mathbf{h})\right).$$

Substituting this into (6) and changing variables, we obtain

$$h(t,x) - h(0,x) \approx \frac{N^3}{N\delta} \sum_{j:\frac{j}{N} \in S_{k(x)}} \int_0^t r^+(j-1,N^4 s) - r^+(j,N^4 s) + r^-(j+1,N^4 s) - r^-(j,N^4 s)ds + M_j(N^4 t), \tag{7}$$

where we have used the notation $r^{\pm}(k,t)$ in place of $r^{\pm}(k,\mathbf{h}_N^t)$.

The random variables $r^{\pm}(N^4 s, j)$ are local, depending only on z_i for $i = j-1, j, j+1$. The local equilibrium measure, which we describe in detail in the following section, varies smoothly along the domain. Thus, the random variables $r^{\pm}(N^4 s, j)$ for $\frac{j}{N} \in S_{k(x)}$ have nearly the same distribution as that of

$r^{\pm}(N^4s, k(x))$, the central point of the window $S_{k(x)}$. Moreover, they should have low correlation with one another (also discussed in the following section), which implies that the $M_j(N^4t)$ have low correlation with one another as well. We thus expect a law of large numbers to hold, allowing us to replace the average of M_j over $j \in S_{k(x)}$ by 0, its expectation, and to make the substitution

$$\frac{1}{N\delta} \sum_{j:\frac{j}{N} \in S_{k(x)}} r^{\pm}(j, N^4 s) \approx \mathbb{E}\left[r^{\pm}(k(x), N^4 s)\right].$$

Also,

$$\frac{1}{N\delta} \sum_{j:\frac{j}{N} \in S_{k(x)}} r^{+}(j-1, N^4 s) \approx \mathbb{E}\left[r^{+}(k(x)-1, N^4 s)\right],$$

and similarly for $r^-(j+1)$. We thus have

$$h(t,x) - h(0,x) \approx N^3 \int_0^t \mathbb{E}\left[r^+(k(x)-1, N^4 s) - r^+(k(x), N^4 s) + r^-(k(x)+1, N^4 s) - r^-(k(x), N^4 s)\right] ds. \tag{8}$$

4.2 Local equilibrium measure

We now turn to characterizing the local equilibrium measure according to which the height profiles are distributed at time N^4s, for s away from 0. Such a measure is defined to have maximum entropy among all measures that have certain prescribed local averages, which is equivalent to the measure whose Kullback-Leibler divergence with respect to the global equilibrium measure p_0 is minimal among this subset of measures. This idea is described in more generality and made rigorous by Roux and Weare in [8]. In our case, these local averages should be the values of $\nabla h(s, \frac{i}{N})$, the surface gradient of the macroscopic profile the microscopic process is converging to. This means that in a neighborhood of $\frac{i}{N}$, we expect the values of the microscopic height gradients to cluster around $\nabla h(s, \frac{i}{N})$. To be more precise, the local equilibrium measure p_λ - known as the optimal twist measure - is the solution to the constrained minimization problem

$$\begin{aligned} &\underset{p}{\text{minimize}} && KL(p||p_0) \\ &\text{subject to} && \sum_{\mathbf{h} \in \mathbb{Z}_N^m} p(\mathbf{h}) = 1, \sum_{\mathbf{h} \in \mathbb{Z}_N^m} (h_{i+1} - h_i) p(\mathbf{h}) = \nabla h(t, \frac{i}{N}), i = 1, ..., N. \end{aligned}$$

We solve this problem using Lagrange multipliers. Defining $\mathbf{p} = (p(\mathbf{h}))_{\mathbf{h} \in \mathbb{Z}_N^m}$ (a sequence, since the state space is countable), we have

$$\begin{aligned} \mathbf{0} &= \nabla_{\mathbf{p}} \sum_{\mathbf{h} \in \mathbb{Z}_N^m} -p(\mathbf{h}) \log \frac{p(\mathbf{h})}{p_0(\mathbf{h})} + \alpha \left(\sum_{\mathbf{h} \in \mathbb{Z}_N^m} p(\mathbf{h}) - 1 \right) + \sum_{i=1}^N \lambda_i \left(\sum_{\mathbf{h} \in \mathbb{Z}_N^m} (h_{i+1} - h_i) p(\mathbf{h}) - \nabla h(t, \frac{i}{N}) \right) \\ &= \left(-\log p(\mathbf{h}) + \log p_0(\mathbf{h}) + \alpha + \sum_{i=1}^N \lambda_i (h_{i+1} - h_i) - 1 \right)_{\mathbf{h} \in \mathbb{Z}_N^m}. \end{aligned} \tag{9}$$

We thus see that

$$p_\lambda(\mathbf{h}) \propto p_0(\mathbf{h}) \exp[\sum_{i=1}^N \lambda_i (h_{i+1} - h_i)] \propto \exp[\sum_{i=1}^N -K(h_{i+1} - h_i)^2 + \lambda_i (h_{i+1} - h_i)].$$

To find the λ_i's we must solve the equation

$$\nabla h(t, \frac{i}{N}) = \mathbb{E}_\lambda[h_{i+1} - h_i] = \frac{\sum_{\mathbf{h} \in \mathbb{Z}_N^m} (h_{i+1} - h_i) \exp[\sum_{i=1}^N -K(h_{i+1} - h_i)^2 + \lambda_i (h_{i+1} - h_i)]}{\sum_{\mathbf{h} \in \mathbb{Z}_N^m} \exp[\sum_{i=1}^N -K(h_{i+1} - h_i)^2 + \lambda_i (h_{i+1} - h_i)]}. \tag{10}$$

It is helpful to change variables in the sum over profiles $\mathbf{h} \in \mathbb{Z}_N^m$ appearing in (10) and instead consider gradient-tuples. The space of gradient profiles corresponding to the height profiles in $\mathbb{Z}_N^m$ is

$$\nabla \mathbb{Z}_N^m := \{\mathbf{z} = (z_1, z_2, ..., z_{N-1}) \in \mathbb{Z}^{N-1} : \sum_{i=1}^{N-1} (N-i) z_i \equiv m \mod N\},$$

where the z_i's are defined as before, i.e. $z_i = h_{i+1} - h_i, i = 1, ..., N-1$ (note that $z_N = h_1 - h_N = -\sum_{i=1}^{N-1} z_i$). To see why the modulus restriction is necessary, note that given an arbitrary $N-1$-tuple $\mathbf{z}$, we have $h_i = h_1 + z_1 + ... + z_{i-1}, i > 1$, so that

$$m = \sum_{i=1}^{N} h_i = N h_1 + \sum_{i=1}^{N-1} (N-i) z_i$$

has an integer solution h_1 if and only if $\sum_{i=1}^{N-1} (N-i) z_i \equiv m \mod N$.

It can be shown by arguments similar to those appearing in Section 6.1 (i.e. by considering the sums as expectations of a function of independent random variables $Z_1, ..., Z_{N-2}$), that for large N, λ_i is well-approximated by the solution of

$$\nabla h(t, \frac{i}{N}) = \mathbb{E}_{\lambda_i}[Z_i] = \frac{\sum_{n \in \mathbb{Z}} n e^{-Kn^2 + \lambda_i n}}{\sum_{n \in \mathbb{Z}} e^{-Kn^2 + \lambda_i n}},$$

which is what (10) would reduce to if the z_i were independent.

Consider the general inverse equation for λ as a function of u:

$$u = \frac{\sum_{z \in \mathbb{Z}} z e^{-Kz^2 + \lambda(u) z}}{\sum_{z \in \mathbb{Z}} e^{-Kz^2 + \lambda(u) z}} = \frac{d}{d\lambda} \log \mathcal{Z}_\lambda,$$

where $\mathcal{Z}_\lambda := \sum_{z \in \mathbb{Z}} e^{-Kz^2 + \lambda z}$. One can easily show that $\lambda(u) = \nabla \sigma_D(u)$, where

$$\sigma_D(u) = \sup_{\eta \in \mathbb{R}} \{\eta u - \log \mathcal{Z}_\eta\},$$

the Legendre tranform of $\log \mathcal{Z}_\eta$. The function σ_D is the surface tension obtained by Marzuola in [4]; we see now the connection between it and the local equilibrium measure. Another way to think about the surface tension is as follows: in the presence of a restorative force field $\nabla \sigma_D(u)$, the Hamiltonian of a "system" of one height gradient becomes $H(z) = z^2 - \frac{\nabla \sigma_D(u)}{K} z$, and the average value of the height gradient with probability distribution $p(z) \propto e^{-KH(z)}$ is u.

We now briefly recall the model considered by Nishikawa in [5] in which the heights are allowed to be continuous. In this case, the change of variables

$$\{(h_1, ..., h_N) \in \mathbb{R}^N : \sum_{i=1}^{N} h_i = m\} \mapsto \{(z_1, ..., z_{N-1}) \in \mathbb{R}^{N-1}\}$$

is bijective. Since there is one less restriction on the relationship between the z_j, we expect the solution to the equation reduced to the single z_i variable to be an even better approximation for λ_i, i.e.

$$\nabla h(t, \frac{i}{N}) = \mathbb{E}_\lambda[z_i] = \frac{\int_{\mathbb{R}} x e^{-Kx^2 + \lambda_i x}}{\int_{\mathbb{R}} e^{-Kx^2 + \lambda_i x}}, \tag{11}$$

which has solution $\lambda_i = 2K \nabla h(t, \frac{i}{N})$. The function $\sigma_C(u)$ obtained in [5] is constructed analogously to $\sigma_D(u)$, and is simply $\sigma_C(u) = Ku^2$. We explore the relationship between σ_D, σ_C, and K, at the end of this section.

To summarize, the unnormalized optimal twist measure at time $N^4 s$ is given by

$$p_\lambda(\mathbf{h}) = \exp\left[-H(\mathbf{h}) + \sum_{i=1}^{N} \lambda_i z_i\right], \lambda_i = \nabla \sigma_D(\nabla h(s, \frac{i}{N})).$$

From now on, expectation with respect to the optimal twist measure will be denoted by $\langle \cdot \rangle_\lambda$.

4.3 Rate expectation

We now return to computing the expectation of the rates in (8) with respect to the optimal twist measure. We have

$$\langle r^+(j)\rangle_\lambda = \frac{\sum_{\mathbf{h}\in\mathbb{Z}_N^m} r^+(j,\mathbf{h})p_\lambda(\mathbf{h})}{\sum_{\mathbf{h}\in\mathbb{Z}_N^m} p_\lambda(\mathbf{h})}.$$

The summand in the numerator takes the form

$$\begin{aligned}
&r^+(j,\mathbf{h})p_\lambda(\mathbf{h})\\
&= \exp\left[-(H(J_j^{j+1}\mathbf{h}) - H(h))/2 - H(\mathbf{h}) + \sum_{i=1}^N \lambda_i z_i\right]\\
&= \exp\left[-(H(J_j^{j+1}\mathbf{h}) + H(h))/2 + \sum_{i=1}^N \lambda_i z_i\right]\\
&= \exp\left[-\frac{K}{2}((z_{j-1}-1)^2 + z_{j-1}^2 + (z_j+2)^2 + z_j^2 + (z_{j+1}-1)^2 + z_{j+1}^2)\right]\\
&\times \exp\left[-K\sum_{i\neq j-1,j,j+1} z_i^2 + \sum_{i=1}^N \lambda_i z_i\right]\\
&= \exp\big[-K(z_{j-1}^2 - z_{j-1} + \frac{1}{2} + z_j^2 + 2z_j + 2 + z_{j+1}^2 - z_{j+1} + \frac{1}{2}) - K\sum_{i\neq j-1,j,j+1} z_i^2 + \sum_{i=1}^N \lambda_i z_i\big]\\
&= \exp\big[-3K - K\sum_{i=1}^N z_i^2 + (\lambda_{j-1}+K)z_{j-1} + (\lambda_j - 2K)z_j + (\lambda_{j+1}+K)z_{j+1} + \sum_{i\neq j-1,j,j+1} \lambda_i z_i\big]
\end{aligned} \tag{12}$$

Thus,

$$\langle r^+(j)\rangle_\lambda = e^{-3K}\frac{\sum_{\mathbf{h}\in\mathbb{Z}_N^m} \exp[\sum_{i=1}^N -Kz_i^2 + \tilde{\lambda}_i z_i]}{\sum_{\mathbf{h}\in\mathbb{Z}_N^m} \exp[\sum_{i=1}^N -Kz_i^2 + \lambda_i z_i]},$$

where $\tilde{\lambda}_{j-1} = \lambda_{j-1} + K$, $\tilde{\lambda}_j = \lambda_j - 2K$, $\tilde{\lambda}_{j+1} = \lambda_{j+1} + K$, and $\tilde{\lambda}_i = \lambda_i$ for all other i. Multiplying the numerator and denominator by $\exp\left[-\frac{1}{4K}\sum_{i=1}^N \lambda_i^2\right]$, we obtain

$$\langle r^+(j)\rangle_\lambda = e^{-3K}\frac{\sum_{\mathbf{h}\in\mathbb{Z}_N^m} \exp[s_j - K\sum_{i\neq j-1,j,j+1}(z_i - \lambda_i/2K)^2]}{\sum_{\mathbf{h}\in\mathbb{Z}_N^m} \exp[-K\sum_{i=1}^N (z_i - \lambda_i/2K)^2]}, \tag{13}$$

where

$$\begin{aligned}
s_j = &-Kz_{j-1}^2 + (\lambda_{j-1}+K)z_{j-1} - \frac{\lambda_{j-1}^2}{4K}\\
&- Kz_j^2 + (\lambda_{j-1} - 2K)z_j - \frac{\lambda_j^2}{4K}\\
&- Kz_{j+1}^2 + (\lambda_{j+1}+K)z_{j+1} - \frac{\lambda_{j+1}^2}{4K}.
\end{aligned} \tag{14}$$

Completing each of the three squares in s_j yields

$$
\begin{aligned}
s_j &= -K\left(z_{j-1} - \frac{\lambda_{j-1}+K}{2K}\right)^2 - K\left(z_j - \frac{\lambda_j - 2K}{2K}\right)^2 - K\left(z_{j+1} - \frac{\lambda_{j+1}+K}{2K}\right)^2 \\
&+ \frac{(\lambda_{j-1}+K)^2}{4K} - \frac{\lambda_{j-1}^2}{4K} + \frac{(\lambda_j - 2K)^2}{4K} - \frac{\lambda_j^2}{4K} + \frac{(\lambda_{j+1}+K)^2}{4K} - \frac{\lambda_{j+1}^2}{4K} \\
&= -K\left(z_{j-1} - \frac{\lambda_{j-1}+K}{2K}\right)^2 - K\left(z_j - \frac{\lambda_j - 2K}{2K}\right)^2 - K\left(z_{j+1} - \frac{\lambda_{j+1}+K}{2K}\right)^2 \\
&+ \frac{3}{2}K + \frac{1}{2}(\lambda_{j-1} - 2\lambda_j + \lambda_{j+1})
\end{aligned}
\tag{15}
$$

Substituting this expression into (13), we obtain

$$
\langle r^+(j)\rangle_\lambda = \exp[-\frac{3}{2}K + \frac{1}{2}(\lambda_{j-1} - 2\lambda_j + \lambda_{j+1})]\frac{\sum_{\mathbf{h}\in\mathbb{Z}_N^m}\exp[-K\sum_{i=1}^N (z_i - (\frac{\lambda_i}{2K} + c_i))^2]}{\sum_{\mathbf{h}\in\mathbb{Z}_N^m}\exp[-K\sum_{i=1}^N (z_i - \frac{\lambda_i}{2K})^2]}, \tag{16}
$$

where $c_{j-1} = c_{j+1} = \frac{1}{2}$, $c_j = -1$, and $c_i = 0$ for all other i.

We now simplify the sums in this expression. Since they have the same form, we concentrate on the sum in the denominator, and the numerator will be similar. As in the previous section, change variables from $\mathbf{h} \in \mathbb{Z}_N^m$ to $\mathbf{z} \in \nabla\mathbb{Z}_N^m$. Define $p_l(z) := e^{-K(z-l)^2}$, $l_i := \frac{\lambda_i}{2K}$. Under the change of variables, the sum in the denominator of (16) takes the form

$$
\sum_{\mathbf{h}\in\mathbb{Z}_N^m} p_{l_N}(-\sum_{i=1}^{N-1} z_i)\prod_{i\neq N} p_{l_i}(z_i) \tag{17}
$$

We first work on computing the inner sum over z_{N-1}. Define the variables $S_N = \sum_{i=1}^{N-2} z_i$ and $T_N = \sum_{i=1}^{N-2} i z_i$. The variable z_{N-1} can take values in

$$
\{Nj + m - \sum_{i=1}^{N-2}(N-i)z_i : j \in \mathbb{Z}\} = \{Nj + m + T_N : j \in \mathbb{Z}\}.
$$

Substituting this second expression for z_{N-1} into the part of the product in the summand that depends on it, we obtain

$$
\begin{aligned}
p_{l_{N-1}}(z_{N-1})p_{l_N}\left(-\sum_{i=1}^{N-1} z_i\right) &= p_{l_{N-1}}(z_{N-1})p_{l_N}(-S_N - z_{N-1}) \\
&= \exp\left[-K(Nj + m + T_N - l_{N-1})^2 - K(Nj + m + T_N + S_N + l_N)^2)\right]
\end{aligned}
\tag{18}
$$

Now we make use of the identity

$$
c_1(x + a_1)^2 + c_2(x + a_2)^2 = (c_1 + c_2)(x + \frac{a_1c_1 + a_2c_2}{c_1 + c_2})^2 + \frac{c_1c_2}{c_1 + c_2}(a_1 - a_2)^2
$$

with $x = Nj + m + T_N$ to reexpress (18) as

$$
f_j(S_N, T_N) := \exp\left[-\frac{K}{2}(S_N + l_{N-1} + l_N)^2 - 2K(Nj + m + \frac{l_N - l_{N-1}}{2} + T_N + \frac{S_N}{2})^2\right]. \tag{19}
$$

Defining $f(S_N, T_N) := \sum_{j\in\mathbb{Z}} f_j(S_N, T_N)$, and noting that $z_1, ..., z_{N-2}$ can range freely over $\mathbb{Z}$, we thus have

$$
\sum_{\mathbf{h}\in\mathbb{Z}_N^m} p_{l_N}(-\sum_{i=1}^{N-1} z_i)\prod_{i\neq N} p_{l_i}(z_i) = \sum_{z_1=-\infty}^{\infty} \cdots \sum_{z_{N-2}=-\infty}^{\infty}\prod_{i=1}^{N-2} p_{l_i}(z_i) f(S_N, T_N).
$$

Defining the integer-valued random variables Z_i, with $P(Z_i = n) \propto p_{l_i}(n), n \in \mathbb{Z}$, we thus see that

$$\sum_{\mathbf{h}\in\mathbb{Z}_N^m} p_{l_N}(-\sum_{i=1}^{N-1} z_i) \prod_{i\neq N} p_{l_i}(z_i) = \mathcal{Z}\mathbb{E}\left[f(S_N(Z_1, ..., Z_{N-2}), T_N(Z_1, ..., Z_{N-2}))\right]$$

where the Z_i's are independent and

$$\mathcal{Z} = \prod_{i=1}^{N-2} \sum_{n\in\mathbb{Z}} p_{l_i}(n)$$

is the normalization factor of the distribution of $(Z_1, ..., Z_{N-2})$.

Similarly, the numerator in (16) takes the form

$$\mathcal{Z}^+\mathbb{E}\left[f(S_N(Z_1^+, ..., Z_{N-2}^+), T_N(Z_1^+, ..., Z_{N-2}^+))\right],$$

where $P(Z_i^+ = n) \propto p_{l_i+c_i}(n), n \in \mathbb{Z}$, and $\mathcal{Z}^+$ is the normalization factor of the distribution of $(Z_1^+, ..., Z_{N-2}^+)$ (the + indicates we are computing r^+). Note that, because of the product structure in the normalization factors, and the fact that $c_i = 0$ for all but three values of i, $\frac{\mathcal{Z}^+}{\mathcal{Z}}$ will reduce to

$$\begin{aligned}\frac{\mathcal{Z}^+}{\mathcal{Z}} &= \frac{\sum_{n\in\mathbb{Z}} p_{l_{j-1}+\frac{1}{2}}(n)}{\sum_{n\in\mathbb{Z}} p_{l_{j-1}}(n)} \frac{\sum_{n\in\mathbb{Z}} p_{l_j-1}(n)}{\sum_{n\in\mathbb{Z}} p_{l_j}(n)} \frac{\sum_{n\in\mathbb{Z}} p_{l_{j+1}+\frac{1}{2}}(n)}{\sum_{n\in\mathbb{Z}} p_{l_{j+1}}(n)} \\ &= \frac{\sum_{n\in\mathbb{Z}} e^{-K(n-l_{j-1}-\frac{1}{2})^2}}{\sum_{n\in\mathbb{Z}} e^{-K(n-l_{j-1})^2}} \frac{\sum_{n\in\mathbb{Z}} e^{-K(n-l_{j+1}-\frac{1}{2})^2}}{\sum_{n\in\mathbb{Z}} e^{-K(n-l_{j+1})^2}}.\end{aligned} \tag{20}$$

The middle term cancelled because a shift by one does not change the sum. Letting $\mathbb{E}^+$ be the expectation in the numerator, we have thus shown that

$$\langle r^+(j)\rangle_\lambda = e^{-\frac{3}{2}K}\exp\left[\frac{1}{2}(\lambda_{j-1} - 2\lambda_j + \lambda_{j+1})\right] \frac{\mathcal{Z}^+}{\mathcal{Z}} \frac{\mathbb{E}^+}{\mathbb{E}}. \tag{21}$$

We now compute $\langle r^-(j+1)\rangle_\lambda$, as it is very similar in form to $\langle r^+(j)\rangle_\lambda$. To do so, we simply change $z_{j-1} - 1, z_j + 2, z_{j+1} - 1$ in the third line of (12) to $z_{j-1} + 1, z_j - 2, z_{j+1} + 1$, respectively, and continue changing $-$ to $+$ and vice versa throughout the derivation. We end up with

$$\langle r^-(j+1)\rangle_\lambda = e^{-\frac{3}{2}K}\exp\left[-\frac{1}{2}(\lambda_{j-1} - 2\lambda_j + \lambda_{j+1})\right] \frac{\mathcal{Z}^-}{\mathcal{Z}} \frac{\mathbb{E}^-}{\mathbb{E}}. \tag{22}$$

Note that $\mathcal{Z}^+ = \mathcal{Z}^-$, because the change from $l_{j-1} + \frac{1}{2}$ to $l_{j-1} - \frac{1}{2}$ is equivalent to a shift of n by 1 in the sum (the same goes for l_{j+1}).

We now show that $\mathbb{E}^+ = \mathbb{E}^-$. These expectations differ only in the probability mass functions of the variables $Z_k^{+,-}, k = j-1, j, j+1$. We have that $Z_{j-1}^- = Z_{j-1}^+ - 1$, $Z_j^- = Z_j^+ + 2$, and $Z_{j+1}^- = Z_{j+1}^+ - 1$ in distribution. To see this, note for example that

$$p_{j-1}^+(n) = \frac{e^{-K(n-(l_{j-1}+\frac{1}{2}))^2}}{\sum_n e^{-K(n-(l_{j-1}+\frac{1}{2}))^2}} = \frac{e^{-K(n-1-(l_{j-1}-\frac{1}{2}))^2}}{\sum_n e^{-K(n-1-(l_{j-1}-\frac{1}{2}))^2}} = p_{j-1}^-(n-1).$$

Thus,

$$S_N^- = \sum_{i\neq j-1,j,j+1} Z_i + (Z_{j-1}^+ - 1) + (Z_j^+ + 2) + (Z_{j+1}^+ - 1) = S_N^+$$

and

$$T_N^- = \sum_{i\neq j-1,j,j+1} iZ_i + (j-1)(Z_{j-1}^+ - 1) + j(Z_j^+ + 2) + (j+1)(Z_{j+1}^+ - 1) = T_N^+.$$

Hence, since the expectations are a function of S_N and T_N only, they must be equal.

4.4 The PDE

We substitute the formulas obtained for $\langle r^{\pm}\rangle_\lambda$ into (8), letting Z_j denote $\frac{\mathcal{Z}^{\pm}}{\mathcal{Z}}$, E_j denote $\frac{\mathbb{E}^{\pm}}{\mathbb{E}}$ and D_j denote $\lambda_{j-1} - 2\lambda_j + \lambda_{j+1}$. Recall that λ_j actually depends on s, i.e. $\lambda_j(s) = \nabla\sigma_D(\nabla h(s, \frac{j}{N}))$.

$$
\begin{aligned}
h(t,x) - h(0,x) &\approx N^3 \int_0^t e^{-\frac{3}{2}K} Z_j E_j \left[e^{-\frac{1}{2}D_j} - e^{\frac{1}{2}D_j}\right] - e^{-\frac{3}{2}K} Z_{j-1}E_{j-1}\left[e^{-\frac{1}{2}D_{j-1}} - e^{\frac{1}{2}D_{j-1}}\right] ds \\
&\approx N^3 \int_0^t -e^{-\frac{3}{2}K} Z_j E_j D_j + e^{-\frac{3}{2}K} Z_{j-1}E_{j-1}D_{j-1} ds \qquad (23)\\
&= N^3 \int_0^t -e^{-\frac{3}{2}K}\left[(E_j - E_{j-1})Z_jD_j + (Z_jD_j - Z_{j-1}D_{j-1})E_{j-1}\right] ds
\end{aligned}
$$

Now, for N large,

$$D_i \approx N^{-2}\partial_x^2 \nabla\sigma_D(\nabla h(t, \frac{i}{N})),$$

and $Z_i \approx Z^2(\lambda_i) = Z^2(\nabla\sigma_D(\nabla h(t, \frac{i}{N}))$, where

$$Z(\lambda) := \frac{\sum_{n\in\mathbb{Z}} e^{-K(n-\lambda-\frac{1}{2})^2}}{\sum_{n\in\mathbb{Z}} e^{-K(n-\lambda)^2}}.$$

We will show in the appendix that $E_i \to 1$ for all i, and $N(E_j - E_{j-1}) \to 0$. Thus, N^3 times the first summand in the integral above will vanish, while N^3 times the second summand will approach

$$\partial_x\left[Z^2(\nabla\sigma_D(\nabla h(t,x))\partial_x^2\nabla\sigma_D(\nabla h(t,x))\right].$$

We thus arrive at

$$\partial_t h(t,x) = -e^{-\frac{3}{2}K}\partial_x\left[Z^2(\nabla\sigma_D(\nabla h(t,x))\partial_x^2\nabla\sigma_D(\nabla h(t,x))\right]. \qquad (24)$$

Note that this equation does not immediately fit into the framework of gradient descent in surface free energy described in Section 2. This suggests that the "true" surface tension σ that is appropriate for these rates is not in fact σ_D, but a small correction on σ_D that depends on $Z(\lambda)$. This is not implausible; indeed, recall that the λ_i computed using the Legendre transformation were only approximations of the parameters λ that resulted in the distribution closest to p_0.

4.5 Dependence on K

The inverse temperature K affects the PDE through Z and $\nabla\sigma_D$. The sums appearing in Z can be expressed through the θ_3 function as follows:

$$\sum_{z=-\infty}^{\infty} e^{-K(z-\lambda)^2} = \frac{\sqrt{\pi}}{\sqrt{K}}\theta_3(\pi\lambda, e^{-\frac{\pi^2}{K}}) = \frac{\sqrt{\pi}}{\sqrt{K}}\sum_{n=-\infty}^{\infty} e^{-\frac{\pi^2 n^2}{K}}\cos 2\pi n\lambda. \qquad (25)$$

Thus, for small K we have

$$\sum_{z=-\infty}^{\infty} e^{-K(z-\lambda)^2} \approx \frac{\sqrt{\pi}}{\sqrt{K}}(1 + 2e^{-\frac{\pi^2}{K}}\cos 2\pi\lambda).$$

The second term is an exponentially small correction, so

$$Z(\lambda) = \frac{\sum_{n\in\mathbb{Z}} e^{-K(n-\lambda-\frac{1}{2})^2}}{\sum_{n\in\mathbb{Z}} e^{-K(n-\lambda)^2}} \approx 1$$

for all λ, and one can also see that its derivative is approaching 0.

Moreover, we can show that $\nabla\sigma_D \to \nabla\sigma_C$ as $K \to 0$: We have

$$\begin{aligned}\sum_{z=-\infty}^{\infty}(z-\lambda)e^{-K(z-\lambda)^2} &= \frac{1}{2K}\frac{d}{d\lambda}\Big[\sum_{z=-\infty}^{\infty}e^{-K(z-\lambda)^2}\Big] \\ &= \frac{1}{2K}\frac{d}{d\lambda}\Big[\frac{\sqrt{\pi}}{\sqrt{K}}\sum_{n=-\infty}^{\infty}e^{-\frac{\pi^2n^2}{K}}\cos 2\pi n\lambda\Big] = -\frac{\pi}{K}\frac{\sqrt{\pi}}{\sqrt{K}}\sum_{n=-\infty}^{\infty}ne^{-\frac{\pi^2n^2}{K}}\sin 2\pi n\lambda.\end{aligned} \tag{26}$$

Thus, using $\lambda(u)$ to denote $\frac{\nabla\sigma_D(u)}{2K}$, we have

$$\begin{aligned}u-\lambda(u) &= \frac{\sum_{z=-\infty}^{\infty}(z-\lambda)e^{-K(z-\lambda)^2}}{\sum_{z=-\infty}^{\infty}e^{-K(z-\lambda)^2}} = -\frac{\pi}{K}\frac{\sum_{n=-\infty}^{\infty}ne^{-\frac{\pi^2n^2}{K}}\sin 2\pi n\lambda}{\sum_{n=-\infty}^{\infty}e^{-\frac{\pi^2n^2}{K}}\cos 2\pi n\lambda} \\ &= -\frac{\pi}{K}\frac{2(e^{-\frac{\pi^2}{K}}\sin 2\pi\lambda+...)}{1+2(e^{-\frac{\pi^2}{K}}\cos 2\pi\lambda+...)} \approx \frac{-2\pi}{K}e^{-\frac{\pi^2}{K}}\sin 2\pi\lambda.\end{aligned} \tag{27}$$

Thus we have the implicit equation for $\nabla\sigma_D(u)$,

$$\nabla\sigma_D(u) \approx 2Ku + 4\pi e^{-\frac{\pi^2}{K}}\sin\Big(\frac{\pi}{K}\nabla\sigma_D(u)\Big).$$

Since $|\nabla\sigma_D(u)-2Ku| \le 4\pi e^{-\frac{\pi^2}{K}} << 1$, we may substitute $2Ku$ for $\nabla\sigma_D(u)$ on the righthand side to obtain

$$\nabla\sigma_D(u) \approx 2Ku + 4\pi e^{-\frac{\pi^2}{K}}\sin(2\pi u).$$

For small K therefore, the PDE starts to look like (1) with σ_C, i.e.

$$\partial_t h(t,x) = -2Ke^{-\frac{3}{2}K}\partial_x^4 h(t,x).$$

5 Numerical Simulations and Discussion

To check whether our derivation is correct, we used the Kinetic Monte Carlo (KMC) algorithm to simulate the evolution of the rescaled microscopic process up to time N^4T for a given T. The algorithm takes in the initial profile h_0 at $t_0 = 0$ and proceeds as follows, until t_k exceeds N^4T:

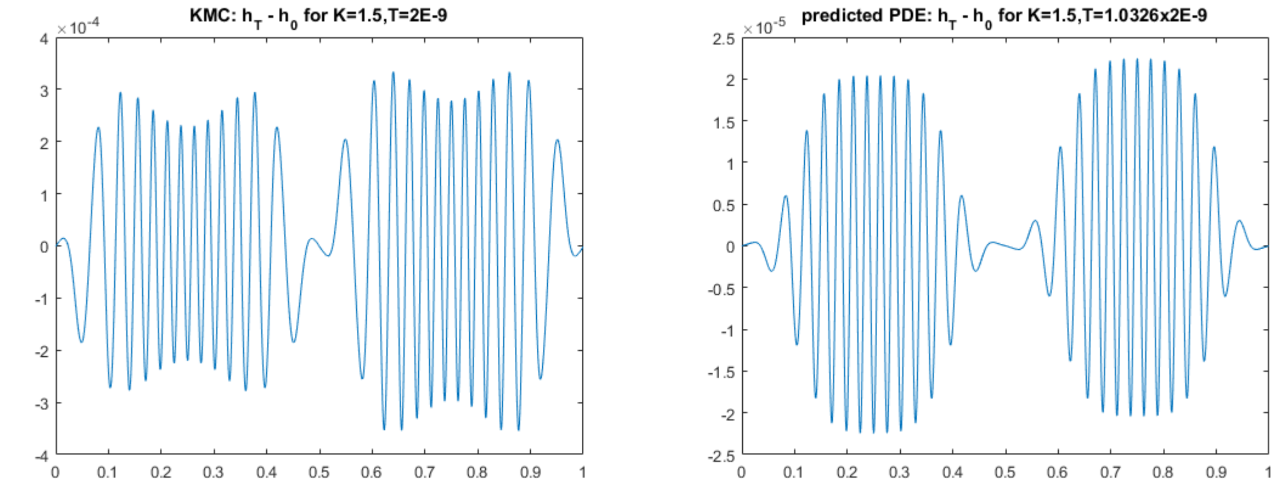
1. Given t_k and $h(t_k,\cdot)$, compute $R_k := \{r^{\pm}(i,h(t_k,\cdot)\}$.
2. Draw dt from the $T_{h_k} \sim \text{Exp}(\sum_{r\in R_k} r)$ distribution and set $t_{k+1} = t_k + dt$.
3. Choose an r_k from the distribution $P(r = r_i) = \frac{r_i}{R_k}$, and define $h(t_{k+1},\cdot)$ to be the result of the transition from $h(t_k,\cdot)$ associated to r_k.

This is an accurate simulation of the Markov process because for small dt we have

$$\begin{aligned}\frac{P(h_{t+dt} = J_i^{i\pm1}h|h_t = h)}{dt} &= \frac{r^{\pm}(i,h)}{\sum_i(r^+(i,h)+r^-(i,h))}\frac{P(T_h\le dt)}{dt} \\ &= \frac{r^{\pm}(i,h)}{\sum_i(r^+(i,h)+r^-(i,h))}\frac{1-\exp\left[-\sum_i(r^+(i,h)+r^-(i,h))dt\right]}{dt} = r^{\pm}(i,h)+o(1)\end{aligned} \tag{28}$$

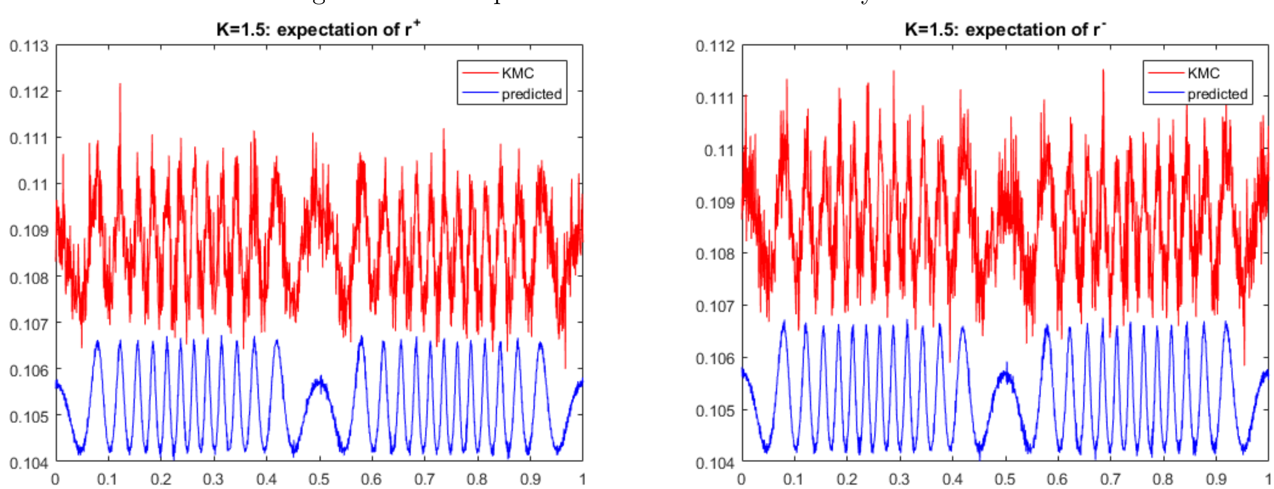
We used $h_0(x) = \sin(2\pi x)$ in all of our simulations. The signature of the microscopic dynamics is best discerned by looking at the change in the profile after a short amount of time. Figure 1 shows a plot of the average of $\mathbf{h}^{N^4t} - \mathbf{h}^0$ over many sample runs of KMC, next to a plot of $h(T,\cdot) - h(0,\cdot)$, where $h(t,x)$ is the solution to the PDE we have derived. The KMC plot was generated by averaging over about 10^7 sample trajectories, with a crystal of size $N = 1500$. We see that the two plots are very different; the reason we compare the two profiles at slightly different times will be explained below.

Figure 1: $h_T - h_0$ from KMC and PDE



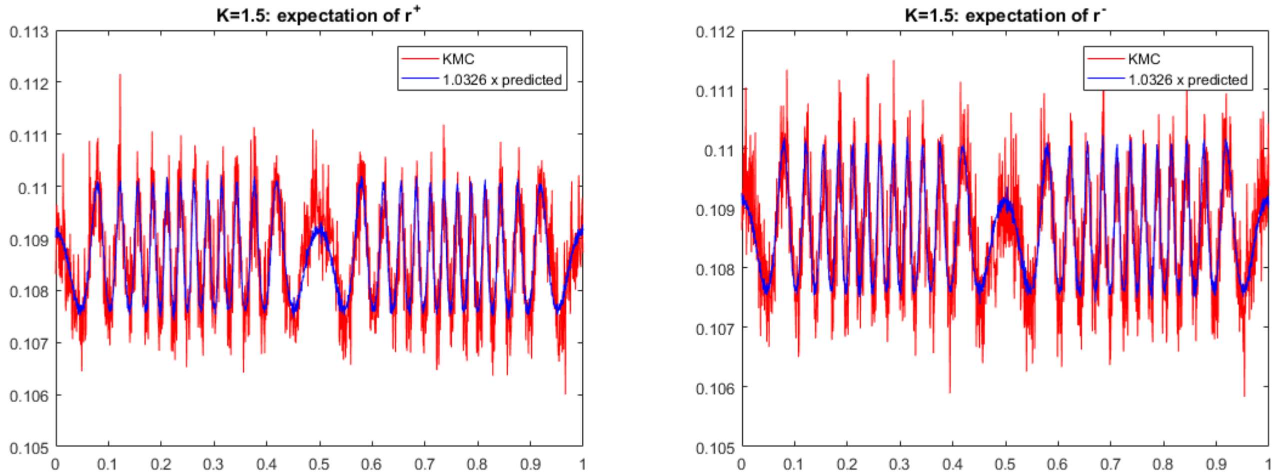
Since computing the rate expectation was the most involved part of the derivation, we investigate whether it was the source of error. We do so by computing the average over many runs of the rates associated to each site of the profile obtained from KMC at time $T = 2 \times 10^{-9}$ and comparing to the formulas for $\langle r^{\pm}(j) \rangle_\lambda$ that we obtained. The plots comparing the rate expectations from KMC and from the theory are shown in Figure 2.

Figure 2: Rate expectations from KMC vs. theory



The plots in Figure 3 seem to suggest that the rates we computed are off only by a multiplicative constant.

Figure 3: Empirical rate expectations from KMC compared to scaled rate expectations from theory

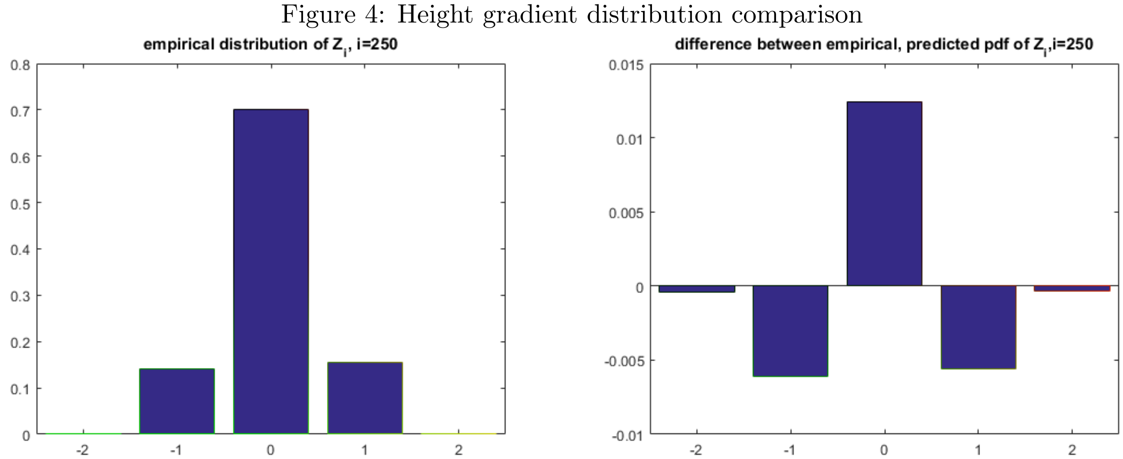


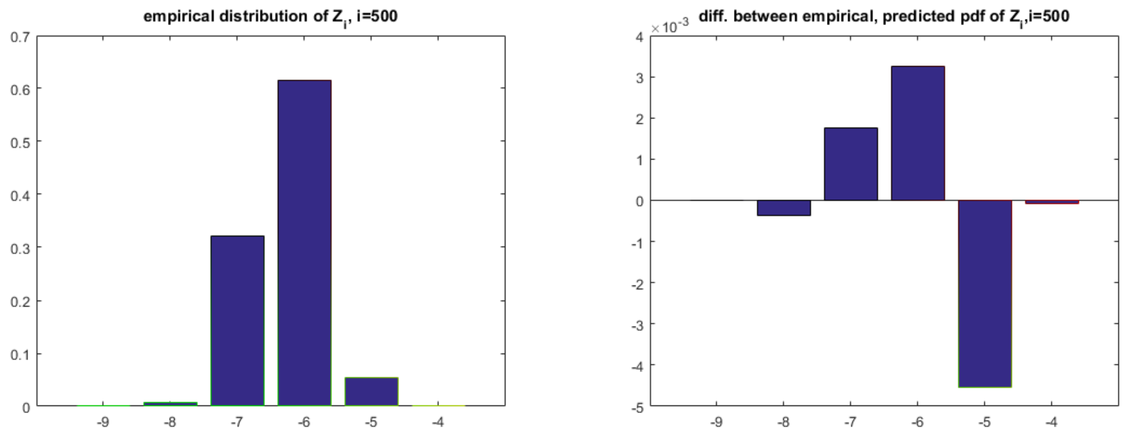
But this cannot be true, since such a scaling error would simply lead to a time scaling difference in the solution to the PDE, and we see this is not the case from the plots of $h_T - h_0$ in Figure 1. One possible explanation for the discrepancy between the predicted and computed rate expectation is that the height profiles were not distributed according to the local equilibrium measure at macroscopic time $T = 2 \times 10^{-9}$. Figure 4 compares the empirical distribution of $z_i = h_{i+1} - h_i$ based on 2×10^5 simulated paths of the microscopic process with $K = 1.5$ and $T = 2 \times 10^{-9}$ to the expected distribution,

$$P(z_i = n) \propto \exp(-Kn^2 + \lambda_i n) \propto \exp(-K(n - \frac{\lambda_i}{2K})^2),$$

where $\lambda_i = \nabla \sigma_D(\nabla h(T, \frac{i}{N}))$. Specifically, we plot n vs $P(z_i = n)$ for those n such that the probability is nonnegligible.

Figure 4: Height gradient distribution comparison

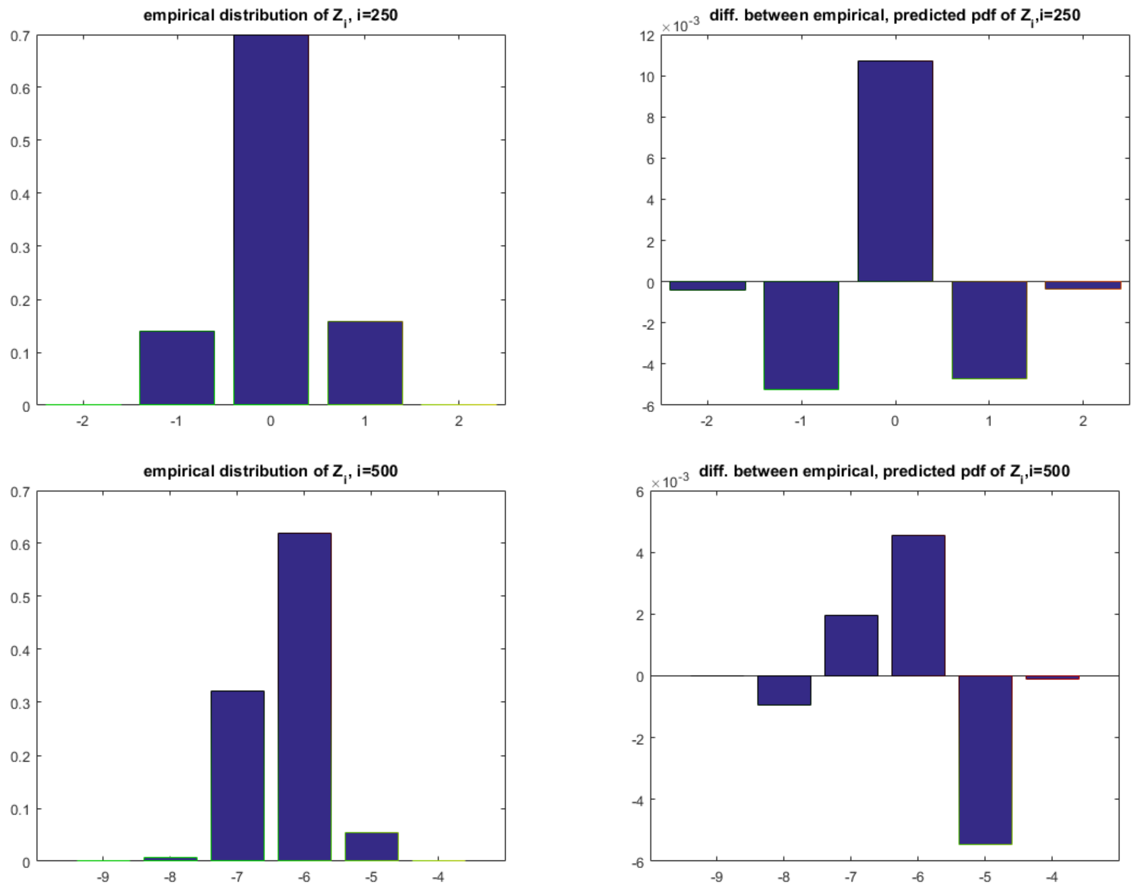




This slight discrepancy in the measure of the marginal distributions of Z_i may arise because λ_i does not equal $\nabla\sigma_D(\nabla h(t, \frac{i}{N}))$ exactly.

We carried out all of these numerical simulations for the process driven by the adatom rates studied in [4] as well. Interestingly, as Figure 5 shows, the difference between the empirical and predicted height gradient distributions was on the same order and had the same shape as that of the process driven by the Metropolis type rates.

Figure 5: Height gradient distribution comparison for process driven by adatom rates



Moreover, the rate expectation also did not coincide exactly with the formula obtained for it in [4],

$\langle r(j) \rangle_\lambda = \frac{1}{2} e^{-(\lambda_j - \lambda_{j-1})}$ (see Figure 6), while the fluxes $h_T - h_0$ from KMC and the derived PDE nearly coincided (see Figure 7).

Figure 6: Adatom rate expectations from KMC vs. theory

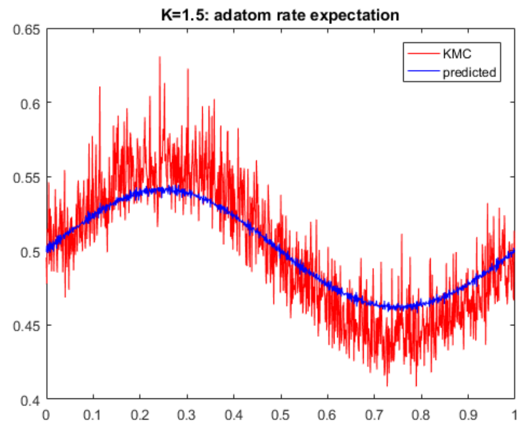
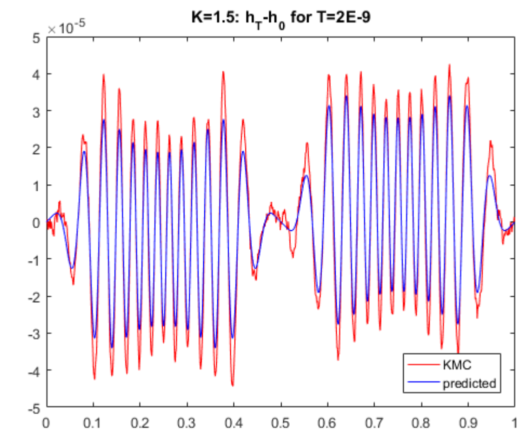


Figure 7: $h_T - h_0$ from KMC and PDE for adatom rate driven process



For the adatom rates, the discrepancy in the rate average may indeed be due only to a small multiplicative constant. Note, in particular, that the formula for the expectation does not have the $Z(\lambda)$ term that the Metropolis-type rate expectation formula has. These rates' additional dependence on λ could make the rate expectation more sensitive to error in λ.

The error in the predicted rate average and flux was smaller for $K = 0.5$; for example, the ratio of the rate average from KMC to the predicted rate average is about 1.0094. Since for smaller K the microscopic system behaves more like its continuous analog, where assuming independence of the Z_i seems more justifiable, this suggests that the delicate calculation of the λ_is or of $\frac{\mathbb{E}^{\pm}}{\mathbb{E}}$ for finite N and integer distributions lies at the heart of these errors. Figures 8 and 9 were generated by averaging over about 10^8 trajectories, with a crystal of size $N = 1000$. The flux $h_{\text{KMC}}(T, \cdot) - h(0, \cdot)$ in Figure 9 was smoothed by convoluting with a Gaussian filter using a window of 30 lattice sites.

Figure 8: Rate expectations from KMC vs. theory

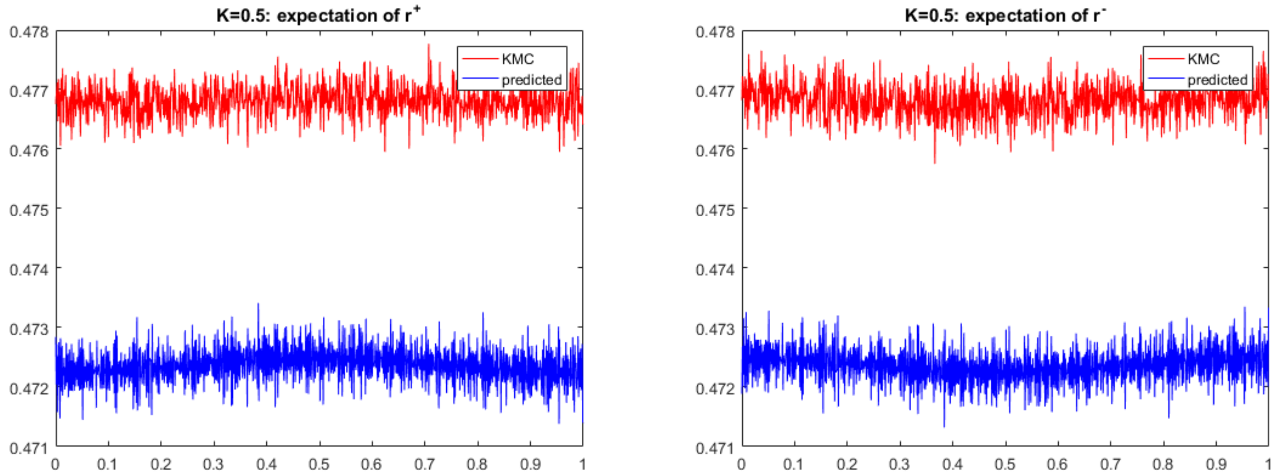
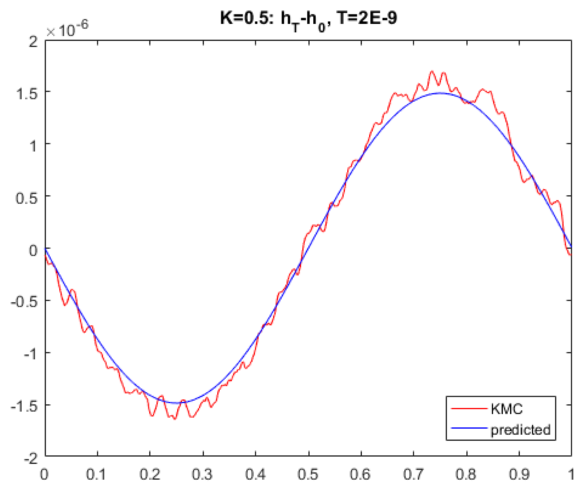


Figure 9: $h_T - h_0$ from KMC and PDE



6 Limit Computations

This section contains the calculations of limits appearing in the derivation of the PDE. The limits required estimation of the distribution of a sum of integer random variables with unbounded variance; since this is not a standard result in the literature, we prove it in the local limit theorems of Section 3. The second local limit theorem was thought to be necessary for the limit calculations, but turned out to be less useful than the first. We include it because the ideas of its proof were the basis for the proof of the first local limit theorem.

6.1 Proving $\mathbb{E}^+/\mathbb{E} \to 1$

Recall that $\mathbb{E}^+$ and $\mathbb{E}$ are expectations of a function of $S_N = \sum_{i=1}^{N-2} Z_i$ and $T_N = \sum_{i=1}^{N-2} iZ_i$, where the Z_i's are independent integer-valued random variables with distribution given by $P(Z_i = n) \propto e^{-K(n-(l_i+c_i))^2}$ and $P(Z_i = n) \propto e^{-K(n-l_i)^2}$ respectively. The function f is given by

$$f(S_N,T_N) = \exp\left[-\frac{K}{2}(S_N+l)^2\right]\sum_{j=-\infty}^{\infty}\exp\left[-2K(Nj+M+T_N+\frac{S_N}{2})^2\right],$$

where we have let $l = l_N + l_{N-1}$ and $M = m + \frac{l_N - l_{N-1}}{2}$ for brevity.

Note that the value of the sum depends only on $T_N + \frac{S_N}{2} \mod N$ or equivalently, $2T_N + S_N \mod 2N$. Define $g(k) = \sum_j \exp\left[-2K(M+\frac{k}{2}+Nj)^2\right], k \in \mathbb{Z}_{2N}$ and $h(m) = \exp\left[-\frac{K}{2}(m+l)^2\right]$. Then

$$\begin{aligned}\mathbb{E}\left[f(S_N,T_N)\right] &= \sum_{k=0}^{2N-1} g(k) \sum_{(m,n):m+2n\equiv k \mod 2N} h(m)P(S_N=m,T_N=n)\\ &= \sum_{k=0}^{2N-1} g(k)\sum_{m\in\mathbb{Z}} h(m)P(S_N=m,T_N\equiv\frac{k-m}{2} \mod N)\end{aligned} \tag{29}$$

Using Theorem 1 of Section 6.3, there exists a $c \in (0,1)$ such that

$$P(S_N=m,T_N\equiv k \text{ mod } N) = \frac{1}{N}P(S_N=m)+\epsilon(k,m) \tag{30}$$

and $|\epsilon(k,m)| \le c^N$ for all $k \in \mathbb{Z}_N$ and $m \in \mathbb{Z}$.

Substituting this expression for the probability and noting that k and m must have the same parity for the probability to be nonzero, we arrive at

$$\begin{aligned}\mathbb{E}\left[f(S_N,T_N)\right] = \frac{1}{N}\sum_{k\in\mathbb{Z}_{2N},k \text{ odd}} g(k)\sum_{m\in\mathbb{Z},m \text{ odd}} h(m)P(S_N=m) + \frac{1}{N}\sum_{k\in\mathbb{Z}_{2N},k \text{ even}} g(k)\sum_{m\in\mathbb{Z},m \text{ even}} h(m)P(S_N=m)\\ + \sum_{k\in\mathbb{Z}_{2N},m\in\mathbb{Z},k\equiv m \text{ mod } 2} g(k)h(m)\epsilon(m,\frac{k-m}{2})\end{aligned} \tag{31}$$

Note that

$$\sum_{k\in\mathbb{Z}_{2N},k \text{ even}} g(k) = \sum_{j\in\mathbb{Z}}\sum_{k'=0}^{N-1}\exp(-2K(Nj+k'+m+\frac{1}{2}(l_N-l_{N-1}))^2) = \sum_{n\in\mathbb{Z}}\exp(-2K(n+\frac{1}{2}(l_N-l_{N-1}))^2),$$

where we have used that m, the sum of all the heights, is an integer. Since $l_N - l_{N-1} \to 0$ as $N \to \infty$, we have

$$\sum_{k\in\mathbb{Z}_{2N},k \text{ even}} g(k) \to \sum_{n\in\mathbb{Z}} e^{-2Kn^2} =: c_0.$$

Similarly,

$$\sum_{k\in\mathbb{Z}_{2N},k \text{ odd}} g(k) \to \sum_{n\in\mathbb{Z}} e^{-2K(n+\frac{1}{2})^2} =: c_1. \tag{32}$$

Further, we can employ a local limit theorem [7] to write $P(S_N = m)$ as

$$P(S_N=m) = \frac{1}{\sqrt{2\pi B_N}}e^{-(m-M_N)^2/(2B_N)} + \epsilon_M, \tag{33}$$

where $M_N = \sum_{j=1}^{N-2}\mathbb{E}Z_j, B_N = \sum_{j=1}^{N-2}\text{Var}Z_j$, and $\sup_m|\epsilon_m| = o(\frac{1}{\sqrt{B_N}})$. Note that $B_N = O(N)$ and $M_N = O(1)$. Indeed, recall that the λ_j were chosen so that $\mathbb{E}Z_j = h'(\frac{j}{N},t)$ where h is the PDE solution. Using the periodicity of h, we have

$$|\frac{1}{N}\sum_{j=1}^{N}\mathbb{E}Z_j| = |\frac{1}{N}\sum_{j=1}^{N}h'(\frac{j}{N},t) - \int_0^1 h'(x,t)dx|,$$

which is on the order of $\frac{1}{N}$ if, for example, h' is Lipschitz continuous. Thus M_N is $O(1)$.

Substituting (33) and (32) into (31) yields

$$\begin{aligned}\mathbb{E}\left[f(S_N,T_N)\right] = &\frac{1}{N}\frac{c_0}{\sqrt{2\pi B_N}}\sum_{m\in 2\mathbb{Z}} e^{-\frac{K}{2}(m+\lambda)^2}e^{-(m-M_N)^2/2B_N}\\ &+\frac{1}{N}\frac{c_1}{\sqrt{2\pi B_N}}\sum_{m\in 2\mathbb{Z}+1} e^{-\frac{K}{2}(m+\lambda)^2}e^{-(m-M_N)^2/2B_N}\\ &+\frac{c_0+c_1}{N}\sum_{m\in\mathbb{Z}} e^{-\frac{K}{2}(m+\lambda)^2}\epsilon_m\\ &+\sum_{k\in\mathbb{Z}_{2N},m\in\mathbb{Z},k\equiv m \bmod 2} g(k)e^{-\frac{K}{2}(m+\lambda)^2}\epsilon(m,\frac{k-m}{2}).\end{aligned} \tag{34}$$

Since λ is $O(1)$ and $e^{-(m-M_N)^2/2B_N}$ is $O(1)$ in any finite neighborhood of $m=0$, the first two terms are of order $\frac{1}{N\sqrt{B_N}}O(1)=O(\frac{1}{N\sqrt{N}})$, while the third term is of order $o(\frac{1}{N\sqrt{N}})$. Also, we have

$$\left|\sum_{k\in\mathbb{Z}_{2N},m\in\mathbb{Z},k\equiv m \bmod 2} g(k)e^{-\frac{K}{2}(m+\lambda)^2}\epsilon(m,\frac{k-m}{2})\right| \le c^N(c_0+c_1)\sum_{m\in\mathbb{Z}} e^{-\frac{K}{2}(m+\lambda)^2} = c^N O(1),$$

which is negligible compared to the other terms.

Thus, we may only leave the first two terms of (34) in the numerator and denominator of $\frac{\tilde{\mathbb{E}}}{\mathbb{E}}$. The only variables depending on the values of the λ_js are M_N and B_N. But $\exp(\frac{(m-M_N)^2}{2B_N})\to 1$ as $N\to\infty$ for any finite m, both in the numerator and denominator. Hence, for large N,

$$\frac{\mathbb{E}^+}{\mathbb{E}} \approx \sqrt{\frac{B_N}{B_N^+}} \to 1 \text{ as } N\to\infty,$$

because only $\mathrm{Var}Z_i, i=j-1,j,j+1$ are different in the summands that make up B_N^+.

6.2 Proving $N(E_j - E_{j-1}) \to 0$

Let $S_N^j = \sum_{i\neq j-1,j,j+1} Z_i + \sum_{i=j-1,j,j+1}\tilde{Z}_i$ (the random variable factoring in to the expectation in the numerator of E_j, and similarly for $j-1$. Let $S_N' = \sum_{i\neq j-2,j-1,j,j+1} Z_i$, $X_j = S_N^j - S_N'$, $X_{j-1} = S_N^{j-1}-S_N'$. We use (30) to write

$$E_j - E_{j-1} = \frac{c_0\sum_{m\in 2\mathbb{Z}} e^{-\frac{K}{2}(m+\lambda)^2}(P(S_N^j=m)-P(S_N^{j-1}=m)) + c_1\sum_{m\in 2\mathbb{Z}+1} e^{-\frac{K}{2}(m+\lambda)^2}(P(S_N^j=m)-P(S_N^{j-1}=m))}{c_0\sum_{m\in 2\mathbb{Z}} e^{-\frac{K}{2}(m+\lambda)^2}P(S_N=m)+c_1\sum_{m\in 2\mathbb{Z}+1} e^{-\frac{K}{2}(m+\lambda)^2}P(S_N=m)}, \tag{35}$$

where we have disregarded the remainder terms because they are negligible in comparison to the other terms. Now, we have

$$P(S_N^j=m) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\phi_{S_N'}(t)\phi_{X_j}(t)e^{-itm}dt,$$

and similarly for $j-1$. Thus,

$$P(S_N^j=m)-P(S_N^{j-1}=m) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\phi_{S_N'}(t)(\phi_{X_j}(t)-\phi_{X_{j-1}}(t))e^{-itm}dt.$$

By the lemma in the next section, we can find $c\in(0,1)$ such that $\sup_i\sup_{\epsilon<|t|<\pi}|\phi_{Z_i}(t)|<c$. Thus, since $\phi_{S_N'}(t)=\prod_{i\neq j-2,\ldots,j+1}\phi_{Z_i}(t)$, there exists a positive constant A such that

$$\left|\frac{1}{2\pi}\int_{\epsilon<|t|<\pi}\phi_{S_N'}(t)(\phi_{X_j}(t)-\phi_{X_{j-1}}(t))e^{-itm}dt\right| \le Ac^N.$$

To compute the remaining part of the integral, we first write $\phi_{S'_N}(t) = \phi_{S'_N - M'_N}(t)e^{iM'_N t}$, where $M'_N = \mathbb{E}S'_N$,

$$\phi_{X_j}(t) = 1 + i\mathbb{E}X_j t - \frac{\mathbb{E}X_j^2}{2}t^2 + o(t^2),$$

and similarly for $\phi_{X_{j-1}}$. Substituting these expressions into the integral, we obtain

$$P(S_N^j = m) - P(S_N^{j-1} = m) \approx \frac{1}{2\pi}\int_{-\epsilon}^{\epsilon} \phi_{S'_N - M'_N}(t)((\mathbb{E}X_j - \mathbb{E}X_{j-1})it - \frac{\mathbb{E}X_j^2 - \mathbb{E}X_{j-1}^2}{2}t^2 + o(t^2))e^{-it(m - M'_N)}dt.$$

Now,

$$\begin{aligned}\phi_{S'_N - M'_N}(t) &= \prod_{i \neq j-2,\ldots,j+1} \phi_{Z_i - \mu_i}(t) = \prod_{i \neq j-2,\ldots,j+1} (1 - \frac{\sigma_i^2}{2}t^2 + o(t^2)) \\ &= \exp\left[\sum_i \log(1 - \frac{\sigma_i^2}{2}t^2 + o(t^2))\right] = \exp\left[-\frac{\sum_i \sigma_i^2}{2}t^2 + \sum_i o(t^2)\right].\end{aligned} \tag{36}$$

We now change variables to $t = \frac{x}{\sqrt{N}}$, which yields

$$\begin{aligned}P(S_N^j = m) - P(S_N^{j-1} = m) \approx \frac{1}{2\pi\sqrt{N}}\int_{-\epsilon\sqrt{N}}^{\epsilon\sqrt{N}} \exp\left[-\frac{\sum_i \sigma_i^2}{2N}x^2 + o(1)\right]((\mathbb{E}X_j - \mathbb{E}X_{j-1})i\frac{x}{\sqrt{N}} - \frac{\mathbb{E}X_j^2 - \mathbb{E}X_{j-1}^2}{2N}x^2 + o(\frac{x^2}{N})) \\ \times e^{-i\frac{x}{\sqrt{N}}(m - M'_N)}dt.\end{aligned} \tag{37}$$

Let $\sigma_N^2 = \frac{\sum_i \sigma_i^2}{2N}$. Using the formulas

$$\int_{-\mathbb{R}} xe^{-ax^2+bx}dx = \frac{\sqrt{\pi}b}{2a^{\frac{3}{2}}}e^{\frac{b^2}{4a}}, \int_{-\mathbb{R}} x^2 e^{-ax^2+bx}dx = \frac{\sqrt{\pi}(2a+b^2)}{4a^{\frac{5}{2}}}e^{\frac{b^2}{4a}}$$

(the part of the integral outside $[-\epsilon\sqrt{N}, \epsilon\sqrt{N}]$ is exponentially small), we obtain

$$\begin{aligned}&P(S_N^j = m) - P(S_N^{j-1} = m) \approx \\ &\quad \frac{1}{2\pi\sqrt{N}}\frac{\sqrt{2\pi}}{N\sigma_N^3}\exp\left[-\frac{(m - M'_N)^2}{2N\sigma_N^2}\right]\left((\mathbb{E}X_j - \mathbb{E}X_{j-1})(m - M'_N) - \frac{\mathbb{E}X_j^2 - \mathbb{E}X_{j-1}^2}{2}(1 - \frac{(m - M'_N)^2}{N\sigma_N^2})\right).\end{aligned} \tag{38}$$

Now, recall from the previous section that $m - M'_N = O(1)$ for any fixed m,and

$$P(S_N = m) \approx \frac{1}{\sqrt{2\pi B_N}}\exp\left[-\frac{(m - M_N)^2}{2B_N}\right].$$

We can simply replace the exponential in the probabilities with 1. Note that B_N is of order N, so that

$$N(E_j - E_{j-1}) \approx O(1)\frac{c_0 \sum_{m \in 2\mathbb{Z}} e^{-\frac{K}{2}(m+\lambda)^2}\left((\mathbb{E}X_j - \mathbb{E}X_{j-1})(m - M'_N) - \frac{\mathbb{E}X_j^2 - \mathbb{E}X_{j-1}^2}{2}\right) + c_1 \sum_{m \in 2\mathbb{Z}+1} \cdots}{c_0 \sum_{m \in 2\mathbb{Z}} e^{-\frac{K}{2}(m+\lambda)^2} + c_1 \sum_{m \in 2\mathbb{Z}+1} e^{-\frac{K}{2}(m+\lambda)^2}}.$$

We now show that $\mathbb{E}X_j - \mathbb{E}X_{j-1}$ and $\mathbb{E}X_j^2 - \mathbb{E}X_{j-1}^2$ go to 0 as $N \to \infty$, which will conclude the proof.

Let $\mu(\lambda)$ be the expectation and $\sigma^2(\lambda)$ the variance of the discrete random variable Z with parameter λ. Then

$$EX_j = \mu(\lambda_{j-2}) + \mu(\lambda_{j-1} - \frac{1}{2}) + \mu(\lambda_j + 1) + \mu(\lambda_{j+1} - \frac{1}{2}) = \mu(\lambda_{j-2}) + \mu(\lambda_{j-1} - \frac{1}{2}) + 1 + \mu(\lambda_j) + \mu(\lambda_{j+1} - \frac{1}{2}).$$

Similarly,

$$EX_{j-1} = \mu(\lambda_{j-2} - \frac{1}{2}) + 1 + \mu(\lambda_{j-1}) + \mu(\lambda_j - \frac{1}{2}) + \mu(\lambda_{j+1}).$$

Thus,

$$EX_j - EX_{j-1} = \Big(\mu(\lambda_{j-2}) - \mu(\lambda_{j-2} - \frac{1}{2})\Big) + \Big(\mu(\lambda_{j-1} - \frac{1}{2}) - \mu(\lambda_{j-1})\Big) + \Big(\mu(\lambda_j) - \mu(\lambda_j - \frac{1}{2})\Big) + \Big(\mu(\lambda_{j+1} - \frac{1}{2}) - \mu(\lambda_{j+1})\Big).$$

But as $N \to \infty$, the λ_is, $i = j-2, j-1, j, j+1$ all approach the same value, so we see that $EX_j - EX_{j-1} \to 0$.

Similarly,

$$\begin{aligned}\mathrm{Var}X_j - \mathrm{Var}X_{j-1} &= \Big(\sigma^2(\lambda_{j-2}) - \sigma^2(\lambda_{j-2} - \frac{1}{2})\Big) + \Big(\sigma^2(\lambda_{j-1} - \frac{1}{2}) - \sigma^2(\lambda_{j-1})\Big) \\ &+ \Big(\sigma^2(\lambda_j) - \sigma^2(\lambda_j - \frac{1}{2})\Big) + \Big(\sigma^2(\lambda_{j+1} - \frac{1}{2}) - \sigma^2(\lambda_{j+1})\Big) \to 0\end{aligned} \tag{39}$$

as $N \to \infty$, because $\sigma^2(\lambda_j + 1) = \sigma^2(\lambda_j)$ (appearing in X_j) and $\sigma^2(\lambda_{j-1} + 1) = \sigma^2(\lambda_{j-1})$ (appearing in X_{j-1}). It follows that

$$EX_j^2 - EX_{j-1}^2 = \mathrm{Var}X_j - \mathrm{Var}X_{j-1} + (EX_j)^2 - (EX_{j-1})^2 \to 0.$$

6.3 Lattice Local Limit Theorems

Theorem 1. *Let $S_N = \sum_{j=1}^n Z_j$ and $T_N = \sum_{j=1}^n jZ_j$, where $P(Z_j = n) \propto e^{-K(n-\lambda_j)^2}$ if $n \in \mathbb{Z}$ and 0 otherwise, and $(\lambda_1, \ldots, \lambda_N) \in \mathbb{R}^N$. Then there exists $c \in (0,1)$ such that*

$$|P(S_N = m, T_N \equiv n \text{ mod } N) - \frac{1}{N} P(S_N = m)| \leq c^N$$

for all $m \in \mathbb{Z}, n = 0, \ldots, N-1$.

Lemma 1. *Define $\phi(\lambda, t) = \mathbb{E}(e^{iZ_\lambda t})$, where Z_λ is the integer-valued random variable for which $P(Z_\lambda = n) \propto e^{-K(n-\lambda)^2}, n \in \mathbb{Z}$. Then*

$$\sup_{\lambda \in \mathbb{R}, \epsilon \leq |t| \leq \pi} |\phi(\lambda, t)| < 1$$

for any $\epsilon \in (0, \pi]$.

Proof. We have

$$|\phi(\lambda, t)| = \frac{|\sum_{n \in \mathbb{Z}} e^{int} e^{-K(n-\lambda)^2}|}{\sum_{n \in \mathbb{Z}} e^{-K(n-\lambda)^2}} = \frac{|\sum_{n \in \mathbb{Z}} e^{i(n-\lfloor\lambda\rfloor)t} e^{-K(n-\lambda)^2}|}{\sum_{n \in \mathbb{Z}} e^{-K(n-\lambda)^2}}.$$

Letting $m = n - \lfloor\lambda\rfloor$ and summing over m in the numerator and denominator, we obtain

$$|\phi(\lambda, t)| = \frac{|\sum_{m \in \mathbb{Z}} e^{imt} e^{-K(m-(\lambda-\lfloor\lambda\rfloor))^2}|}{\sum_{m \in \mathbb{Z}} e^{-K(m-(\lambda-\lfloor\lambda\rfloor))^2}} = |\phi(\lambda \bmod 1, t)|.$$

Thus, we can replace the supremum with λ ranging over $\mathbb{R}$ to $\lambda \in [0,1]$. It is not difficult to show that ϕ is jointly continuous in (t, λ) on the compact set $\{\lambda \in [0,1], \epsilon \leq |t| \leq \pi\}$, so it attains a maximum at some (λ_0, t_0). But for any fixed λ, $|\phi(\lambda, t)| < 1$ whenever $0 < |t| \leq \pi$, so the supremum must be strictly less than 1. □

Proof. Note that

$$1_{\{S_N = m, T_N \equiv n\}} = \big(\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{it(S_N - m)} dt\big)\big(\frac{1}{N} \sum_{z \in U_N} z^{T_N - n}\big),$$

where U_N is the set of Nth roots of unity. Similarly expressing $1_{\{S_N=m\}}$ and taking expectations, we obtain

$$\begin{aligned}
&P(S_N = m, T_N \equiv n \text{ mod } N) - \frac{1}{N}P(S_N = m) \\
&= \frac{1}{2\pi}\mathbb{E}\big[\big(\int_{-\pi}^{\pi} e^{it(S_N-m)}dt\big)\big(\frac{1}{N}\sum_{z\in U_N} z^{T_N-n} - \frac{1}{N}\big)\big] = \frac{1}{2\pi N}\mathbb{E}\big[\big(\int_{-\pi}^{\pi} e^{it(S_N-m)}dt\big)\big(\sum_{z\in U_N, z\neq 1} z^{T_N-n}\big)\big] \\
&= \frac{1}{2\pi N}\sum_{k=1}^{N-1}(e^{\frac{2\pi ik}{N}})^{-n}\int_{-\pi}^{\pi}\mathbb{E}[e^{itS_N}e^{\frac{2\pi ik}{N}T_N}]e^{-itm}dt = \frac{1}{2\pi N}\sum_{k=1}^{N-1}(e^{\frac{2\pi ik}{N}})^{-n}\int_{-\pi}^{\pi}\mathbb{E}\big[\exp\big(\sum_{j=1}^{N} iZ_j(t+\frac{2\pi kj}{N})\big)\big]e^{-itm}dt \\
&= \frac{1}{2\pi N}\sum_{k=1}^{N-1}(e^{\frac{2\pi ik}{N}})^{-n}\int_{-\pi}^{\pi}\prod_{j=1}^{N}\phi_{Z_j}(t+\frac{2\pi kj}{N})e^{-itm}dt.
\end{aligned} \tag{40}$$

Taking absolute values yields

$$\begin{aligned}
&|P(S_N = m, T_N \equiv n \text{ mod } N) - \frac{1}{N}P(S_N = m)| \\
&\leq \frac{1}{2\pi N}\sum_{k=1}^{N-1}\int_{-\pi}^{\pi}\prod_{j=1}^{N}|\phi_{Z_j}(t+\frac{2\pi kj}{N})|dt \leq \frac{1}{N}\sum_{k=1}^{N-1}\sup_{t\in[-\pi,\pi]}\prod_{j=1}^{N}|\phi_{Z_j}(t+\frac{2\pi kj}{N})|.
\end{aligned} \tag{41}$$

Fix $0 < \epsilon << 1$ and cover the torus $[-\pi,\pi]$ (with endpoints identified) by disjoint intervals $I_1,\ldots,I_N$ of length $\frac{2\pi}{N}$ so that $[-\epsilon,\epsilon]$ lies entirely within some $N(\epsilon)$ intervals. By the lemma, there exists $c < 1$ such that $|\phi_{Z_j}(t)| \leq c$ for all j and $\epsilon \leq |t| \leq \pi$. Now, fix $t_0 \in [-\pi,\pi]$ and $1 \leq k \leq N-1$. Let $d = \gcd(k,N)$. Then the points $t_0 + \frac{2\pi k}{N},\ldots,t_0+2\pi k$ take each of $\frac{N}{d}$ values d times, with distance $\frac{2\pi d}{N}$ between neighboring values. Now, if $\frac{2\pi d}{N} > 2\epsilon$, then at most one of these values can be in the interval $[-\epsilon,\epsilon]$, and thus $t_0 + \frac{2\pi kj}{N} \notin [-\epsilon,\epsilon]$ for at least $N-d$ js. Since $\gcd(k,N) \leq \frac{N}{2}$, it follows that

$$\prod_{j=1}^{N}|\phi_{Z_j}(t_0+\frac{2\pi kj}{N})| \leq c^{\frac{N}{2}}.$$

If $\frac{2\pi d}{N} \leq 2\epsilon$, then at most $\lceil\frac{2\epsilon N}{2\pi d}\rceil \leq \frac{\epsilon N}{\pi d}+1$ of these values belong in $[-\epsilon,\epsilon]$, so $t_0+\frac{2\pi kj}{N} \notin [-\epsilon,\epsilon]$ for at least $N - d(\frac{\epsilon N}{\pi d}+1) \geq N(\frac{1}{2}-\frac{\epsilon}{\pi})$ js. Combining with the previous case, we see that

$$\prod_{j=1}^{N}|\phi_{Z_j}(t_0+\frac{2\pi kj}{N})| \leq c^{N(\frac{1}{2}-\frac{\epsilon}{\pi})}$$

for all t_0 and k. Thus

$$|P(S_N = m, T_N \equiv n \text{ mod } N) - \frac{1}{N}P(S_N = m)| \leq c^{(\frac{1}{2}-\frac{\epsilon}{\pi})N},$$

as desired.

□

Theorem 2. *Let $Z_j, j = 1,..,N$ be i.i.d random variables taking integer values such that $\mathbb{E}Z_1 = \mu, \mathrm{Var}Z_1 = \sigma^2$. Further, assume that there exist no $m,n \in \mathbb{Z}$, $n > 1$, such that $Z_1 \in m + n\mathbb{Z}$. Let $S_N = \sum_{j=1}^{N} jZ_j$. Then*

$$P(S_N = m) = \frac{1}{N\sqrt{N}}\big(\frac{\sqrt{3}}{\sqrt{2\pi}\sigma}e^{-\frac{3}{2\sigma^2N^3}(m-N\mu)^2} + o(1)\big).$$

Proof. Since S_N takes only integer values,

$$P(S_N = m) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi_{S_N}(t)e^{-imt}dt = \frac{1}{2\pi}\int_{-\pi}^{\pi}\prod_{j=1}^{N}\phi_{Z_j}(jt)e^{-imt}dt$$
$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\prod_{j=1}^{N}\phi_{Z_j-\mu}(jt)e^{it(N\mu-m)}dt. \tag{42}$$

From now on, we will use ϕ to denote $\phi_{Z_j-\mu}$. Let $t = \frac{x}{N\sqrt{N}}$. Then, changing variables,

$$P(S_N = m) = \frac{1}{2\pi}\frac{1}{N\sqrt{N}}\int_{-\pi N\sqrt{N}}^{\pi N\sqrt{N}}\prod_{j=1}^{N}\phi(\frac{j}{N}\frac{x}{\sqrt{N}})e^{ix(N\mu-m)/N\sqrt{N}}dx.$$

We now study the product of characteristic functions, $\Phi_N(t) := \prod_{j=1}^{N}\phi(jt)$ for $t = \frac{x}{N\sqrt{N}}$ and x fixed. Taylor expanding $\log\phi(t)$ to second order about $t = 0$, we get

$$\log\phi(t) = -\frac{\sigma^2}{2}t^2 + o(t^2).$$

Thus,

$$\log\Phi_N(\frac{x}{N\sqrt{N}}) = -\frac{\sigma^2}{2}\Big(\frac{1}{N}\sum_{j=1}^{N}(\frac{j}{N})^2\Big)x^2 + \sum_{j=1}^{N}o(\frac{1}{N}).$$

Carrying out the sum of squares, we see that $\Phi_N(\frac{x}{N\sqrt{N}}) \to e^{-\frac{1}{6}\sigma^2x^2}$ as $N \to \infty$. We will use this to show that

$$\frac{1}{2\pi}\int_{-\pi N\sqrt{N}}^{\pi N\sqrt{N}}\Phi_N(\frac{x}{N\sqrt{N}})e^{ix(N\mu-m)/N\sqrt{N}}dx = \frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-\frac{1}{6}\sigma^2x^2}e^{ix(N\mu-m)/N\sqrt{N}}dx + o(1), \tag{43}$$

which is sufficient because the integral on the right comes out to

$$\frac{\sqrt{3}}{\sqrt{2\pi}\sigma}e^{-\frac{3}{2\sigma^2N^3}(m-N\mu)^2},$$

as desired. First note that, since the integral on the right in (43) is finite over $\mathbb{R}$, it is sufficient to show that

$$\int_{-\pi N\sqrt{N}}^{\pi N\sqrt{N}}|\Phi_N(\frac{x}{N\sqrt{N}}) - e^{-\frac{1}{6}\sigma^2x^2}|dx \to 0.$$

Since we have pointwise convergence, we can apply the Lebesgue Dominated Convergence Theorem once we find an integrable function f such that $|\Phi_N(\frac{x}{N\sqrt{N}})| < f(x)$ for all N and all $x \in [-\pi N\sqrt{N}, \pi N\sqrt{N}]$. It will be convenient to let $2\pi t = \frac{x}{N\sqrt{N}}$, so that t ranges over $[-\frac{1}{2}, \frac{1}{2}]$. We split the domain into three sets and bound $|\Phi_N|$ by an integrable function on each of these domains.

First consider the domain $D_1 := \{\epsilon \le |t| \le \frac{1}{2}\}$ for some ϵ small, corresponding to $D_1' := \{2\pi N\sqrt{N}\epsilon \le |x| \le \pi N\sqrt{N}\}$. Note that

$$|\phi_{Z_j-\mu}(2\pi t)| = |\phi_{Z_j}(2\pi t)| = |\sum_{n\in\mathbb{Z}}e^{2\pi itn}P(Z_j = n)| \le 1,$$

with equality holding iff $e^{2\pi itn} = e^{2\pi itm}$ whenever $P(Z_j = n) > 0, P(Z_j = m) > 0$. By the assumption on the values that the Z_j take, this implies that equality holds iff $t = k \in \mathbb{Z}$. Thus, by continuity and periodicity of $|\phi(2\pi t)|$, there exists $c < 1$ such that

$$|\phi(2\pi t)| < c \text{ for all } t \in A := \bigcup_{k\in\mathbb{Z}}[k-\frac{1}{2}, k-\epsilon] \cup [k+\epsilon, k+\frac{1}{2}] = \{t : t \bmod 1 \in [\epsilon, 1-\epsilon]\}$$

We will give a uniform bound on $|\Phi_N(2\pi t)| = \prod_{j=1}^N \phi(2\pi jt)$ for all $t \in A$ (which contains D_1).To obtain this bound, we estimate the number of j's less than N for which $jt \in A$.

Define a cycle as a block of consecutive j's such that jt traverses A and A^C once, ending when jt enters A again. Note that since $t \mod 1 \geq \epsilon = O(1)$, a cycle takes $O(1)$ steps (i.e. will contain $O(1)$ consecutive j's). Hence there will be $O(N)$ cycles, so jt will land in both A and A^C $O(N)$ times. If t is irrational, by the ergodicity of the transformation $x \in [0,1] \mapsto x + t \bmod 1$, there will be approximately $(1-2\epsilon)N$ j's for which $jt \in A$. If t is rational, there will still be at least $\frac{rN}{3}$ such j's (where rN is the number of cycles) because in each cycle, $jt \in A$ at least once and in A^C at most twice (since $t \mod 1 \geq \epsilon$). Thus, $|\Phi_N(2\pi t)| \leq c^{\frac{rN}{3}}$ for $t \in D_1$. Switching back to x, we can find a δ such that

$$|\Phi_N(\frac{x}{N\sqrt{N}})| \leq c^{\frac{rN}{3}} \leq \frac{\delta}{x^2} 1_{\{|x|\geq 1\}}$$

for all $x \in D_1'$.

Next, consider the region $D_2 = \{\frac{1-\epsilon}{N} \leq |t| \leq \epsilon\}$ corresponding to $D_2' = \{2\pi N\sqrt{N}\frac{1-\epsilon}{N} \leq |x| \leq 2\pi N\sqrt{N}\epsilon\}$. From the symmetry of the set A with respect to 0, we will obtain the same bound for t positive and negative, so we restrict ourselves to t positive. Note that, unlike $t \in D_1$, there may not be $O(N)$ cycles through A and A^C for $t \in D_2$, so we have to more carefully estimate the number of j's for which $jt \in A$. jt will fall in A for $j \in S_m = \{\lfloor\frac{m-1+\epsilon}{t}\rfloor + 1, ..., \lfloor\frac{m-\epsilon}{t}\rfloor\}, m = 1, ...$ until these values exceed N. Let k be the maximum m for which $S_m \subset \{1, 2, ..., N\}$. Note that $k \geq 1$, since $1t \leq \epsilon$ and $Nt \geq 1-\epsilon$. We have

$$|S_m| = \lfloor\frac{m-\epsilon}{t}\rfloor - \lfloor\frac{m-1+\epsilon}{t}\rfloor \geq \frac{1-2\epsilon}{t} - 1,$$

so

$$|\{j : jt \in A\}| \geq k(\frac{1-2\epsilon}{t}) - k.$$

We now bound this count from below by a value depending only on N.

From the definition of k, it follows that

$$\frac{k-\epsilon}{t} - 1 \leq \lfloor\frac{k-\epsilon}{t}\rfloor \leq N < \lfloor\frac{k+1-\epsilon}{t}\rfloor \leq \frac{k+1-\epsilon}{t}. \tag{44}$$

From the left side of (44), we deduce

$$\frac{k-\epsilon}{N+1} \leq t \implies k \leq (N+2)\epsilon, \tag{45}$$

where we have used that $t < \epsilon$. From the right side of (44), we obtain

$$\frac{k(1-2\epsilon)}{t} \geq \frac{k}{k+1-\epsilon}(1-2\epsilon)N \geq \frac{1-2\epsilon}{2-\epsilon}N, \tag{46}$$

where we have used that $k \geq 1$. Combining (45) and (46) gives

$$k(\frac{1-2\epsilon}{t}) - k \geq (\frac{1-2\epsilon}{2-\epsilon} - \epsilon)N - 2\epsilon.$$

Hence,

$$|\Phi_N(2\pi t)| \leq (c^{-2\epsilon})c^{r_2 N}$$

for $r_2 = \frac{1-2\epsilon}{2-\epsilon} - \epsilon > 0$ if, for example, $\epsilon < \frac{1}{4}$. Switching back to x, we can bound $|\Phi_N(\frac{x}{N\sqrt{N}})|$ on D_2' with the same function as on D_1' (perhaps with a different constant δ).

Finally, consider the domain $D_3 = \{0 \leq |t| \leq \frac{1-\epsilon}{N}\}$, corresponding to $D_3' = \{0 \leq |x| \leq 2\pi N\sqrt{N}\frac{1-\epsilon}{N}\}$.Since $\phi(t) = 1 - \frac{\sigma^2}{2}t^2 + o(t^2)$, we can find $\delta > 0$ such that

$$|\phi(2\pi t)| \leq 1 - \frac{(2\pi\sigma)^2}{4}t^2 \leq e^{-\frac{(2\pi\sigma)^2}{4}t^2}$$

for all $|t| < \delta$. For all $t \in D_3$ and $j \leq j_0 := \lfloor \frac{\delta}{1-\epsilon} N \rfloor$, we have $|jt| < \delta$. Choose $c > 0$ such that $j_0 \geq cN$. Then

$$|\Phi_N(2\pi t)| \leq \prod_{j=1}^{j_0} |\phi(2\pi jt)| \leq e^{-(2\pi\sigma)^2 j_0(j_0+1)(2j_0+1)t^2/24} \leq e^{-(2\pi\sigma)^2 j_0^3 t^2/12} \leq e^{-(2\pi\sigma)^2 c^3 N^3 t^2/12}.$$

Substituting $2\pi t = \frac{x}{N\sqrt{N}}$, we get

$$|\Phi_N(\frac{x}{N\sqrt{N}})| \leq e^{-\sigma^2 c^3 x^2/12}$$

for $x \in D_3'$. We have bounded $|\Phi_N(\frac{x}{N\sqrt{N}})|$ by an integrable function over each of the domains, and can now apply LDCT to finish the proof.

□

References

[1] T. Funaki, *Stochastic Interface Models*, Lectures on Probability Theory and Statistics, Lecture Notes in Math., **1869**, Springer, Berlin (2005), 103-274.

[2] C. Kipnis and C. Landim *Scaling Limits of Interacting Particle Systems*, Springer (1999).

[3] J. Krug, H.T. Dobbs and S. Majaniemi, *Adatom mobility for the solidon-solid model*, Z. Phys. B **97** (1994), 281-291.

[4] J.L. Marzuola and J. Weare, *The relaxation of a family of broken bond crystal surface models*, Phys. Rev. E **88**(2013), 032403

[5] T. Nishikawa, *Hydrodynamic limit for the Ginzburg-Landau* $\nabla\phi$ *interface model with a conservation law*, J. Math. Sci. Univ. Tokyo **9** (2002), 481-519.

[6] G. Pavliotis, *Stochastic Processes and Applications*, Springer (2014).

[7] V.V. Petrov, *Sums of independent random variables* , Springer (1975).

[8] B. Roux and J. Weare, *On the statistical equivalence of restrained-ensemble simulations with the maximum entropy method*, The Journal of Chemical Physics **138**(2013).