論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Stochastic Heat Diffusion Modelling and the Discrete Laplace-Beltrami Operator on Non-Uniformly Gridded Spaces
著者(和文)	FrannekLukasJohann
Author(English)	Lukas Johann Frannek
出典(和文)	学位:博士（学術）, 学位授与機関:東京工業大学, 報告番号:甲第9674号, 授与年月日:2014年9月25日, 学位の種別:課程博士, 審査員:早川　朋久,木村　康治,井村　順一,天谷　賢治,中尾　裕也
Citation(English)	Degree:, Conferring organization: Tokyo Institute of Technology, Report number:甲第9674号, Conferred date:2014/9/25, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Stochastic Heat Diffusion Modelling and the Discrete Laplace-Beltrami Operator on Non-Uniformly Gridded Spaces

by

Lukas Frannek

Submitted to Department of Mechanical and Environmental Informatics
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

Tokyo Institute of Technology
August 2014

Abstract

The solution to the heat equation, the heat kernel, is a useful tool in various scientific contexts. Analytical solutions may be difficult to obtain due to the complexity of the partial differential equations in many engineering applications. Beside deterministic techniques such as the finite element and the finite difference methods, current stochastic techniques deal mostly with uniform grids. This thesis presents a new technique to approximate the diffusion of heat on non-uniformly gridded spaces with stochastic processes. Specifically, we use a Markov chain-driven random walk of a multitude of heat particles, where the infinitesimal generator of the Markov chain is characterized by the metric of the non-uniform grid imposed on a manifold. We show that the statistical properties of the distribution of particles on exemplary grids equals the theoretical distribution of heat on the same manifold, for a certain characterization of the Markov chain. When viewing a grid as a graph, the infinitesimal generator of the random walk on a graph is called graph Laplacian or discrete Laplace-Beltrami operator. Hence our framework also shows how to choose the weights of the graph Laplacian which itself can be used in numerous other applications.

Contents

List of Figures

Chapter 1

Introduction

In mathematics, diffusion equations are known as parabolic partial differential equations [1] and they can be found in numerous scientific applications. In finance, for instance, the Black-Scholes equation can be transformed into a diffusion equation [2]. The Fokker-Planck equation is also a diffusion equation and describes the evolution of the distribution function of a Brownian particle over time [3, 4]. The heat equation, the main equation in this research, is a special case of the Fokker-Planck equation which is also known as the Kolmogorov forward equation [5–8].

With the heat equation we are able to describe the heat distribution on surfaces and in solids [9]. There also are numerous applications of the heat equation in chemical sciences and material sciences, such as the modelling of particle diffusion in polymers [10]. Another application of the heat equation can be found in financial engineering, where option price models often have the same form as the heat equation [11]. Lastly, very important fields in which the theory of heat diffusion is used are video and image processing as well as machine-learning [12]. "Kernel methods", for instance, are a class of algorithms used in machine learning, where the solution to the heat equation is one of various kernels [13, 14]. Another example is the graph Laplacian which can be interpreted as the discrete version of the continuous Laplace-Beltrami operator which is used in the heat equation [15–19].

The research explained in this thesis aims to approximate the diffusion of heat through probabilistic means. Specifically, we view heat diffusion as the random walk of a multitude of heat particles [20]. The physical background is the vibration of atoms due to temperature differences and the excitation of nearby atoms. The natural phenomenon of

the agitation of nearby atoms can be viewed as a random process where, on a molecular level, heat particles randomly jump to all directions. It is therefore very natural to represent this phenomenon with particles which jump between nodes on a grid imposed on a manifold, where this movement is characterized by the metric defined on the manifold, i.e., the distances between nodes. If the jump distance is far the probability that the heat particle undertakes this jump is lower; if the distance is short the probability is higher.

In our framework, the random walk is driven by a continuous-time Markov chain and the transitions of particles per unit time between nodes is expressed by the transition rates. The transition rates are the elements of the infinitesimal generator of the continuous-time Markov chain and in this thesis we give the expressions for the transition rates such that certain statistical properties of the random variable of Markov chain are guaranteed for a given manifold. Since each instance of a Markov chain is viewed as a heat particle and the grid nodes are viewed as states of the Markov chain, we model the diffusion of heat by using a multitude of particles. We can then further assess the distribution of particles with statistical tools. We use the Gaussian distribution for comparison since the solution to the heat equation is a Gaussian function, or in other words, the Gaussian distribution with a certain mean and variance. This strong connection between random walks, Brownian motion, and diffusion processes will be discussed in more detail later on.

For regular Euclidean spaces with zero curvature, the Laplace operator is used in the heat equation. Since we aim to approximate the diffusion of heat on arbitrary curved spaces, we consider the Laplace-Beltrami operator instead of the Laplace operator. The infinitesimal generator of the Markov chain may be interpreted as the graph Laplacian or discrete Laplace-Beltrami operator. When we view the infinitesimal generator as the graph Laplacian, the transition rates mentioned above become the weights of the graph Laplacian.

In this thesis we introduce a new stochastic framework for the approximation of the diffusion of heat. Our main interest lies in the fundamental connection between a random walk of particles on non-uniformly spaced nodes, the discrete Laplace-Beltrami operator, and the diffusion of heat. Regarding the random walk, this leads to the question of how we have to characterize the movement of particles on a microscopic level such that the diffusion of heat on a macroscopic level is approximated. We pose the question if

limitations regarding the node spacing arise due to the space discretization, since heat diffuses isotropically in all directions. To approximate heat diffusion on a macroscopic level there are many well established deterministic techniques available. The main goal in this research is not the comparison with existing deterministic methods but rather the development of a new framework for the approximation of diffusion processes. The Markov Chain Monte Carlo method for instance is a tool used to solve complicated integrals [21,22]. Generally, with the Markov Chain Monte Carlo method, a posterior target distribution is approximated by using samples of that distribution. The framework in this research works in a similar way by using samples of a distribution, i.e., a heat particle, to approximate the distribution of heat particles. Later in this chapter we explain that the heat distribution, and solution to the heat equation, is a Gaussian distribution and also the probability density function of a Brownian particle. Consequently, it follows that each heat particle can be interpreted as a Brownian particle. The interesting question is that, for the random variable of the Markov chain, can we still obtain the same statistical properties that the probability density function of a Brownian particle has?

Other stochastic models for the approximation of diffusion processes use the randomness of, for instance, the thermal diffusivity or a random spacing of nodes. Unlike compartment-based models [23] where the probability of a certain amount of particles in a compartment of fixed size is considered, we consider the random movement of the particles themselves on a microscopic level. To illustrate the movement on a microscopic level, we utilize stochastic simulations of a multitude of particles. On a macroscopic level we want to determine the infinitesimal generator of the Markov chain, which allows us to employ deterministic techniques to find the distribution of heat. In essence, on a macroscopic level, we want to transform the heat equation, which is a partial differential equation, to an ordinary differential equation. Again, the question we pose is how we have to characterize a Markov chain-driven random walk on a non-uniformly spaced nodes, such that the random variable of the Markov chain has the same statistical properties as a Brownian particle on a microscopic level and such that the infinitesimal generator of the Markov chain can be used to deterministically obtain a heat distribution which is a good approximation of the analytical heat distribution on a macroscopic level.

Reference [8] describes the use of Poisson processes to model heat diffusion. The

Markov chain used in this research is a generalization of the Poisson process. Additionally, unlike frameworks such as in references [23,24], we investigate non-uniformly placed nodes. The main advantage of the usage of non-uniformly placed nodes is the ability of a smart placement of nodes on the manifold. We can utilize more nodes on segments of the grid where a higher accuracy is needed than on segments where this is not the case. On a manifold with varying curvature and in segments of that manifold with high curvature, we may require more nodes than in segments with low curvature on that same manifold. If the metric is not fully known, straight line approximations can be used by drawing a straight line between neighboring nodes. The lengths of these straight lines between each pair of neighboring nodes are the Euclidean distances. The error between the real geodesic distance between neighboring nodes on a curve and the distance between neighboring nodes when using straight line approximations is smaller if nodes are placed closer together. Segments with lower curvature on the same manifold do not need nodes with a short distance between them because the error between the real geodesic distance and the length of the straight line is already small even when nodes are placed far apart. In the case where the amount of employable nodes is limited, this smart placement of nodes reduces the amount of nodes needed for a good approximation of the diffusion of heat, especially on complicated manifolds where we are given limited information regarding the curvature. Note that in this thesis we do not give a treatise of grid generation techniques.

From a computational cost point of view, overall our framework offers a fast way to approximate the diffusion of heat on manifolds where the analytical solution to the heat equation is very hard to find or simply not available at all. Our framework is a straightforward technique and gives reasonably good approximations by guaranteeing certain statistical properties of the distribution of heat particles. With regard to deterministic numerical techniques such as the finite element method, good results cannot be guaranteed. This is due to the fact that the finite element method uses piecewise polynomial interpolation, i.e., each element of a surface is approximated by polynomials instead of functions that reflect the real curvature on a surface. And with deterministic methods, statistical properties of distributions are not considered. Furthermore, the fact that we use particle simulations enables the usage of parallel computing which leads to faster computation. The finite element method in particular offers the possibility of parallel computing as well. The problem

is that, in order to implement the parallel computation of the approximation of the diffusion of heat, additional special numerical solution methods are needed. Reference [25] writes about this point in more detail. Furthermore [25] claims that, when partitioning a grid in order to compute the results for the finite element methods, due to the boundaries between elements of the grid, an organized inter-node communication has to be set up. In our framework on the other hand, we can simulate each particle independently of other particles.

Another important point to note is that, in many deterministic numerical methods, the handling of spaces with no boundary conditions is difficult. The reason is that this requires the solution of equations with infinite-dimensional infinitesimal generator matrices. Reference [26] gives a particular example where a different stochastic methods is used to solve that kind of problem. As we show in later chapters, our framework does not have this limitation.

The notation in this thesis is fairly standard. Specifically, $\mathbb{R}$ denotes the set of real numbers, $\mathbb{R}_+$ denotes the set of positive real numbers, and $\mathbb{I}$ denotes the set of integer numbers. Furthermore, we write $(\cdot)^{\mathrm{T}}$ for transpose, I_2 for the identity matrix of dimension 2×2, $\langle \cdot, \cdot \rangle$ for the inner product, $\mathbb{P}[\cdot]$ for the probability of a random variable, and $\mathbb{E}[\cdot]$ for the expectation of a random variable.

1.1 Diffusion of Heat

In this section we give the foundation for the description of the diffusion of heat. In later chapters we specialize the heat equation, including the Laplace-Beltrami operator, for the respective manifold we are dealing with. We also give a short mention of deterministic techniques for the approximation of the diffusion of heat.

1.1.1 Heat Equation

The diffusion of heat is described by the heat equation given by

$$\frac{\partial u}{\partial t} = \zeta \nabla^2 u, \tag{1.1}$$

where u denotes the temperature, t denotes the time, ∇^2 denotes the Laplace-Beltrami operator, and ζ denotes the thermal diffusivity. Note that u does not have to be smooth. Equation (1.1) is the general expression for the heat equation and can be specialized to any dimension. Note that when considering arbitrary manifolds with varying curvature, instead of the regular Laplace operator, the Laplace-Beltrami operator is used in the heat equation. The information about the curvature and parametrization on the manifold are embedded in the Laplace-Beltrami operator. The Laplace-Beltrami operator is defined as

$$\nabla^2 = \frac{1}{\sqrt{|g|}} \sum_{i,j} \frac{\partial}{\partial x^i} (\sqrt{|g|} g^{ij} \frac{\partial}{\partial x^j}), \tag{1.2}$$

where g_{ij} is the (i,j)th element of the metric tensor matrix g defined on a given manifold and g^{ij} is the (i,j)th element of the inverse metric tensor matrix g^{-1}, the value $|g|$ denotes the absolute value of its determinant, and x^i denotes the ith local coordinates on the manifold. With the metric tensor g we are able to determine the geodesic distance between any two points on a given manifold. The geodesic distance can be generally computed with

$$(\mathrm{d}s)^2 = \sum_{i,j} g_{ij} \mathrm{d}x_i \mathrm{d}x_j, \tag{1.3}$$

where s is the geodesic distance. To find the distance between the positions a and b on a circle which was parameterized with angle θ, i.e., only one local coordinate $\mathrm{d}x_i = \mathrm{d}x_j = \mathrm{d}\theta$, we use (1.3) and get

$$s = \int_a^b \sqrt{g} \mathrm{d}\theta. \tag{1.4}$$

The geodesic distance is important because, in our framework, it determines the distance that a heat particle actually has to travel, whereas the Euclidean distance is not necessarily the travel distance. If the distance to one node is small compared to the distance to other nodes, naturally a heat particle is more likely to jump to the nearest node.

We view the heat equation from a more mathematical point of view. Specifically, we assume constant thermal conductivity, constant mass density and constant specific heat, and incorporate all these parameters into one parameter, the thermal diffusivity ζ.

1.1.2 Deterministic Methods

Since the heat equation is a partial differential equation, it may be difficult to obtain analytical solution when we consider complicated spaces, boundary conditions, initial conditions or spaces with non-zero curvature. We may also simply not be able to find the solution at all. In these cases numerical methods are used to still approximate the solution to the heat equation. Examples of deterministic methods include the finite element method [1, 27], the finite difference method [1] or the lattice Boltzman method [28, 29]. Specifically, for curved spaces, there are techniques as described in [30–32]. The disadvantage is that a satisfactory solution cannot be guaranteed, as stated in the previous section. Additionally, the error can become big if the grid is not refined enough. In our framework we can obtain a good solution by guaranteeing certain statistical properties of the distribution of particles, if the metric defined on the manifold is given. And even when we have limited information about the metric we can still use straight line approximations as mentioned in the previous section.

Another difference is the way results are obtained. In deterministic methods we consider a particular node on the grid and a particular time and then compute the temperature at the position of the node at that specific time. With our framework, and generally in stochastic simulations, we consider a histogram of particle positions. In other words, there is a certain probability that a particle is in a histogram segment at a given time. As mentioned in the introduction, this emphasizes the phenomenon of heat transfer that takes place on a molecular level and the idea that it is natural to model this phenomenon with random walks.

1.2 Random Walks

In the previous section we showed how we describe the diffusion of heat from a deterministic point of view. In this section we talk about the background of using stochastic processes to model heat diffusion. This includes the description of random walks as well as a formal treatment of the continuous-time Markov chain. We have to slightly change the latter as we go on to other manifolds in the following chapters in order to make the differences between manifolds clearer. We also give an explanation of the connection

between Brownian motion, heat diffusion and the continuous-time Markov chain, since this connection represents an important part in this research.

Random walks are the basic models for the description of stochastic behavior and are therefore used in many fields [33,34]. Generally, a random walk describes a path consisting of a succession of random jumps. In this thesis the random walk takes place on nodes placed on a given manifold. Furthermore, it is driven by a continuous-time Markov chain which is characterized by the metric, i.e., the distance between nodes. In this research we show how the metric determines the probability that the random walk makes a jump to a specific node from an given node on the same manifold.

1.2.1 Continuous-time Markov Chain

In this section we give a definition of the continuous-time Markov chain. Let $w(t)$ be the random variable of a continuous-time, time-homogeneous Markov chain on a countable space X. The variables x_i, $i = \ldots, -2, -1, 0, 1, 2, \ldots$, denote generic states of X. The Markov chain has the Markov property which tells us that the next state transition only depends on the current state and not on any previous state. We express the conditional probability of the time-homogeneous, continuous-time Markov chain as [35]

$$\mathbb{P}[w(s+t) = x_i | w(s) = x_j] = \mathbb{P}[w(t) = x_i | w(0) = x_j], \quad i, j \in \mathbb{I}, \tag{1.5}$$

for all $s \in \mathbb{R}_+$. We also define

$$p_{j,i}(t) \triangleq \mathbb{P}[w(t) = x_i | w(0) = x_j]. \tag{1.6}$$

The variables $p_{j,i}(t)$ are the elements of matrix $P(t)$ that is called the transition probability matrix and that describes the transitions of the Markov chain. As we will see in later chapters, the form of $P(t)$ depends on the actual manifold where the random walk takes place. The element $p_{j,i}(t)$ in (1.6) is the entry in row j and column i of $P(t)$. Associated with matrix $P(t)$ is the infinitesimal generator Q for the Markov chain (1.5), which

is defined as [35]

$$q_{j,i} = \begin{cases} \lim_{t\to 0} \frac{p_{j,i}(t)}{t}, & j \neq i, \\ \lim_{t\to 0} \frac{p_{j,j}(t)-1}{t}, & j = i. \end{cases} \tag{1.7}$$

Note that $q_{j,i} \geq 0$, $j \neq i$, and $q_{j,j} \leq 0$. The elements $q_{j,i}$ of Q are called transition rates and indicate the rate at which $w(t)$ transitions from state x_j to state x_i. Transitions from a state to itself are not permissible. Furthermore, Q satisfies the Kolmogorov forward equation given by

$$\frac{\mathrm{d}P(t)}{\mathrm{d}t} = P(t)Q, \quad P(0) = I, \tag{1.8}$$

with the solution

$$P(t) = e^{Qt}, \tag{1.9}$$

where I indicates the identity matrix.

In characterizing the probability that $w(t)$ is in state x_i at time t, given that $w(t)$ was in state x_0 at $t = 0$, we look at the entry $(e^{Qt})_{0,i}$. In this research we assume that the Markov chain always starts in state x_0. We consider the case where there is a point source of heat particles at time 0 located at the position corresponding to state x_0. Thus, we define the transition probability

$$p_i(t) \triangleq p_{0,i}(t) = (e^{Qt})_{0,i}, \tag{1.10}$$

which can be calculated as [36], [37]

$$p_i(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} (Q^k)_{0,i}. \tag{1.11}$$

1.2.2 Brownian Motion

Brownian motion is described as the random motion of particles suspended in a fluid. Einstein showed that, in one dimension, a Brownian particle doing Brownian motion has the probability of being at position x in time t after starting at position 0

$$p(x,t) = \frac{1}{(2\pi t)^{\frac{1}{2}}} e^{-\frac{x^2}{2t}}. \tag{1.12}$$

By relating the spatial displacements of the Brownian particle and time differences, Einstein showed that Brownian motion can be modelled as a diffusion equation, where (1.12) is the solution to the diffusion equation. Additionally, (1.12) is a Gaussian function and more precisely the expression for a Gaussian distribution with mean 0 and variance t which also solves the heat equation (1.1) in one dimension. As we show in later chapters, this similarity between the solution to the heat equation and the Gaussian distribution carries over into higher dimensions. The particle with probability density function (1.12) does a standard Brownian motion. That means that the corresponding Gaussian distribution has mean 0 and variance t. Therefore, the Gaussian distribution is the solution to one specific heat equation with a specific thermal diffusivity ζ. The solution to the heat equation with different thermal diffusivity is then not interpreted as the probability density function of a standard Brownian particles. It is interpreted as the probability density function of a Brownian particle doing Brownian motion multiplied by a constant.

Brownian motion has a continuous path, whereas the random walk, driven by a Markov chain, takes place in discrete space. Due to the known connection between the Markov chains and diffusion [38–40], the main point we build our proofs upon is the fact that even though we have a discrete state space we can show that the random variable of the Markov chain has the same mean and variance as the continuous Gaussian distribution which solves the heat equation.

1.3 Outline of the Thesis

We will now give a short outline of the thesis content. In Chapter 2 we introduce the basic framework and notation for approximation of the diffusion of heat. As we consider different manifolds the definition of the Markov chain in Chapter 1 will be extended accordingly. The technique we use for the proof to show that the random variable of the continuous-time Markov chain is distributed according to a Gaussian distribution is also extended.

In Chapter 3 we extend this framework to two-dimensional lattices. In two-dimensions it is important to choose an appropriate lattice. There are certain limitations we are confronted with, since we cannot guarantee that the random variable of the Markov chain

has a Gaussian distribution, if the nodes on the lattice are unfavorably placed.

Chapter 4 introduces the idea of smallest unit based lattices which allow for more types of lattices to be used with our framework.

Chapter 5 deals with the approximation of the diffusion of heat on closed curves. Important extensions are circular boundary conditions and the consideration of the Laplace-Beltrami operator, which becomes necessary since we now deal with spaces with non-zero curvature. We also introduce notions used in the field of directional statistics. The reason is that the notions of mean and variance are not applicable on closed curves. In that case we will utilize the trigonometric moments of the random variable of the Markov chain and other tools provided by directional statistics.

Chapter 6 shows empirical results for the sphere. The main problem on the sphere is the definition of a lattice. On the closed curve we can wrap the Gaussian distribution on the line around the curve. There is no rigorous definitions of a wrapped Gaussian distribution in two dimensions in a form which allows us to obtain theoretical results.

In Chapter 6 and 7 we have a detailed look at this problem. The last chapter gives a conclusion of our results as well as an outlook for further research directions.

Chapter 2

One-Dimensional Heat Diffusion Modelling

2.1 Introduction

In this chapter we develop the base for our framework for the approximation of the diffusion of heat. We introduce notions and expressions which are extended in later chapters. We give the definition of the continuous-time Markov chain on the infinite straight line and show how to characterize a Markov chain in order to approximate the diffusion of heat on the infinite straight line. This is done by determining the transition rates of the infinitesimal generator of the Markov chain. We mathematically describe the diffusion of heat in one dimension and show how the moments of the random variable of the Markov chain, the mean and variance, are related to the physical process. In addition, we provide an exemplary placement of nodes as well as simulations to demonstrate the efficacy of our framework.

2.2 Markov-Chain-Driven Random Walk on the Infinite Straight Line

In this section, we consider a Markov-chain-driven random walk on non-uniformly distributed nodes on the infinite straight line where all the particles are initially placed at node $x_0(= 0)$ on the real axis. Specifically, the nodes x_i, $i \in \mathbb{I}$, are unequally placed on

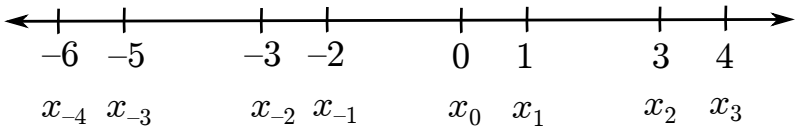

Figure 2.1: Example one of non-uniformly placed nodes on the real axis

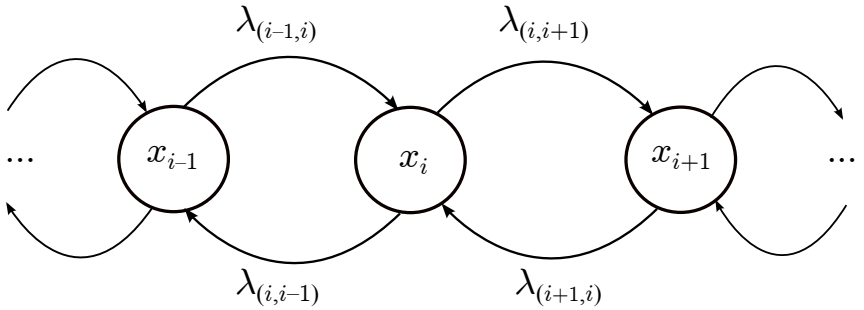

Figure 2.2: State-transition diagram of the continuous-time Markov chain on nodes in example one

the real axis as shown, e.g., in Figure 2.1, and let the particles randomly jump between the neighboring nodes.

In this case, the stochastic process of the successive random jumps is cast as a continuous-time Markov chain whose state transition diagram is described in Figure 2.2, where the transition rate from state x_i to its neighboring states are denoted by $\lambda_{i,i+1}$ and $\lambda_{i,i-1}$. The nodes placed on the manifold are interpreted as states of the Markov chain. The values a node x_i can take are positions on the infinite straight line, as shown in Figure 2.1. Note that the behavior of this random walk is captured by a continuous-time Markov chain with the infinitesimal generator Q given by (2.1) below. The row of Q in (2.1) involving transition rates $\lambda_{0,1}$ and $\lambda_{0,-1}$ is associated with state transitions from x_0 to states x_1 and x_{-1}, and the column involving transition rates $\lambda_{0,1}$ and $\lambda_{2,1}$ is associated with state transitions to x_1 from states x_0 and x_2. The element associated to the particular row and

$$Q = \begin{bmatrix} \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \\ \ddots & \lambda_{-2,-1} & 0 & 0 & 0 & 0 & \cdots \\ \cdots & -(\lambda_{-1,-2}+\lambda_{-1,0}) & \lambda_{-1,0} & 0 & 0 & 0 & \cdots \\ \cdots & \lambda_{0,-1} & -(\lambda_{0,-1}+\lambda_{0,1}) & \lambda_{0,1} & 0 & 0 & \cdots \\ \cdots & 0 & \lambda_{1,0} & -(\lambda_{1,0}+\lambda_{1,2}) & \lambda_{1,2} & 0 & \cdots \\ \cdots & 0 & 0 & \lambda_{2,1} & -(\lambda_{2,1}+\lambda_{2,3}) & \lambda_{2,3} & \ddots \\ & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix} \tag{2.1}$$

column mentioned above represents therefore the transition rate for going from state x_0 to state x_1.

Note that we can write the time derivative of (1.11) as

$$\frac{\mathrm{d}p_i(t)}{\mathrm{d}t} = v_0^{\mathrm{T}} e^{Qt} Q v_i. \tag{2.2}$$

Furthermore, note that Qv_i is an infinite-dimensional column vector given by

$$Qv_i = [\ldots, 0, \lambda_{i-1,i}, -(\lambda_{i,i-1} + \lambda_{i,i+1}), \lambda_{i+1,i}, 0, \ldots]^{\mathrm{T}}, \tag{2.3}$$

representing the column of Q whose elements are the transition rates to state x_i. Note that the term Qv_i can be expressed as a function of v_i, v_{i-1} and v_{i+1}, i.e., we multiply v_{i+1} with $\lambda_{i+1,i}$, v_{i-1} with $\lambda_{i-1,i}$ and v_i with $-(\lambda_{i,i-1} + \lambda_{i,i+1})$, and sum them. Hence, the expression for Qv_i is then rewritten as

$$Qv_i = \lambda_{i-1,i} v_{i-1} - (\lambda_{i,i-1} + \lambda_{i,i+1}) v_i + \lambda_{i+1,i} v_{i+1}. \tag{2.4}$$

From substituting (2.4) into (2.2), it follows that

$$\begin{aligned} \frac{\mathrm{d}p_i(t)}{\mathrm{d}t} &= v_0^{\mathrm{T}} e^{Qt} (\lambda_{i-1,i} v_{i-1} - (\lambda_{i,i-1} + \lambda_{i,i+1}) v_i + \lambda_{i+1,i} v_{i+1}) \\ &= \lambda_{i-1,i} (v_0^{\mathrm{T}} e^{Qt} v_{i-1}) - (\lambda_{i,i-1} + \lambda_{i,i+1}) (v_0^{\mathrm{T}} e^{Qt} v_i) + \lambda_{i+1,i} (v_0^{\mathrm{T}} e^{Qt} v_{i+1}). \end{aligned} \tag{2.5}$$

The expression $v_0^{\mathrm{T}} e^{Qt} v_{i-1}$ corresponds to (1.10), evaluated for state x_{i-1} instead of x_i. We can therefore rewrite (2.5) as

$$\frac{\mathrm{d}p_i(t)}{\mathrm{d}t} = \lambda_{i-1,i} p_{i-1}(t) - (\lambda_{i,i-1} + \lambda_{i,i+1}) p_i(t) + \lambda_{i+1,i} p_{i+1}(t). \tag{2.6}$$

Note that the expression in (2.6) is also known as the master equation.

2.3 The Heat Equation in One Dimension

In one dimension we write the heat equation as

$$\frac{\partial u}{\partial t} = \zeta \frac{\partial^2 u}{\partial x^2}. \tag{2.7}$$

Note that the Laplace-Beltrami operator reduces to the regular Laplace operator $\frac{\partial^2}{\partial x^2}$ on the infinite straight line. The solution to the heat equation with initial condition $u(x, 0) = \delta(0)$, where δ denotes the Dirac delta function, and no boundary condition is expressed as

$$u(x, t) = \frac{1}{(4\pi\zeta t)^{\frac{1}{2}}} e^{-\frac{x^2}{4\zeta t}}. \tag{2.8}$$

This specific Gaussian function describes the Gaussian distribution with mean 0 and variance $2\zeta t$. It therefore follows that in order to approximate the heat distribution on the infinite straight line, the mean and variance obtained from the distribution of particles should also have mean 0 and variance $2\zeta t$. Note that if we set ζ to $\frac{1}{2}$, (2.8) has the same form as the probability density function of a Brownian particle doing standard Brownian motion in (1.12).

2.4 Mean and Variance of the Continuous-Time Markov Chain

For the diffusion modelling on non-uniform grids we use a random walk driven by a continuous-time Markov chain, and from (1.12) it follows that each continuous-time Markov chain should have a variance of $2\zeta t$ and mean 0 at time t. For simplicity, the diffusivity constant ζ is without loss of generality set to $\frac{1}{2}$.

Theorem 2.1. Let $w(t)$ be a random variable of a continuous-time Markov chain whose state-transition diagram is given by Figure 2.2. Furthermore, let $w(0) = x_0 (= 0)$ and let

$$\lambda_{i,i+1} \triangleq \frac{1}{(x_{i+1} - x_i)(x_{i+1} - x_{i-1})}, \tag{2.9}$$

$$\lambda_{i,i-1} \triangleq \frac{1}{(x_i - x_{i-1})(x_{i+1} - x_{i-1})}. \tag{2.10}$$

Then for any time $t > 0$, the random variable $w(t)$ has a variance t and mean 0.

Proof. The variance of $w(t)$ at time t can be calculated using (1.10) and the well known equation for the variance

$$\begin{aligned}\mathbb{E}[(w(t) - \mathbb{E}[w(t)])^2] &= \mathbb{E}[w^2(t)] - \mathbb{E}[w(t)]^2 \\ &= \sum_{i=-\infty}^{\infty} x_i^2 \mathbb{P}[w(t) = x_i | w(0) = x_0] \\ &\quad - \Big(\sum_{i=-\infty}^{\infty} x_i \mathbb{P}[w(t) = x_i | w(0) = x_0] \Big)^2. \end{aligned} \tag{2.11}$$

First, consider $\mathbb{E}[w(t)]^2$. In order to calculate the mean $\mathbb{E}[w(t)]$ we will take the first time derivative

$$\begin{aligned}\frac{\mathrm{d}\mathbb{E}[w(t)]}{\mathrm{d}t} &= \sum_{i=-\infty}^{\infty} x_i \frac{\mathrm{d}\mathbb{P}[w(t) = x_i | w(0) = x_0]}{\mathrm{d}t} \\ &= \sum_{i=-\infty}^{\infty} x_i \frac{\mathrm{d}p_i(t)}{\mathrm{d}t}, \end{aligned} \tag{2.12}$$

as a start. Using (2.6) we can rewrite the expression (2.12) as

$$\begin{aligned}\frac{\mathrm{d}\mathbb{E}[w(t)]}{\mathrm{d}t} &= \sum_{i=-\infty}^{\infty} x_i \Big(\lambda_{i-1,i} p_{i-1}(t) - (\lambda_{i,i-1} + \lambda_{i,i+1}) p_i(t) + \lambda_{i+1,i} p_{i+1}(t) \Big) \\ &= \sum_{i=-\infty}^{\infty} x_i \lambda_{i-1,i} p_{i-1}(t) - \sum_{i=-\infty}^{\infty} x_i (\lambda_{i,i-1} + \lambda_{i,i+1}) p_i(t) \\ &\quad + \sum_{i=-\infty}^{\infty} x_i \lambda_{i+1,i} p_{i+1}(t). \end{aligned} \tag{2.13}$$

Since we have an infinite sum we can increase the running index i by 1 in the first term of (2.13) and decrease it by 1 in the third term to get

$$\begin{aligned}\frac{\mathrm{d}\mathbb{E}[w(t)]}{\mathrm{d}t} &= \sum_{i=-\infty}^{\infty} x_{i+1} \lambda_{i,i+1} p_i(t) - \sum_{i=-\infty}^{\infty} x_i (\lambda_{i,i-1} + \lambda_{i,i+1}) p_i(t) + \sum_{i=-\infty}^{\infty} x_{i-1} \lambda_{i,i-1} p_i(t) \\ &= \sum_{i=-\infty}^{\infty} p_i(t) \big(x_{i+1} \lambda_{i,i+1} - x_i (\lambda_{i,i-1} + \lambda_{i,i+1}) + x_{i-1} \lambda_{i,i-1} \Big). \end{aligned} \tag{2.14}$$

Now, it follows from (2.9) and (2.10) that

$$x_{i+1}\lambda_{i,i+1} - x_i(\lambda_{i,i-1} + \lambda_{i,i+1}) + x_{i-1}\lambda_{i,i-1} = 0, \tag{2.15}$$

so that (2.14) simplifies to

$$\frac{\mathrm{d}\mathbb{E}[w(t)]}{\mathrm{d}t} = 0. \tag{2.16}$$

After integrating (2.16) with the condition $\mathbb{E}[w(0)] = 0$, $\mathbb{E}[w(t)]$ and therefore $\mathbb{E}[w(t)]^2$ both equal 0 for all $t \geq 0$. Now, consider $\mathbb{E}[w^2(t)]$. Note that the time derivative of $\mathbb{E}[w^2(t)]$ can be found by substituting x_i^2 for x_i in (2.14), which leads to

$$\frac{\mathrm{d}\mathbb{E}[w^2(t)]}{\mathrm{d}t} = \sum_{i=-\infty}^{\infty} p_i(t)\Big(x_{i+1}^2\lambda_{i,i+1} - x_i^2(\lambda_{i,i-1} + \lambda_{i,i+1}) + x_{i-1}^2\lambda_{i,i-1}\Big). \tag{2.17}$$

If we substitute (2.9) and (2.10) into (2.17), the expression in the brackets in (2.17) simplifies to

$$x_{i+1}^2\lambda_{i,i+1} - x_i^2(\lambda_{i,i-1} + \lambda_{i,i+1}) + x_{i-1}^2\lambda_{i,i-1} = 1, \tag{2.18}$$

such that we obtain the time derivative of $\mathbb{E}[w^2(t)]$ as

$$\begin{aligned} \frac{\mathrm{d}\mathbb{E}[w^2(t)]}{\mathrm{d}t} &= \sum_{i=-\infty}^{\infty} p_i(t) \\ &= 1. \end{aligned} \tag{2.19}$$

With the given condition $\mathbb{E}[w(0)] = 0$, we write the integral of (2.19) as

$$\int_0^t \mathrm{d}\tau = t \tag{2.20}$$

Since $\mathbb{E}[w^2(t)] = t$ and $\mathbb{E}[w(t)]^2 = 0$, (2.11) equals t. □

The result in Theorem 2.1 shows that, from any given state x_i, the intensity for going one step forward to state x_{i+1} can be determined by substituting the position of nodes with non-uniform spacing into (2.9). At the same time the theoretical variance t at time t and mean 0 are met in order to allow us to model the diffusion of heat with this framework.

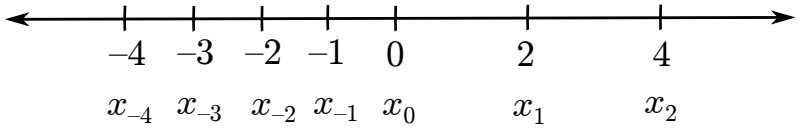

Figure 2.3: Example two of non-uniformly placed nodes on the real axis

Similarly we can determine the transition rate for going one step backward to state x_{i-1} from any given state x_i by substituting the positions of nodes with non-equidistant spacing into (2.10).

Note that if the spacing of nodes is uniform, the transition rates simplify to

$$\lambda_{i,i+1} = \lambda_{i,i-1} = \frac{1}{2(\Delta x)^2}, \quad i \in \mathbb{I}, \tag{2.21}$$

where Δx is the distance between any two neighboring nodes. Using (2.9) and (2.10), we can find all $\lambda_{i,j}$'s in (2.1) by changing the indices i and j. Through the expression of all $\lambda_{i,j}$'s by the position of neighboring nodes, we can completely characterize the continuous-time Markov Chain. Substituting (2.9) and (2.10) into (2.6) gives us the expression for $\frac{\mathrm{d}p_i(t)}{\mathrm{d}t}$

$$\begin{aligned}\frac{\mathrm{d}p_i(t)}{\mathrm{d}t} &= \frac{1}{(x_i - x_{i-1})(x_i - x_{i-2})}p_{i-1}(t) + \frac{x_{i+1} - x_{i-1}}{(x_{i+1} - x_i)(x_i - x_{i-1})(x_{i+1} - x_{i-1})}p_i(t) \\ &\quad + \frac{1}{(x_{i+1} - x_i)(x_{i+2} - x_i)}p_{i+1}(t).\end{aligned} \tag{2.22}$$

An interesting remark is that $\lambda_{i,i+1}$ does not equal $\lambda_{i+1,i}$, which one would expect. If we consider the diffusion of heat particles, the potential difference between $\lambda_{i,i+1}$ and $\lambda_{i+1,i}$ becomes plausible since, due to the non-uniformly placed nodes, a particle at node x_{i+1} may have a far distance to the next node x_{i+2} while at node x_i the distance to the previous node x_{i-1} may be near. Thus, in that case, the transition rates are not equal. It is also interesting to mention that (2.22) has similarities to polynomials for interpolation, e.g., Lagrange interpolation [1], or other commonly used numerical schemes [41], which are usually used to approximate derivatives of functions or PDEs by using unequally spaced points.

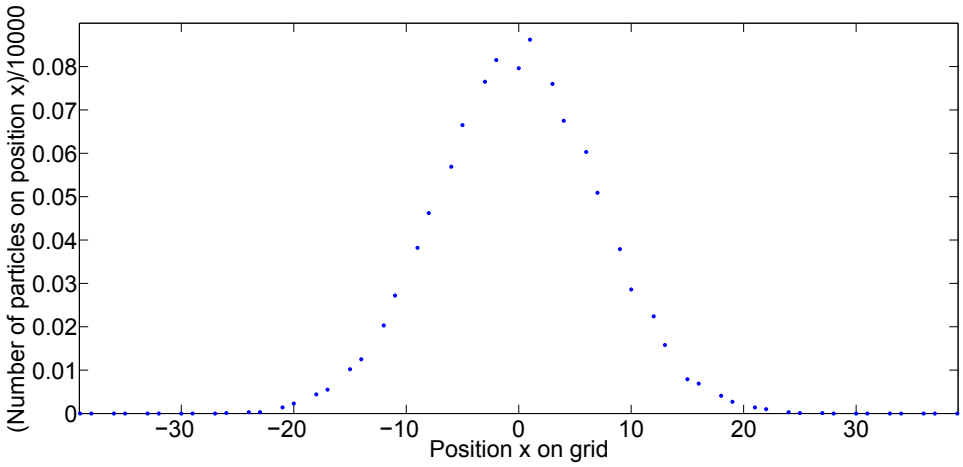

Figure 2.4: Distribution of 10000 Markov Chains at t=50s

2.5 Simulation and Discussion

In this section we present several numerical simulations to demonstrate the utility of the proposed framework. Specifically, consider the Markov-chain-driven random walk on non-uniformly placed nodes given in Figure 2.1. In this case, the infinitesimal generator for the continuous-time Markov chain is characterized by the transition intensities

$$\lambda_{i,i+1} = \begin{cases} \frac{1}{6}, & i = \text{odd}, \\ \frac{1}{3}, & i = \text{even}, \end{cases} \tag{2.23}$$

$$\lambda_{i,i-1} = \begin{cases} \frac{1}{3}, & i = \text{odd}, \\ \frac{1}{6}, & i = \text{even}. \end{cases} \tag{2.24}$$

Starting node for each Markov chain is state x_0, or position 0. Figures 2.4–2.6 show the distribution of the positions of 10000 simulated continuous-time Markov chains, after 50s, 500s and 5000s, respectively. Figures 2.7 and 2.8 show the distributions of the positions of 1000 and 100000 simulated continuous-time Markov chains after 500s, respectively. All figures show the number of particles at each position, divided by the total number of particles employed, or in other words, the percentage of particles at that position after time t, when all particles were concentrated at node $x_0(= 0)$ at time 0.

As expected, a longer simulation time leads to a wider spread or diffusion of parti-

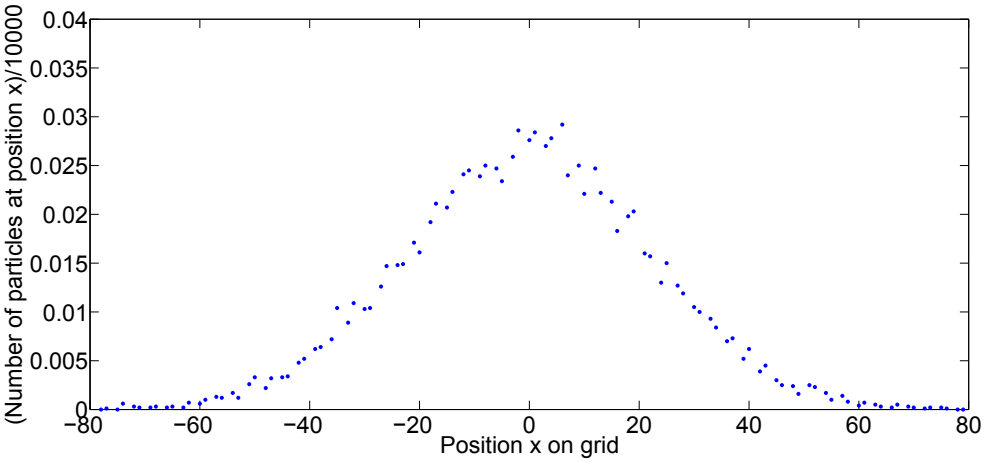

Figure 2.5: Distribution of 10000 Markov Chains at t=500s

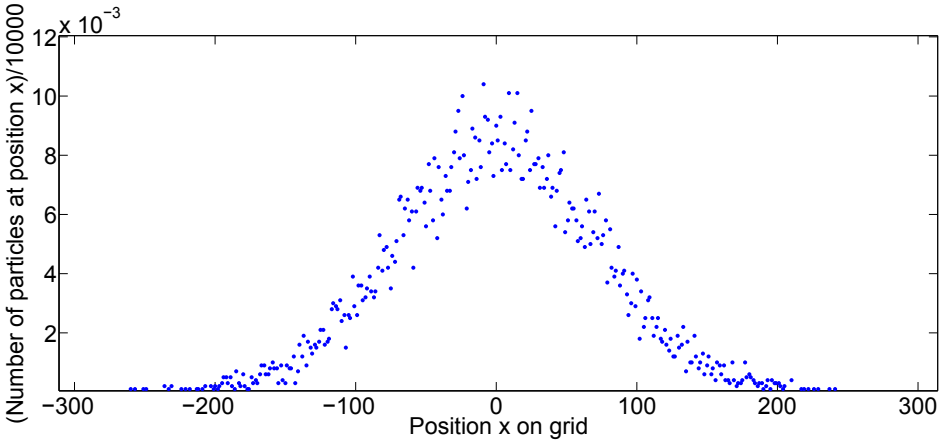

Figure 2.6: Distribution of 10000 Markov Chains at t=5000s

cles. Furthermore, the variance and mean of the distributions of particles have a smaller deviation from t and 0, if more particles are employed.

Let us now have a closer look at the intensities. As mentioned before, the intensities from going, for example, from state x_0 to state x_1 and from state x_1 to state x_0, are possibly different, depending on the distances between the nodes. In the example used in the simulation, both transition rates are $\frac{1}{3}$. But if we imagine nodes placed in the form of $x_{-2} = -4, x_{-1} = -3, x_0 = 0, x_1 = 2, x_2 = 3, x_3 = 6$, we get a transition rate for going from state x_0 to state x_1 of $\frac{1}{10}$. The transition rate for going from state x_1 to state x_0 is $\frac{1}{6}$.

Another example is nodes with a spacing as shown in Figure 2.3. At node x_0, we can see the change of the spacing, i.e., the region with distance 1 between each pair of

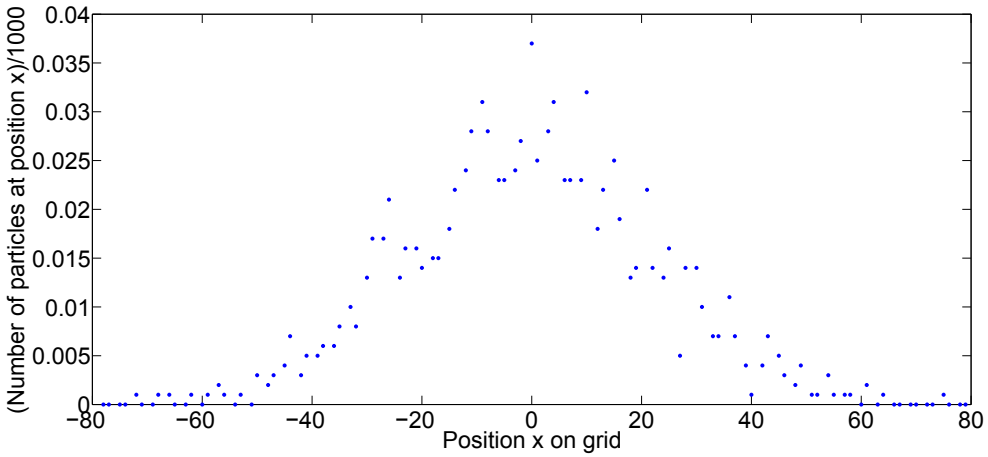

Figure 2.7: Distribution of 1000 Markov Chains at t=500s

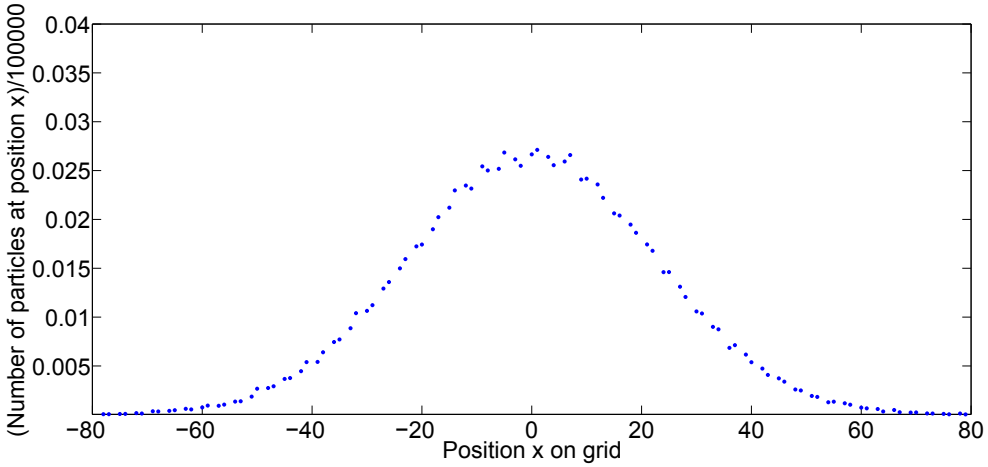

Figure 2.8: Distribution of 100000 Markov Chains at t=500s

neighboring states, to the region with distance 2 between each pair of neighboring states. The transition rates are given by

$$\lambda_{i,i+1} = \begin{cases} \frac{1}{6}, & i = 0, \\ \frac{1}{8}, & i = 1, 2, ..., \\ \frac{1}{2}, & i = -1, -2, ..., \end{cases} \tag{2.25}$$

$$\lambda_{i,i-1} = \begin{cases} \frac{1}{3}, & i = 0, \\ \frac{1}{8}, & i = 1, 2, ..., \\ \frac{1}{2}, & i = -1, -2, ... \end{cases} \tag{2.26}$$

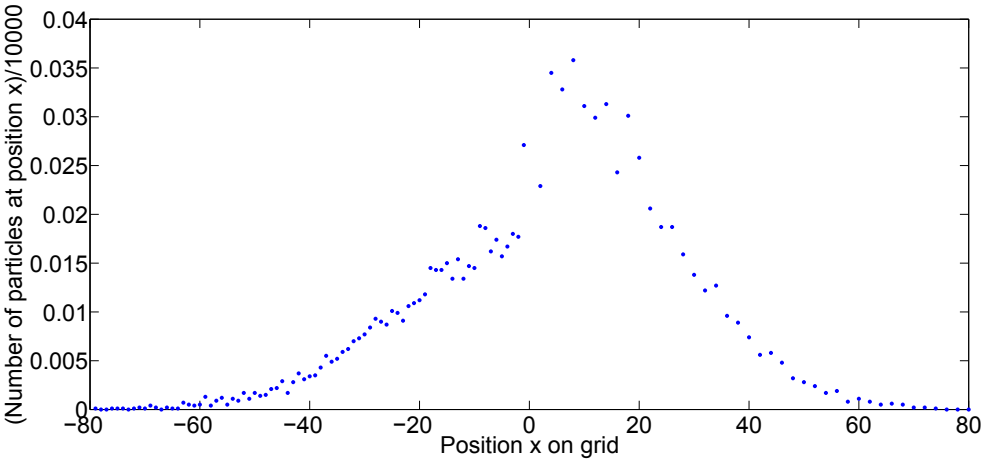

Figure 2.9: Distribution of 10000 Markov Chains at t=500s

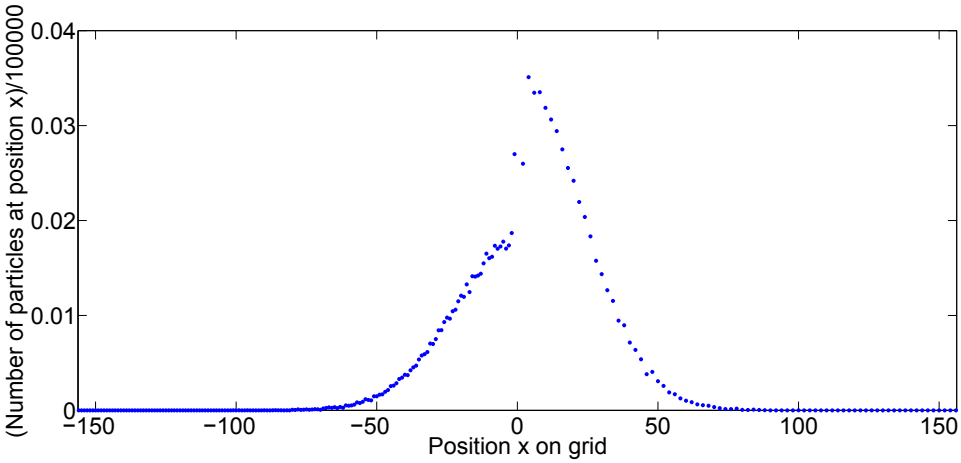

Figure 2.10: Distribution of 100000 Markov Chains at t=500s

Figures 2.9 and 2.10 show the distributions, on nodes with the spacing as explained, for different amounts of heat particles. Similar to our first example, the more continuous-time Markov chains we simulate, the less are the deviations from variance t and mean 0.

For the nodes shown in Figure 2.3, the number of particles on each side of node $x_0(=0)$ has to be equal, since the distribution of particles should approximate the Gaussian distribution. The distance between each pair of neighboring states on the right side of x_0 is twice the distance of each pair of neighboring states on the left side. Due to the doubled distances, each node on the right side of x_0 has to hold more particles so that the number of particles on each side of node $x_0(=0)$ is equal. The plots in Figures 2.9 and 2.10 reflect this idea.

The histogram in Figure 2.11 (resp., Figure 2.12) show the distribution of particles with

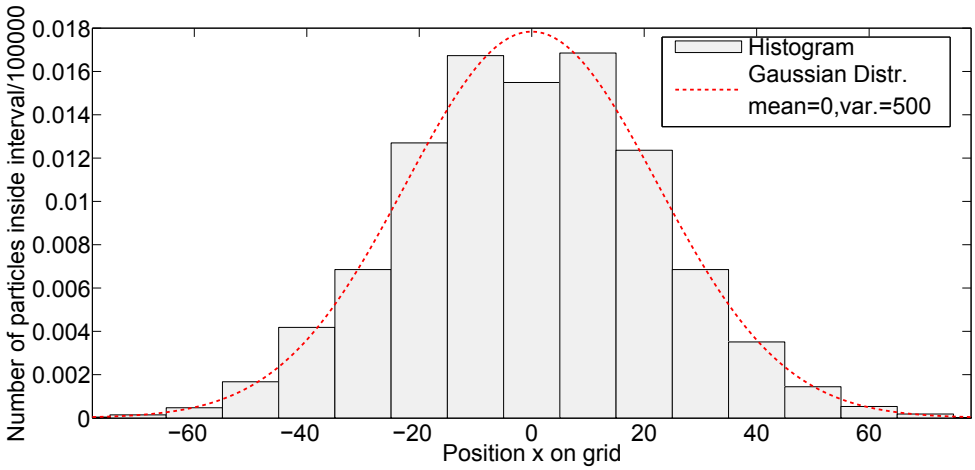

Figure 2.11: Histogram for example one

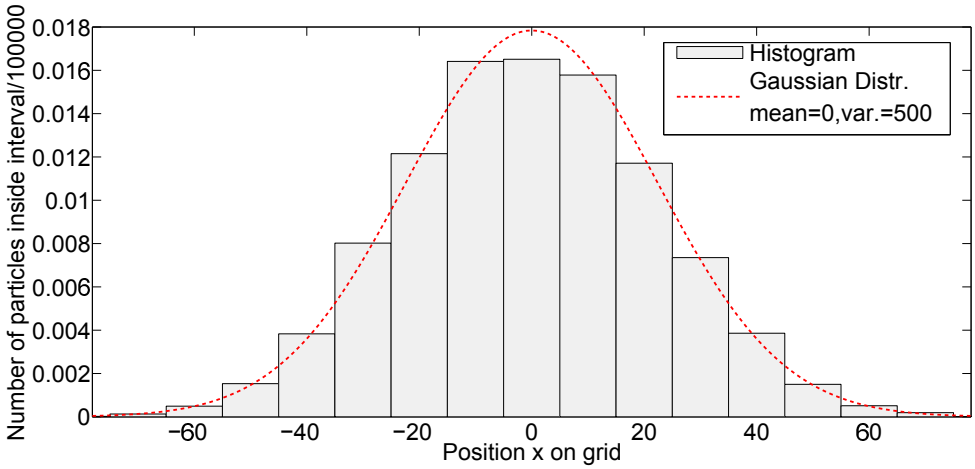

Figure 2.12: Histogram for example two

respect to certain intervals for the example corresponding to Figure 2.1 (resp., Figure 2.3). Their shapes follow the corresponding Gaussian distribution with mean 0 and variance t, and it shows that the number of particles left and right of x_0 are the same for both examples. In this context it is interesting to note that the transition rates are quartered if the distance between nodes double, as in (2.25) and (2.26), indicating less transitions of the continuous-time Markov chain per time unit. This is how the continuous-time Markov chain takes the irregularity into account. However, thinking of heat particles again, it does not represent a slower diffusion, since naturally, it takes longer to go a longer distance than a shorter distance.

2.6 Remarks

This framework can be used not only on the infinite straight line but also on the infinite open curve. In this case we would view the nodes x_i not as positions but distances from the origin. On the infinite straight line the positions correspond the distances but this is generally not the case on the open curve. Another interesting remark is the fact that we can really have any spacing between nodes and still obtain transition rates. In the next chapter we show that, based on the chosen lattice, we cannot always obtain transition rates.

For the one-dimensional case we saw that we have two unknowns, the transition rates to each of the two neighbors, and two constraints, the mean and variance. This leads to unique transition rates if we consider the first and second moment only. An interesting question is why higher moments are not considered? It is quite possible to consider higher moments. However, it more natural to choose the first and the second moments since we only need the mean and the variance to characterize the Gaussian distribution. The third and fourth moments of a random variable, which is distributed according to a Gaussian distribution with mean 0 and variance t, are given as 0 and $3t^2$, respectively. For the second example in this chapter, at $t = 50$ we obtain a third moment of 0.0957 and a fourth moment of 7880, which is quite close to the theoretical third moment of 0 and the theoretical fourth moment of 7500.

As a last remark, it may be interesting to determine the number particles n utilized in the particle simulation such that a certain accuracy of the mean and variance is guaranteed. To determine the necessary number of particles we consider confidence intervals [42]. For instance, we define a 95% confidence that the simulated mean of particle positions is within (L, U) of the theoretical mean as

$$(L, U)_{95\%} = \overline{x} \pm z_{\mathrm{c}} \frac{S_{\mathrm{x}}}{\sqrt{n}}, \tag{2.27}$$

where L is the lower bound, U is the upper bound, $\overline{x}$ is the sample mean, S_{x} is the sample standard deviation and z_{c} are the confidence coefficients. The confidence coefficients are for a Gaussian distribution and vary with confidence levels. Now, by considering the

confidence interval to represent twice the maximum error we can define

$$e_{\max} = \frac{z_{\mathrm{c}} S_{\mathrm{x}}}{\sqrt{n}}. \tag{2.28}$$

The variable $e_{\max}$ denotes the probability that the sample mean is not within the boundaries L and U. The least amount of particles necessary for a given confidence level is then

$$n = \left(\frac{z_{\mathrm{c}} S_{\mathrm{x}}}{e_{\max}}\right)^2. \tag{2.29}$$

As an example, if we wanted a 95% confidence that the sample mean is within ± 1.96 of the true mean at time 5, we would need to employ 7683 particles.

2.7 Comparison to Finite Differences Method

The framework presented in this thesis is a new way to approximate the heat distribution on a given manifold. For this reason it is interesting to give a basic comparison to existing deterministic methods regarding run time and accuracy. The deterministic method we use for the comparison is the finite differences method, and the example grid is the same as in Figure 2.1. Specifically, we use the space discretization

$$\frac{\partial^2 u}{\partial x^2}(x_i, t_n) \approx \frac{(x_i - x_{i-1})u_{i-1,n} - (x_{i+1} - x_{x-1})u_{i,n} + (x_{i+1} - x_i)u_{i-1,n}}{0.5(x_{i+1} - x_i)(x_i - x_{i-1})(x_{i+1} - x_{i-1})}. \tag{2.30}$$

For the time discretization we use

$$\frac{\partial u}{\partial t}(x_i, t_{n+\frac{1}{2}}) \approx \frac{u_{i,n+1} - u_{i,n}}{\Delta t}, \tag{2.31}$$

the forward Euler time integration, i.e., we only use the temperature at time step n to calculate the temperatures at the next time step $n+1$. The spatial derivative is evaluated at time n and the time derivative is evaluated at time $n+\frac{1}{2}$. We can then substitute (2.30) and (2.31) into the heat equation and solve for $u_{i,n+1}$. If we set the thermal diffusivity ζ

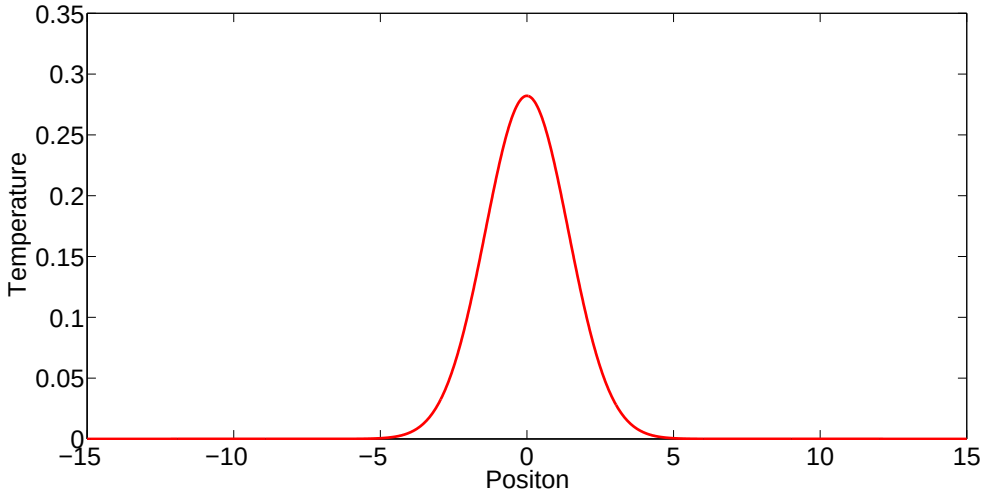

Figure 2.13: Analytical heat distribution at $t = 2$

to 1 we get

$$u_{i,n+1} \approx u_{i,n} + \Delta t \frac{(x_i - x_{i-1})u_{i-1,n} - (x_{i+1} - x_{x-1})u_{i,n} + (x_{i+1} - x_i)u_{i-1,n}}{0.5(x_{i+1} - x_i)(x_i - x_{i-1})(x_{i+1} - x_{i-1})}. \quad (2.32)$$

The difference when comparing our framework with the finite differences method is the way we display results. We consider histograms of particle positions and then put a curve on top of the histogram while the finite differences considers the temperature at specific points only. Figure 2.13 shows the analytical distribution of heat on example grid one at $t = 2$ with an open boundary and the initial condition $u(x, 0) = \delta(x)$. Figure 2.14 shows the finite differences solution with the same boundary and initial condition and Figure 2.15 shows the histogram of particle positions where all $10,000$ particles are at position $x = 0$ at time 0.

When considering the computation time of our framework and the finite differences method it is important to note that the run time heavily depends on which of the many implementations of the code we use for either method. Particularly for the finite differences method a lot of very fast implementations are available. Furthermore it depends on what part of the, theoretically, infinite grid we consider for the finite differences method. A bigger grid results in a longer computation time for the finite differences method, while for our stochastic framework it does not matter. The time step size for the finite differences method, which for this particular case should be less or equal than 0.5 to avoid

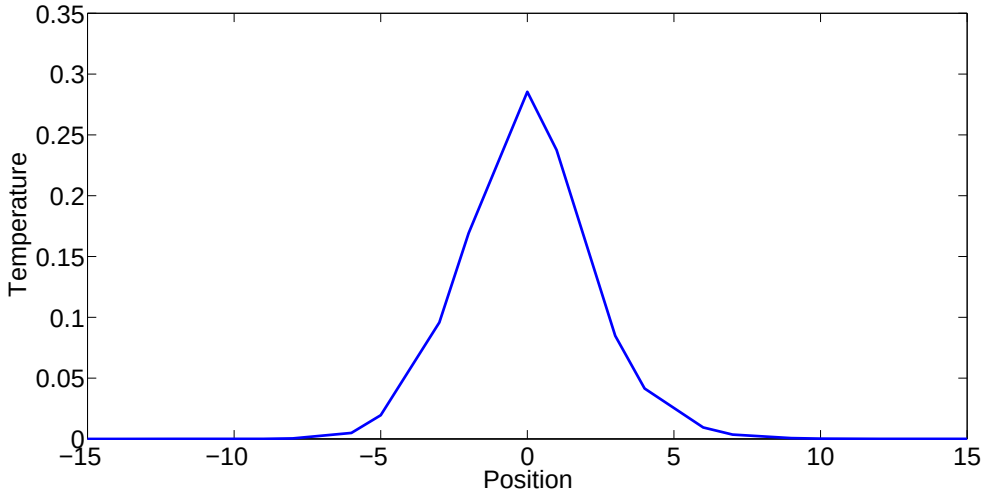

Figure 2.14: Finite differences approximation of the heat distribution on example grid one at $t = 2$

instability, is also an important factor. In the stochastic simulation the time step is drawn from a exponential distribution with grid dependent rate parameter. Table 2.1 shows the computation times for the finite differences method for different grid sizes and time step sizes. The computation time for the stochastic simulation of $10,000$ particles is $0.46s$, for 1000 particles it takes $0.041s$. However, we used a machine with only one core for the stochastic particle simulation. If we were able to use a machine with, for instance, 10 cores, we could use a parallel computing implementation of our simulation code. The computation time difference when employing $10,000$ and 1000 particles is significant and on par with the computation time for the finite differences method. Of course we cannot simply divide the computation time for $10,000$ particles by 10 to obtain the computation time for the parallel computing implementation. However, the computation time is still in that region.

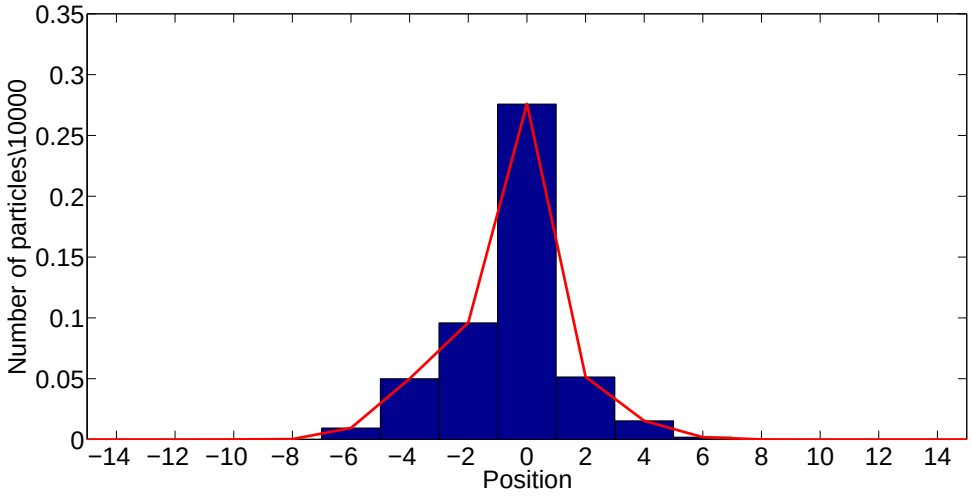

Figure 2.15: Histogram of particle positions on example grid one at $t = 2$

Table 2.1: Computation time for FDM for different time steps Δt and different intervals L

Interval L	**Δ t=0.5**	**Δ t=0.25**
50	0.02s	0.064s
100	0.058s	0.084s
200	0.085s	0.15s
500	0.119s	0.164s

Chapter 3

Heat Diffusion Modelling on Two-Dimensional Lattices

3.1 Introduction

In this chapter we deal with the extension of the framework we developed in Chapter 2. In two dimensions we describe the position of nodes and particles with two-dimensional vectors. The extension of our framework to two dimensions is unfortunately not an easy task. The main change is the decision on what type of lattice should be used, and indeed we are confronted with limitations based on the lattices we choose. In view of possible later extensions of our framework, we mainly consider a triangular lattice. This type of lattice consists of nodes where each node has six neighbors. When considering surfaces on curved spaces, it is more natural to use triangular lattices on those surfaces than, say, square lattices. This becomes evident when we consider the sphere. We can not create a lattice on the sphere which consists of squares, without stretching some edges more than others. With triangular lattices we are able to stretch some edges more than others. In this chapter we will first slightly change the definition of the continuous-time Markov chain and then give a proof that the random variable of the Markov chain has the same mean and covariance matrix as the Gaussian distribution in two dimensions, where the Gaussian distribution satisfies the two-dimensional heat equation.

3.2 Markov-Chain-Driven Random Walk on the Triangular Lattice

For the two-dimensional case we modify the previous definition of the continuous-time Markov chain in order to clearly reflect the fact that we deal with two-dimensional lattices in the notation. Let $w(t)$ be the random variable of a continuous-time, time homogeneous Markov chain on a countable space X. The variables $x_{i,j}$, $i, j = \ldots, -2, -1, 0, 1, 2, \ldots$, of dimension 2×1, denote generic states of X. We express the conditional probability of the time-homogeneous, continuous-time Markov chain as

$$\mathbb{P}[w(s+t) = x_{i,j} | w(s) = x_{k,l}] = \mathbb{P}[w(t) = x_{i,j} | w(0) = x_{k,l}], \quad i, j, k, l \in \mathbb{I}. \tag{3.1}$$

We also define

$$p_{(k,l),(i,j)}(t) \triangleq \mathbb{P}[w(t) = x_{i,j} | w(0) = x_{k,l}]. \tag{3.2}$$

Associated with the infinite-dimensional matrix $P(t)$, where the pair (k, l) is an index for a row of $P(t)$ and the pair (i, j) is an index for a column of $P(t)$, given by $p_{(k,l),(i,j)}(t)$ in (3.2), the infinitesimal generator Q for the Markov chain (3.1) is defined as

$$q_{(k,l),(i,j)} = \begin{cases} \lim_{t \to 0} \frac{p_{(k,l),(i,j)}(t)}{t}, & (k,l) \neq (i,j), \\ \lim_{t \to 0} \frac{p_{(k,l),(k,l)}(t) - 1}{t}, & (k,l) = (i,j). \end{cases} \tag{3.3}$$

Note that $q_{(k,l),(i,j)} \geq 0, (k,l) \neq (i,j)$, and $q_{(k,l),(k,l)} \leq 0$. The elements $q_{(k,l),(i,j)}$ of Q are called transition rates or intensities and indicate the rate at which $w(t)$ transitions from state $x_{k,l}$ to state $x_{i,j}$. Furthermore, Q satisfies the Kolmogorov forward equation given by

$$\frac{\mathrm{d}P(t)}{\mathrm{d}t} = P(t)Q, \quad P(0) = I, \tag{3.4}$$

with the solution

$$P(t) = e^{Qt}, \tag{3.5}$$

where I indicates the infinite-dimensional identity matrix.

If we want to look at the probability that $w(t)$ is in state $x_{i,j}$ at time t, given that

$w(t)$ was in state $x_{0,0}$ in time 0, we look at the corresponding row and column of e^{Qt}. Assuming that the initial state is $x_{0,0}$, we write the probability that the continuous-time Markov chain is in state $x_{i,j}$ at time t as

$$\begin{aligned} p_{i,j}(t) \triangleq p_{(0,0),(i,j)}(t) &= \mathbb{P}[w(t) = x_{i,j} | w(0) = x_{0,0}] \\ &= (e^{Qt})_{(0,0),(i,j)}, \end{aligned} \tag{3.7}$$

which can be calculated as

$$p_{i,j}(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} (Q^k)_{(0,0),(i,j)}. \tag{3.8}$$

Similarly to the case on the infinite straight line, we then consider a Markov-chain-driven random walk on a triangular lattice where all the particles are initially placed at node $x_{0,0}$ in the two-dimensional space. This type of lattice consists of diagonal and horizontal lines, where all horizontal edges have length a and all diagonal edges have length b. We then place nodes $x_{i,j}$, $i, j \in \mathbb{I}$, on the triangular lattice as shown in Figure 3.1. Changing the index i is equivalent to jumping along the horizontal lines. Changing the index j is equivalent to jumping along the ascending diagonal lines. Furthermore, each node has exactly six edges. In this case, if we let the particles randomly jump between the neighboring six lattice nodes, the state transition diagram describing the stochastic process of the successive random jumps to neighboring nodes is described in Figure 3.2, where the transition rates from state $x_{i,j}$ to its neighboring states are denoted by $\lambda_{(i,j),(i+1,j)}$, $\lambda_{(i,j),(i-1,j)}$, $\lambda_{(i,j),(i,j+1)}$, $\lambda_{(i,j),(i,j-1)}$, $\lambda_{(i,j),(i+1,j-1)}$, and $\lambda_{(i,j),(i-1,j+1)}$. Note that the behavior of this random walk is captured by a continuous-time Markov chain with the infinitesimal generator Q given by (3.6), with $\beta_{k,l} = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \lambda_{(k,l),(i,j)}$. The row with transition rates $\lambda_{(0,0),(i,j)}$

$$Q = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \cdots & -\beta_{-1,1} & \lambda_{(-1,1),(-1,0)} & 0 & \lambda_{(-1,1),(0,0)} & \lambda_{(-1,1),(0,1)} & 0 & \cdots & \\ \cdots & \lambda_{(-1,0),(-1,1)} & -\beta_{-1,0} & \lambda_{(-1,0),(0,-1)} & \lambda_{(-1,0),(0,0)} & 0 & 0 & \cdots & \\ \cdots & 0 & \lambda_{(0,-1),(-1,0)} & -\beta_{0,-1} & \lambda_{(0,-1),(0,0)} & 0 & 0 & \cdots & \\ \cdots & \lambda_{(0,0),(-1,1)} & \lambda_{(0,0),(-1,0)} & \lambda_{(0,0),(0,-1)} & -\beta_{0,0} & \lambda_{(0,0),(0,1)} & \lambda_{(0,0),(1,0)} & \cdots & \\ \cdots & \lambda_{(0,1),(-1,1)} & 0 & 0 & \lambda_{(0,1),(0,0)} & -\beta_{0,1} & \lambda_{(0,1),(1,0)} & \cdots & \\ \cdots & 0 & 0 & 0 & \lambda_{(1,0),(0,0)} & \lambda_{(1,0),(0,1)} & -\beta_{1,0} & \cdots & \\ & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix} \tag{3.6}$$

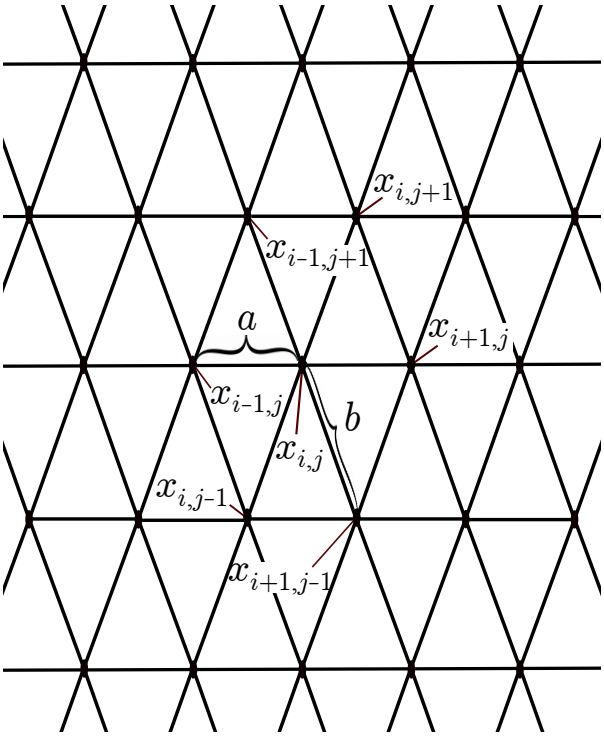

Figure 3.1: Example of a triangular lattice on the infinite plane, where an increase of index i by 1 corresponds to a jump along the horizontal edges to the right and a decrease of index i by 1 corresponds to a jump along the horizontal edges to the left. An increase of index j by 1 corresponds to a jump along the ascending edges and a decrease of index j by 1 corresponds to a jump along the descending edges

of (3.6) is associated with the transition rates from state $x_{0,0}$ to all other states $x_{i,j}$ and the column with transition rates $\lambda_{(k,l),(0,-1)}$ is associated with the transition rates to state $x_{0,-1}$ from all other states $x_{k,l}$. The element in the above mentioned row and column represents therefore the transition rate for going from state $x_{0,0}$ to state $x_{0,-1}$. The order of how we index the rows and columns in (3.6) is not of importance for the purpose of this research.

Note that due to the symmetric structure of this type of lattice we have only two different transition rates. Even though completely irregular lattices with different transition rates and lengths for each edge are certainly desirable, the class of lattices discussed here also leads to some interesting results. Furthermore, note that the time derivative of (3.7)

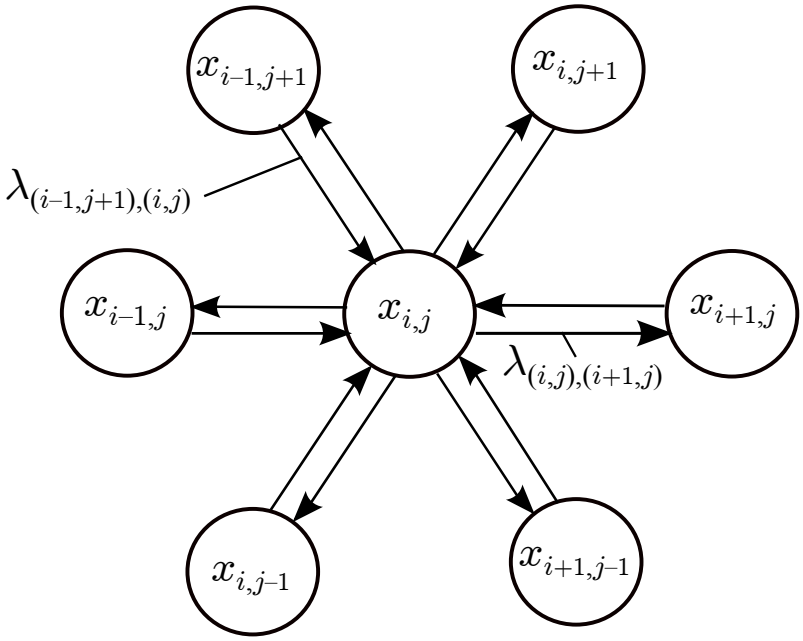

Figure 3.2: State-transition diagram of the continuous-time Markov chain on the example triangular lattice

is given by

$$\frac{\mathrm{d}p_{i,j}(t)}{\mathrm{d}t} = (e^{Qt}Q)_{(0,0),(i,j)}. \tag{3.9}$$

From any given node particles can jump only to one of the six neighboring nodes. The transition rates to the other nodes are therefore 0. Due to this structure of Q and the Kolmogorov forward equation, we can rewrite (3.9) as a combination of transition rates and transition probabilities to each neighboring node given by

$$\begin{aligned}\frac{\mathrm{d}p_{i,j}(t)}{\mathrm{d}t} &= \lambda_{(i+1,j),(i,j)}p_{i+1,j}(t) + \lambda_{(i,j+1),(i,j)}p_{i,j+1}(t) + \lambda_{(i+1,j-1),(i,j)}p_{i+1,j-1}(t) \\ &\quad -(\lambda_{(i,j),(i+1,j)} + \lambda_{(i,j),(i,j+1)} + \lambda_{(i,j),(i+1,j-1)} + \lambda_{(i,j),(i-1,j)} \\ &\quad +\lambda_{(i,j),(i-1,j+1)} + \lambda_{(i,j),(i,j-1)})p_{i,j}(t) + \lambda_{(i-1,j),(i,j)}p_{i-1,j}(t) \\ &\quad +\lambda_{(i-1,j+1),(i,j)}p_{i-1,j+1}(t) + \lambda_{(i,j-1),(i,j)}p_{i,j-1}(t).\end{aligned} \tag{3.10}$$

3.3 The Heat Equation on the Plane

In two dimensions the heat equation is expressed by

$$\frac{\partial u}{\partial t} = \zeta\Big(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}\Big)u, \tag{3.11}$$

where x is a two-dimensional vector with elements x_1 and x_2, indicating the position in $\mathbb{R}^2$. This time the Laplace-Beltrami operator becomes the regular Laplace operator in two dimensions. The initial condition is $u(x,0) = \delta(x)$, where in this chapter $\delta(x)$ represents the Dirac delta function in two dimensions. When we assume no boundary conditions, we express the solution to (3.11) as

$$u(x,t) = \frac{1}{2\pi t(\det D)^{\frac{1}{2}}} e^{-\frac{1}{2t}x^{\mathrm{T}}D^{-1}x}, \tag{3.12}$$

where D is a matrix whose elements represent the thermal diffusivity. Similar to the 1D-case, (3.12) is also a Gaussian function. Specifically, it is a Gaussian distribution of the form

$$p(x) = \frac{1}{2\pi(\det\Sigma)^{\frac{1}{2}}} e^{-\frac{1}{2}x^{\mathrm{T}}\Sigma^{-1}x}, \tag{3.13}$$

where Σ is the covariance matrix of the Gaussian distribution in two dimensions. Hence, (3.12) is a Gaussian distribution with $\Sigma = tD$ and $\mu = 0$. The theoretical mean and covariance matrix of the Gaussian distribution is then taken as the requirement for the mean and covariance matrix of the random variable of the Markov chain, i.e., mean 0 and covariance matrix tD. Note that if we set D to $\frac{1}{2}I_2$, (3.12) has the same form as the probability density function for a Brownian particle doing standard Brownian motion on the plane.

3.4 Mean and Covariance of the Continuous-Time Markov Chain

As in Chapter 2 we now proceed with the proof that the random variable of the continuous-time Markov chain has the theoretical mean and covariance matrix of a Gaussian distribution in two dimensions. Without loss of generality we set the thermal diffusivity to $\frac{1}{2}I_2$.

Theorem 3.1. Let $w(t)$ be a random variable of a continuous-time Markov chain whose state-transition diagram is given by Figure 3.2. Furthermore, let $w(0) = x_{0,0}(= (0,0))$ and

let

$$\begin{aligned}\lambda_{(i,j),(i+1,j)} &= \lambda_{(i,j),(i-1,j)} \triangleq \lambda^{(1)},\\ \lambda_{(i,j),(i,j-1)} &= \lambda_{(i,j),(i-1,j+1)} = \lambda_{(i,j),(i+1,j-1)}\\ &= \lambda_{(i,j),(i,j+1)} \triangleq \lambda^{(2)},\end{aligned} \tag{3.14}$$

where

$$\lambda^{(1)} \triangleq \frac{2b^2 - a^2}{2a^2(4b^2 - a^2)}, \tag{3.15}$$

$$\lambda^{(2)} \triangleq \frac{1}{2(4b^2 - a^2)}. \tag{3.16}$$

If $\sqrt{2}b > a$, then the random variable $w(t)$ has a covariance matrix $\frac{1}{2}tI_2$ and mean 0 for all time $t > 0$.

Proof. The mean of $w(t)$ can be computed as

$$\mathbb{E}[w(t)] = \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{i,j}\mathbb{P}[w(t) = x_{i,j}|w(0) = x_{0,0}]. \tag{3.17}$$

In order to compute $\mathbb{E}[w(t)]$, we take the time derivative of $\mathbb{E}[w(t)]$ so that

$$\begin{aligned}\frac{\mathrm{d}\mathbb{E}[w(t)]}{\mathrm{d}t} &= \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{i,j}\frac{\mathrm{d}\mathbb{P}[w(t) = x_{i,j}|w(0) = x_{0,0}]}{\mathrm{d}t}\\ &= \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{i,j}\frac{\mathrm{d}p_{i,j}(t)}{\mathrm{d}t}.\end{aligned} \tag{3.18}$$

Using (3.10) we can rewrite expression (3.18) as

$$\begin{aligned}\frac{\mathrm{d}\mathbb{E}[w(t)]}{\mathrm{d}t} &= \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{i,j}\Big(\lambda_{(i+1,j),(i,j)}p_{i+1,j}(t) + \lambda_{(i,j+1),(i,j)}p_{i,j+1}(t)\\ &\quad +\lambda_{(i+1,j-1),(i,j)}p_{i+1,j-1}(t) - (\lambda_{(i,j),(i+1,j)} + \lambda_{(i,j),(i,j+1)} + \lambda_{(i,j),(i+1,j-1)}\\ &\quad +\lambda_{(i,j),(i-1,j)} + \lambda_{(i,j),(i-1,j+1)} + \lambda_{(i,j),(i,j-1)})p_i(t) + \lambda_{(i-1,j),(i,j)}p_{i-1,j}(t)\\ &\quad +\lambda_{(i-1,j+1),(i,j)}p_{i-1,j+1}(t) + \lambda_{(i,j-1),(i,j)}p_{i,j-1}(t)\Big)\\ &= \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{i,j}\lambda_{(i+1,j),(i,j)}p_{i+1,j}(t) + \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{i,j}\lambda_{(i,j+1),(i,j)}p_{i,j+1}(t)\end{aligned}$$

$$
\begin{aligned}
&+\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{i,j}\lambda_{(i+1,j-1),(i,j)}p_{i+1,j-1}(t) - \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{i,j}(\lambda_{(i,j),(i+1,j)}\\
&+\lambda_{(i,j),(i,j+1)} + \lambda_{(i,j),(i+1,j-1)} + \lambda_{(i,j),(i-1,j)} + \lambda_{(i,j),(i-1,j+1)}\\
&+\lambda_{(i,j),(i,j-1)})p_i(t) + \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{i,j}\lambda_{(i-1,j),(i,j)}p_{i-1,j}(t)\\
&+\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{i,j}\lambda_{(i-1,j+1),(i,j)}p_{i-1,j+1}(t)\\
&+\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{i,j}\lambda_{(i,j-1),(i,j)}p_{i,j-1}(t). \qquad (3.19)
\end{aligned}
$$

Since we have an infinite sum, we can decrease the running index i by 1 in the first term of (3.19), decrease j by 1 in the second term, decrease i by 1 and increase j by 1 in the fourth term, increase i in the fifth term, increase i and decrease j in the sixth term and decrease j in the seventh term to get

$$
\begin{aligned}
\frac{\mathrm{d}\mathbb{E}[w(t)]}{\mathrm{d}t} &= \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{i-1,j}\lambda_{(i,j),(i-1,j)}p_{i,j}(t) + \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{i,j-1}\lambda_{(i,j),(i,j-1)}p_{i,j}(t)\\
&+\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{i-1,j+1}\lambda_{(i,j),(i-1,j+1)}p_{i,j}(t) - \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{i,j}(\lambda_{(i,j),(i+1,j)}\\
&+\lambda_{(i,j),(i,j+1)} + \lambda_{(i,j),(i+1,j-1)} + \lambda_{(i,j),(i-1,j)} + \lambda_{(i,j),(i-1,j+1)}\\
&+\lambda_{(i,j),(i,j-1)})p_i(t)\\
&+\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{i+1,j}\lambda_{(i,j),(i+1,j)}p_{i,j}(t)\\
&+\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{i+1,j-1}\lambda_{(i,j),(i+1,j-1)}p_{i,j}(t)\\
&+\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{i,j+1}\lambda_{(i,j),(i,j+1)}p_{i,j}(t)\\
&= \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} p_{i,j}(t)\Big(x_{i-1,j}\lambda_{(i,j),(i-1,j)} + x_{i,j-1}\lambda_{(i,j),(i,j-1)}\\
&+x_{i-1,j+1}\lambda_{(i,j),(i-1,j+1)} - (\lambda_{(i,j),(i+1,j)} + \lambda_{(i,j),(i,j+1)} + \lambda_{(i,j),(i+1,j-1)}\\
&+\lambda_{(i,j),(i-1,j)} + \lambda_{(i,j),(i-1,j+1)} + \lambda_{(i,j),(i,j-1)})x_{i,j} + x_{i+1,j}\lambda_{(i,j),(i+1,j)}\\
&+x_{i+1,j-1}\lambda_{(i,j),(i+1,j-1)} + x_{i,j+1}\lambda_{(i,j),(i,j+1)}\Big). \qquad (3.20)
\end{aligned}
$$

Substituting (3.14) and inserting the actual coordinates from Figure 3.1 into (3.20) gives

us

$$\begin{aligned}
\frac{\mathrm{d}\mathbb{E}[w(t)]}{\mathrm{d}t} &= \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} p_{i,j}(t)\Big(\Big(\begin{bmatrix}-a\\0\end{bmatrix}+\begin{bmatrix}a\\0\end{bmatrix}\Big)\lambda^{(1)} \\
&\quad +\Big(\begin{bmatrix}-\frac{a}{2}\\ \sqrt{b^2-\frac{a^2}{4}}\end{bmatrix}+\begin{bmatrix}-\frac{a}{2}\\ -\sqrt{b^2-\frac{a^2}{4}}\end{bmatrix} \\
&\quad +\begin{bmatrix}\frac{a}{2}\\ \sqrt{b^2-\frac{a^2}{4}}\end{bmatrix}+\begin{bmatrix}\frac{a}{2}\\ -\sqrt{b^2-\frac{a^2}{4}}\end{bmatrix}\Big)\lambda^{(2)}\Big) \\
&= \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}\begin{bmatrix}0\\0\end{bmatrix} p_{i,j}(t) \\
&= 0. \qquad (3.21)
\end{aligned}$$

Integrating (3.21) leads to $\mathbb{E}[w(t)] \equiv 0$. Hence, since the covariance is calculated with $\mathbb{E}[w(t)w^{\mathrm{T}}(t)] - \mathbb{E}[w(t)]\mathbb{E}[w^{\mathrm{T}}(t)]$, we can simplify it to $\mathbb{E}[w(t)w^{\mathrm{T}}(t)]$, [43].

As in the calculation for the mean, we compute the time derivative of $\mathbb{E}[w(t)w^{\mathrm{T}}(t)]$ first, to find $\mathbb{E}[w(t)w^{\mathrm{T}}(t)]$. An expression for the time derivative of $\mathbb{E}[w(t)w^{\mathrm{T}}(t)]$ can be easily obtained by substituting $x_{i,j}x_{i,j}^{\mathrm{T}}$ for $x_{i,j}$ in (3.20)

$$\begin{aligned}
\frac{\mathrm{d}\mathbb{E}[w(t)w^{\mathrm{T}}(t)]}{\mathrm{d}t} &= \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} p_{i,j}(t)\Big(x_{i-1,j}x_{i-1,j}^{\mathrm{T}}\lambda_{(i,j),(i-1,j)} + x_{i,j-1}x_{i,j-1}^{\mathrm{T}}\lambda_{(i,j),(i,j-1)} \\
&\quad +x_{i-1,j+1}x_{i-1,j+1}^{\mathrm{T}}\lambda_{(i,j),(i-1,j+1)} \\
&\quad -(\lambda_{(i,j),(i+1,j)} + \lambda_{(i,j),(i,j+1)} + \lambda_{(i,j),(i+1,j-1)} + \lambda_{(i,j),(i-1,j)} \\
&\quad +\lambda_{(i,j),(i-1,j+1)} + \lambda_{(i,j),(i,j-1)})x_{i,j}x_{i,j}^{\mathrm{T}} + x_{i+1,j}x_{i+1,j}^{\mathrm{T}}\lambda_{(i,j),(i+1,j)} \\
&\quad +x_{i+1,j-1}x_{i+1,j-1}^{\mathrm{T}}\lambda_{(i,j),(i+1,j-1)} + x_{i,j+1}x_{i,j+1}^{\mathrm{T}}\lambda_{(i,j),(i,j+1)}\Big). \qquad (3.22)
\end{aligned}$$

Inserting (3.14) and the coordinates as we did in (3.21) gives us

$$\begin{aligned}
\frac{\mathrm{d}\mathbb{E}[w(t)w^{\mathrm{T}}(t)]}{\mathrm{d}t} &= \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} p_{i,j}(t)\Big(\begin{bmatrix}2a^2 & 0\\ 0 & 0\end{bmatrix}\lambda^{(1)} \\
&\quad +\begin{bmatrix}a^2 & 0\\ 0 & 4(b^2-\frac{a^2}{4})\end{bmatrix}\lambda^{(2)}\Big). \qquad (3.23)
\end{aligned}$$

We can then insert (3.15) and (3.16) to simplify (3.23) to

$$\frac{\mathrm{d}\mathbb{E}[w(t)w^{\mathrm{T}}(t)]}{\mathrm{d}t} = \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} p_{i,j}(t)\begin{bmatrix}\frac{1}{2} & 0\\ 0 & \frac{1}{2}\end{bmatrix}. \tag{3.24}$$

Since $\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} p_{i,j}(t) = 1$, we integrate (3.24) and get

$$\mathbb{E}[w(t)w^{\mathrm{T}}(t)] = \begin{bmatrix}\frac{1}{2} & 0\\ 0 & \frac{1}{2}\end{bmatrix} t. \tag{3.25}$$

□

Related to Figure 3.1, we can stretch or shrink the lattice along the vertical axis, i.e., increase or decrease the ratio $\frac{b}{a}$, and still obtain transition rates, as long as $\sqrt{2}b > a$. Illustrated on the example lattice in Figure 3.1, this conditions has the meaning of a minimum angle of 45 degrees between ascending edges and horizontal edges. If this condition is not satisfied we cannot compute the transition rates for such a lattice, or in other words, we cannot model the Markov chain as a Brownian particle.

Related to the physical phenomenon of heat diffusion, in order to model the symmetric diffusion of heat we have to have particle moving into all directions. If the above condition $\sqrt{2}b > a$ is not satisfied, particles can not move "up". The horizontal part of their movement is too big compared to the vertical movement.

3.5 Simulation and Discussion

To illustrate the utility of the proposed framework, we consider the lattice shown in Figure 3.1 with $b = 1.5$ and $a = 1$. From Theorem 3.1 it follows that the infinitesimal generator of the continuous-time Markov chain is characterized by the transition rates $\lambda^{(1)} = \frac{7}{32}$ and $\lambda^{(2)} = \frac{1}{16}$. The initial state in each simulation is state $x_{0,0}$ or position $(0, 0)$.

Figures 3.3–3.10 show the histograms of the positions of particles, at different times and different amounts of employed particles as well as multivariate Gaussian distributions with mean 0 and covariance $\frac{1}{2}tI_2$. Evidently, if we look at the diffusion process at a later time, particles have diffused farther, leading to a wider distribution of particles. If we lay

Table 3.1: Mean and covariance after 500s for n Markov chains

	n=1000	n=10,000	n=100,000
Mean	[0.232 0.490]	[0.085 -0.064]	[0.001 -0.022]
Covariance	diag[250.83,275.21]	diag[249.94,256.14]	diag[249.84,248.81]

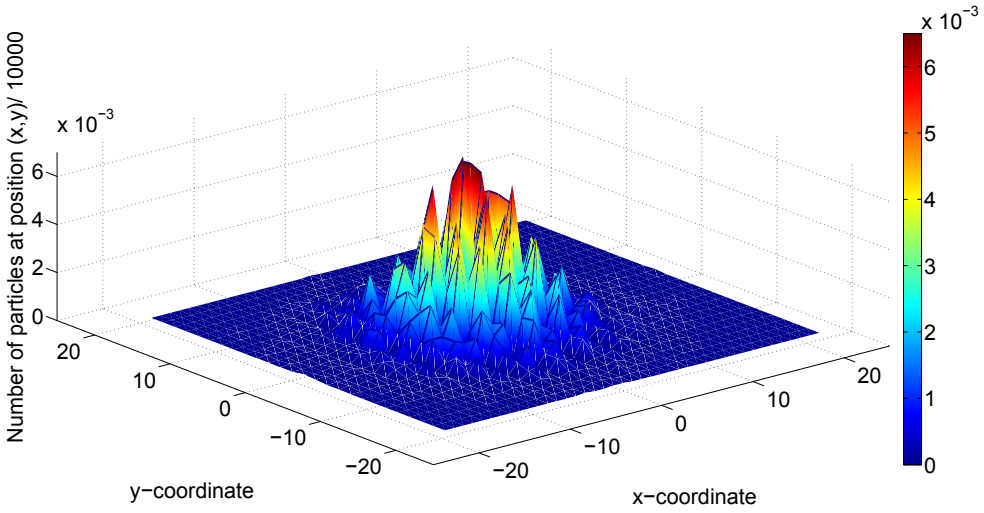

Figure 3.3: Distribution of 10,000 Markov chains at t=50s

the respective multivariate Gaussian distribution on top of a histogram, we receive a very good match in terms of height and width of the distributions.

As in Chapter 2, lower numbers of employed particles result in less accurate results regarding the matching of shapes as well as a higher deviation of the mean and covariance from the ideal mean of 0 and covariance of $\frac{1}{2}tI_2$ (see Table 3.1). As mentioned in Chapter 1, a large number of particles is supposed to be used. In Figures 3.9 and 3.10, 1000 and 100,000 particles were employed, and we can see that the Gaussian bell curve is much easier to identify in the shape of the histogram, if we employ 100,000 particles.

3.6 Other Lattice Structures

As an example of other lattice structures we consider the square lattice given in Figure 3.11. Each node has four neighbors, where all edges have length a. The state-transition diagram of the continuous-time Markov chain on this lattice is given in Figure 3.12.

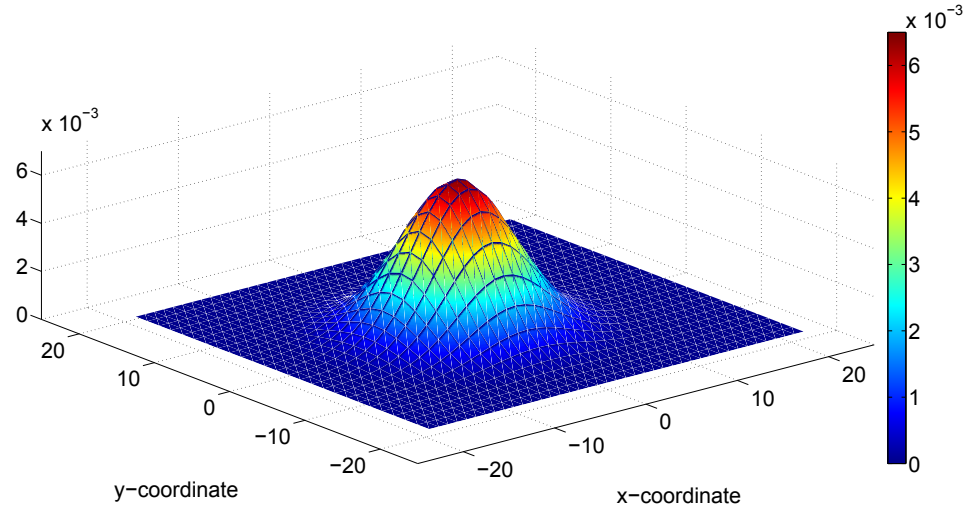

Figure 3.4: Multivariate Gaussian at t=50s

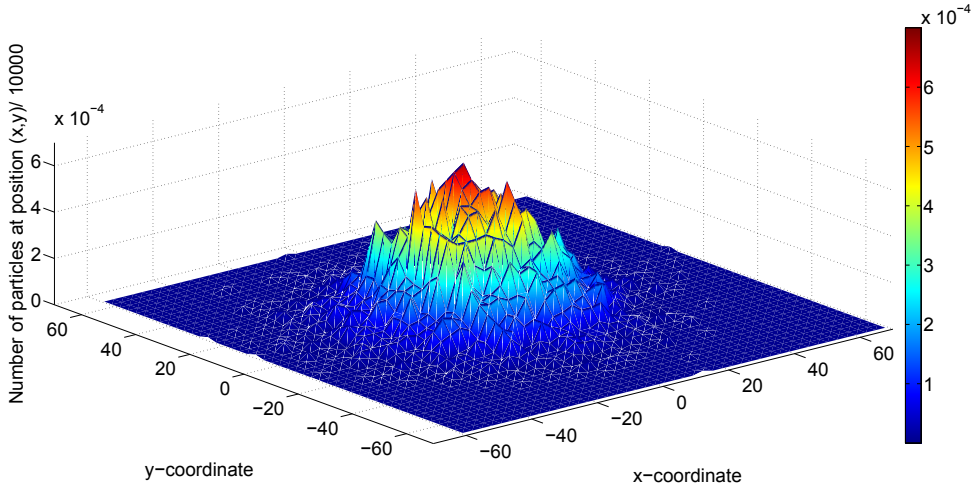

Figure 3.5: Distribution of $10,000$ Markov chains at t=500s

Theorem 3.2. Let $w(t)$ be a random variable of a continuous-time Markov chain whose state-transition diagram is given by Figure 3.12. Furthermore, let $w(0) = x_{0,0} (= (0,0))$ and let

$$\lambda_{\mathrm{SQ}} \quad = \quad \frac{1}{4a^2},$$

then the random variable $w(t)$ has a covariance matrix $\frac{1}{2}tI_2$ and mean 0 for all time $t > 0$.

The proof for Theorem 3.2 is analogue to Theorem 3.1 and is therefore omitted. On a square lattice the transition rate becomes $\frac{1}{4a^2}$, one divided by the number of edges times

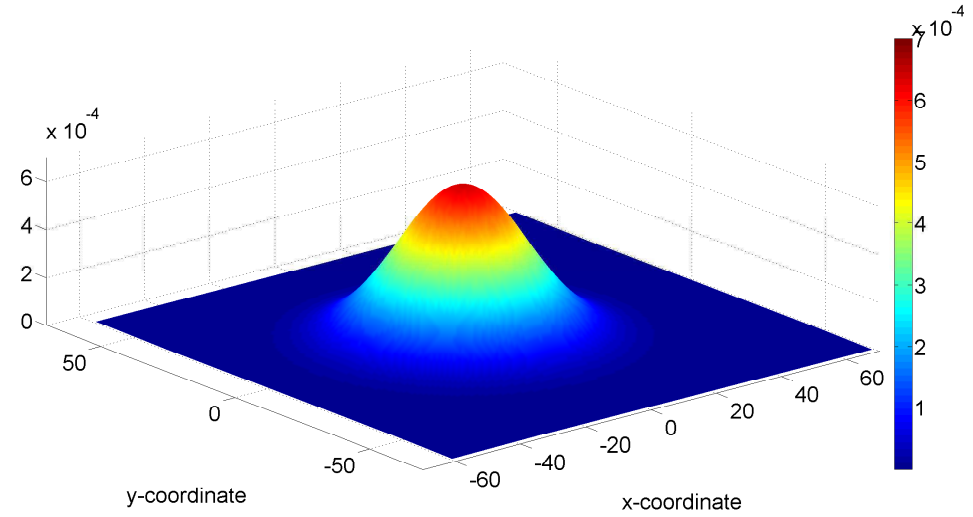

Figure 3.6: Multivariate Gaussian at t=500s

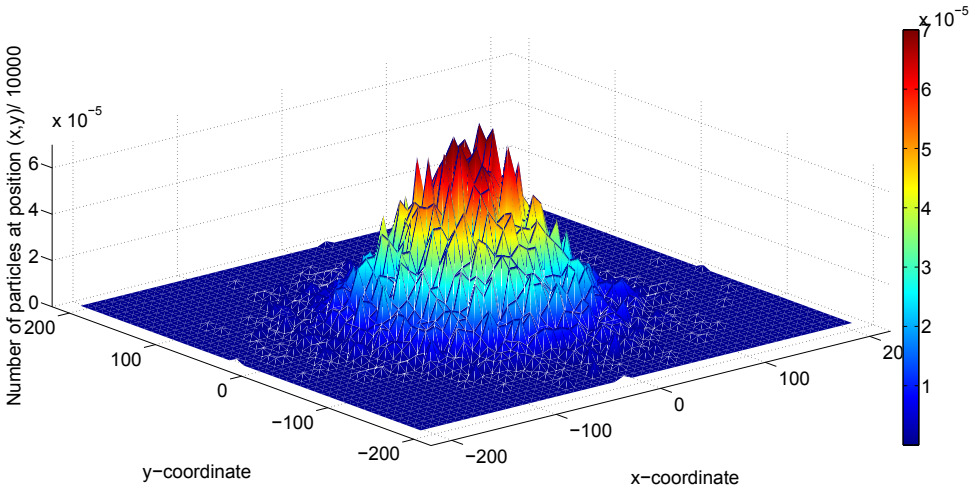

Figure 3.7: Distribution of $10,000$ Markov chains at t=5000s

the edge length squared. If we set all six edges in Figure 3.1 to length a, all transition rates become $\frac{1}{6a^2}$. Intuitively this makes a lot of sense and is even expected.

In this lattice the angle between each pair of neighboring edges is 90 degrees. This fact makes this lattice undesirable when we have to deal with slightly distorted lattices where the angle is not 90 degrees, e.g., when a lattice from a sphere in $\mathbb{R}^3$ is projected onto the plane in $\mathbb{R}^2$.

As another example of other lattice structures we consider the lattice in Figure 3.13. This lattice can be thought of as the lattice given previously in this chapter, skewed to the left. All horizontal edges have the length of a, all ascending diagonal edges have the

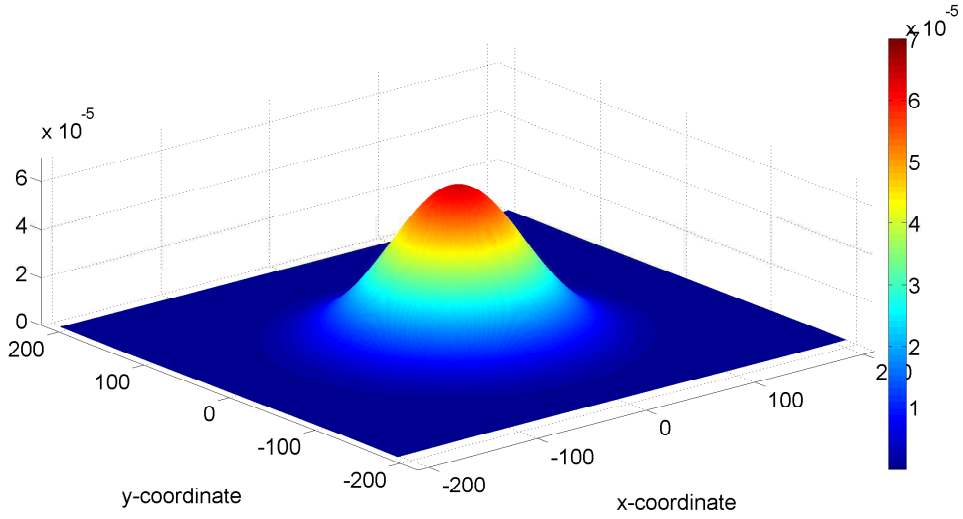

Figure 3.8: Multivariate Gaussian at t=5000s

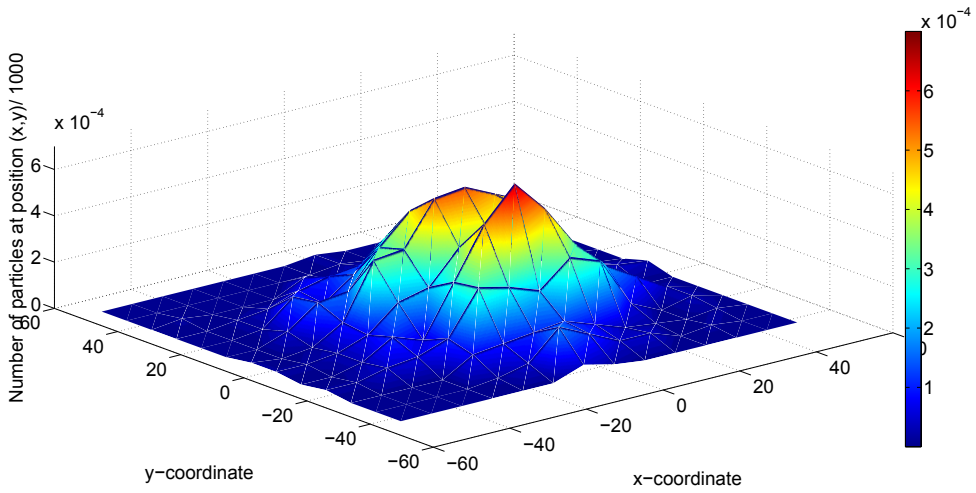

Figure 3.9: Distribution of 1000 Markov chains at t=500s

length of c and all descending edges have the length of b. The vertical distance between two horizontal edges is denoted as y and $\beta \in [0, 1]$ indicates how far the lattice is skewed to the left. As we skew the lattice more the the left, y decreases and β increases. The state-transition diagram is the same as in the previous example in this chapter, given in Figure 3.2.

Theorem 3.3. Let $w(t)$ be a random variable of a continuous-time Markov chain whose state-transition diagram is given by Figure 3.2. Furthermore, let $w(0) = x_{0,0} (= (0, 0))$ and

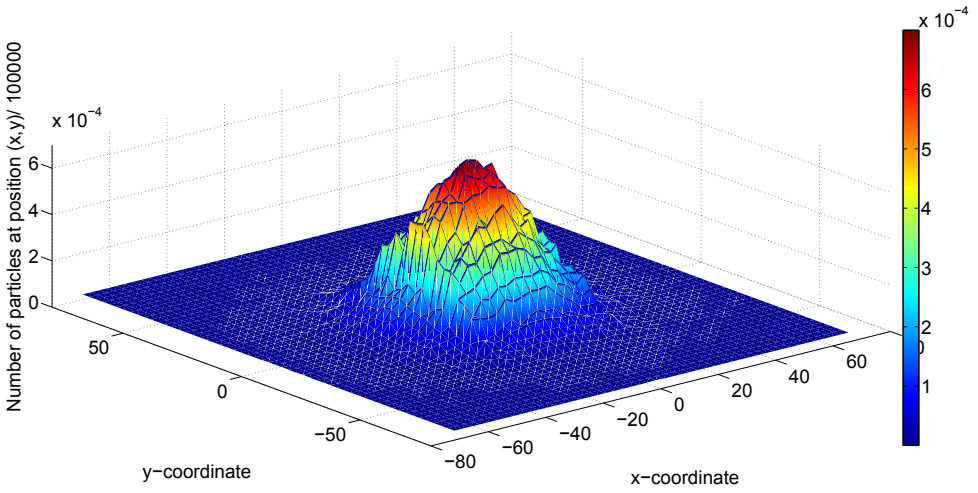

Figure 3.10: Distribution of 100,000 Markov chains at t=500s

let

$$\begin{aligned} \lambda_{(i,j),(i+1,j)} &= \lambda_{(i,j),(i-1,j)} \triangleq \lambda^{(1)}, \\ \lambda_{(i,j),(i,j-1)} &= \lambda_{(i,j),(i,j+1)} \triangleq \lambda^{(2)}, \\ \lambda_{(i,j),(i-1,j+1)} &= \lambda_{(i,j),(i+1,j-1)} \triangleq \lambda^{(3)}, \end{aligned} \tag{3.26}$$

where

$$\lambda^{(1)} \triangleq \frac{\beta^2 - \beta}{4y^2} + \frac{1}{4a^2}, \tag{3.27}$$

$$\lambda^{(2)} \triangleq \frac{\beta}{4y^2}, \tag{3.28}$$

$$\lambda^{(3)} \triangleq \frac{1 - \beta^2}{4y^2}. \tag{3.29}$$

If $b^2 + c^2 - a^2 \geq 0$, then the random variable $w(t)$ has a covariance matrix $\frac{1}{2}tI_2$ and mean 0 for all time $t > 0$.

We omit the proof for this theorem since the derivation of the transition is analogue to the proof of Theorem 3.1. Note that the requirement $b^2 + c^2 - a^2 \geq 0$ translates to a maximum angle between edges b and c of 90 degrees. The lattice in Figure 3.13 extends the class of lattices usable with our framework. There are three different transition rates for this class of lattices. Note that we can recover the same transition rates as in Theorem

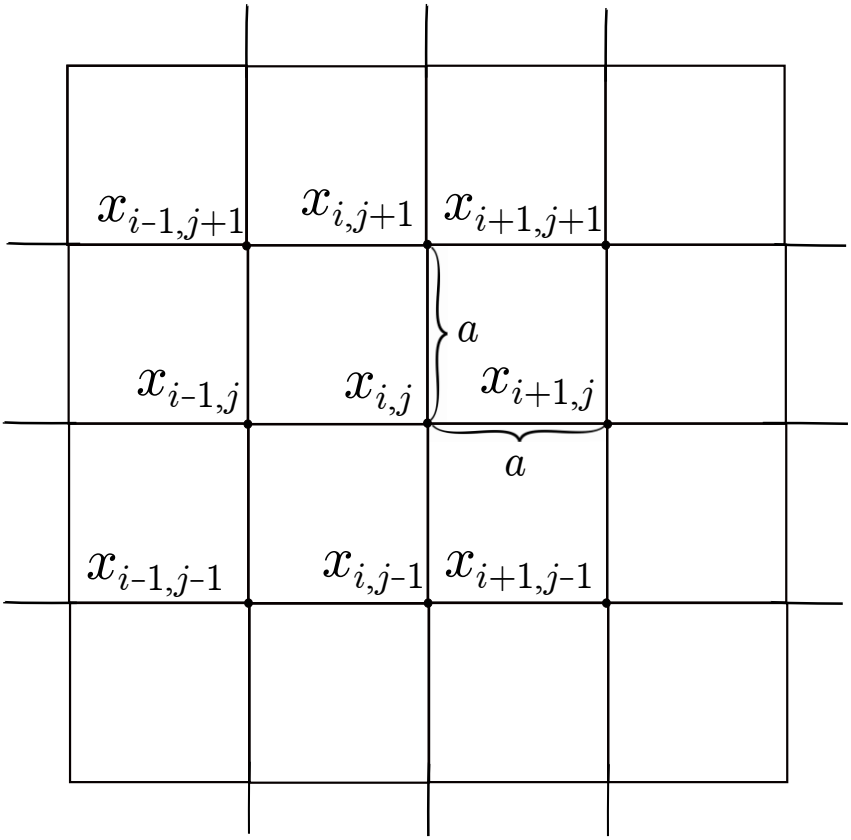

Figure 3.11: Example square lattice on the infinite plane

3.1 if we set β to $\frac{1}{2}$. In that case b equals c such that

$$\lambda^{(1)} = \frac{2b^2 - a^2}{2a^2(4b^2 - a^2)}, \tag{3.30}$$

$$\lambda^{(2)} \triangleq \frac{1}{2(4b^2 - a^2)}, \tag{3.31}$$

$$\lambda^{(3)} \triangleq \frac{1}{2(4b^2 - a^2)}. \tag{3.32}$$

Also note that the nodes in each of the previous lattices on the plane have the same configuration of edges and neighboring nodes. This limits the lattices we can construct on the plane. For instance, we cannot construct a lattice with four different transition rates. This is an obstacle and we show how we can overcome this limitation in the next chapter.

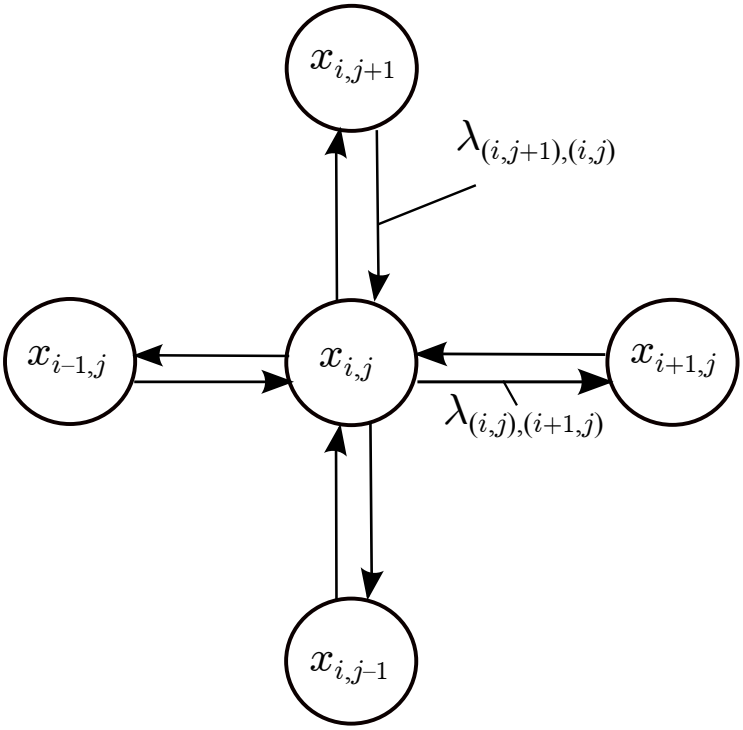

Figure 3.12: State-transition diagram of the continuous-time Markov chain on the example square lattice

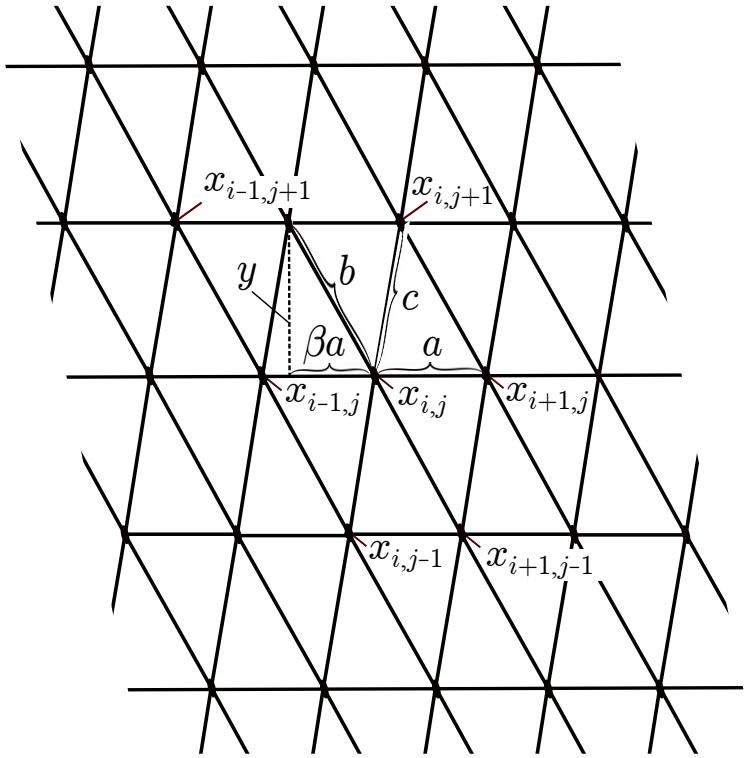

Figure 3.13: Lattice structure with three different transition rates

Chapter 4

Heat Diffusion Modelling on Smallest Unit-Based Lattices

4.1 Introduction

The lattices in Chapter 3 represent only a part of lattices where our framework for the modelling of heat diffusion is usable. In this chapter we consider a new class of lattices. Specifically, we consider a lattice with missing edges. To deal with this class of lattices we subdivide the lattice into units and different types of nodes, where each type of node has a different number of neighbors. For the random variable of a continuous-time Markov chain on such lattices, we then show how can still guarantee the mean and covariance matrix of a Gaussian distribution, such that the solution to the heat equation is approximated.

Note that we use the same definition of the continuous-time Markov chain as in Chapter 3.

4.2 Markov-Chain-Driven Random Walk on a Lattice with Missing Edges

In this section, we discuss Markov-chain-driven random walks on a lattice in the two-dimensional space. For that, nodes $x_{i,j}$, $i, j \in \mathbb{I}$, are placed on the lattice in Figure 4.1. Each particle is initially placed at node $x_{0,0}$ and jumps between neighboring nodes. Each node is connected to five or six neighboring nodes and the connecting edges consist of

diagonal and horizontal lines, where all edges have the same length a. Each node, with its connected edges, is categorized into a type, which is the expression for the configuration of one nodes and its connected edges on a lattice. This idea of a lattice with different types allows for more irregular lattice structures than in Chapter 3. The important feature in the example shown in Figure 4.1 is that some nodes have only five neighbors. To approximate the diffusion of heat on such a lattice we first define the smallest unit of the lattice as shown in Figure 4.2. With this smallest unit, we can create the entire lattice by placing those units next to each other. This idea of breaking the lattice down into units helps to identify the different types of nodes on this lattice. Each type of node has a different arrangement of neighbors and edges connected to it.

As indicated by the numbers next to the nodes in Figure 4.2, we have four types of nodes for this lattice. The node $x_{i,j+2}$ for instance is a type four node, since it has the same configuration of nodes and edges around it as $x_{i,j}$. If we let the particles randomly jump between the neighboring five or six lattice nodes, we can describe the stochastic process of the successive random jumps to neighboring nodes with the state transition diagram in Figure 4.3. The transition rate from state $x_{i,j}$ to state $x_{i+1,j}$ for instance is denoted as $\lambda_{(i,j),(i+1,j)}$, the transition rate from state $x_{i-1,j+1}$ to state $x_{i,j}$ is denoted as $\lambda_{(i-1,j+1),(i,j)}$, and so on. In Figure 4.3 we focus on the shaded nodes, since they represent the four types of nodes we define.

Note that the behavior of the random walk is captured by a continuous-time Markov chain with infinitesimal generator Q, with elements $\lambda_{(i,j),(k,l)}$. Furthermore, note that the time derivative of (1.10) is given by

$$\frac{\mathrm{d}p_{i,j}(t)}{\mathrm{d}t} = (e^{Qt}Q)_{(0,0),(i,j)}. \tag{4.1}$$

From any given node, particles can jump only to the five or six neighboring nodes. The transition rates to the other nodes are therefore 0. Due to this structure of Q and the Kolmogorov forward equation, we can rewrite (4.1) as a combination of transition rates and transition probabilities to each neighboring node. For the sake of the following derivation we change the indexing slightly. Type 1 nodes are indexed as $x_{2i+1,2j}$, type 2 nodes are indexed as $x_{2i,2j+1}$, type 3 nodes are indexed as $x_{2i+1,2j-1}$, and type 4 nodes are indexed

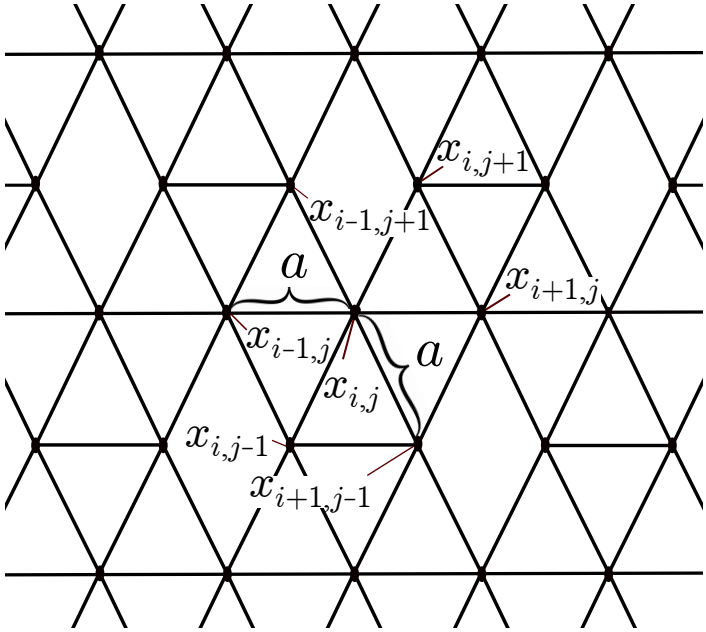

Figure 4.1: Example of a triangular lattice with missing edges on the infinite plane, where an increase of index i by 1 corresponds to a jump of length a to the right and a decrease if index i by 1 corresponds to a jump of length a to the left. An increase of index j by 1 corresponds to a jump of length a along the ascending edges and a decrease of index j by 1 corresponds to a jump of length a along the descending edges.

as $x_{2i,2j}$. With $x_{2i+1,2j}$, for instance, for any choice of i and j, we can only obtain type 1 nodes. With these indices, we write (4.1) for each type of node and get

$$\begin{aligned}
\frac{\mathrm{d}p_{2i+1,2j}(t)}{\mathrm{d}t} &= \lambda_{(2i,2j),(2i+1,2j)}p_{2i,2j}(t) + \lambda_{(2i,2j+1),(2i+1,2j)}p_{2i,2j+1}(t) \\
&\quad +\lambda_{(2i+1,2j-1),(2i+1,2j)}p_{2i+1,2j-1}(t) - (\lambda_{(2i+1,2j),(2i,2j)} \\
&\quad +\lambda_{(2i+1,2j),(2i,2j+1)} + \lambda_{(2i+1,2j),(2i+1,2j-1)} \\
&\quad +\lambda_{(2i+1,2j),(2i+1,2j+1)} + \lambda_{(2i+1,2j),(2i+2,2j)} \\
&\quad +\lambda_{(2i+1,2j),(2i+2,2j-1)})p_{2i+1,2j}(t) + \lambda_{(2i+1,2j+1),(2i+1,2j)}p_{2i+1,2j+1}(t) \\
&\quad +\lambda_{(2i+2,2j),(2i+1,2j)}p_{2i+2,2j}(t) \\
&\quad +\lambda_{(2i+2,2j-1),(2i+1,2j)}p_{2i+2,2j-1}(t), \qquad (4.2) \\
\frac{\mathrm{d}p_{2i,2j+1}(t)}{\mathrm{d}t} &= \lambda_{(2i,2j),(2i,2j+1)}p_{2i,2j}(t) + \lambda_{(2i+1,2j),(2i,2j+1)}p_{2i+1,2j}(t) \\
&\quad +\lambda_{(2i+1,2j+1),(2i,2j+1)}p_{2i+1,2j+1}(t) - (\lambda_{(2i,2j+1),(2i,2j)}
\end{aligned}$$

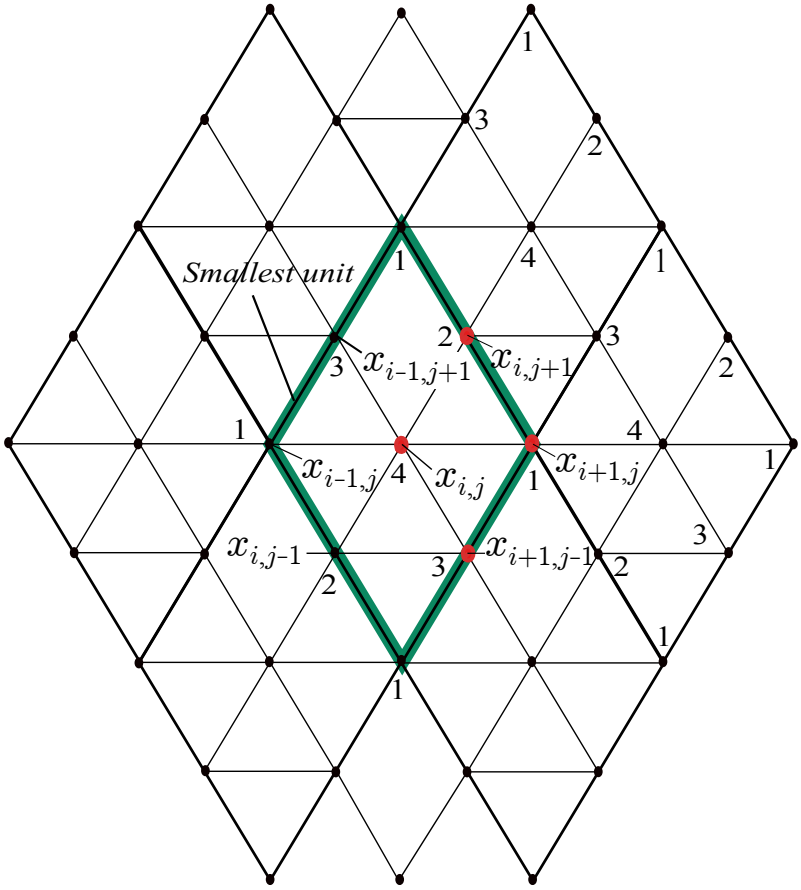

Figure 4.2: Smallest unit and types of nodes on example triangular lattice with missing edges, where the numbers next to the nodes indicate the type of the node

$$
\begin{aligned}
& +\lambda_{(2i,2j+1),(2i+1,2j)} + \lambda_{(2i,2j+1),(2i+1,2j+1)} + \lambda_{(2i,2j+1),(2i,2j+2)} \\
& +\lambda_{(2i,2j+1),(2i-1,2j+2)})p_{2i,2j+1}(t) + \lambda_{(2i,2j+2),(2i,2j+1)}p_{2i,2j+2}(t) \\
& +\lambda_{(2i-1,2j+2),(2i,2j+1)}p_{2i-1,2j+2}(t), \qquad (4.3) \\
\frac{\mathrm{d}p_{2i+1,2j-1}(t)}{\mathrm{d}t} &= \lambda_{(2i,2j),(2i+1,2j-1)}p_{2i,2j}(t) + \lambda_{(2i+1,2j),(2i+1,2j-1)}p_{2i+1,2j}(t) \\
& +\lambda_{(2i,2j-1),(2i+1,2j-1)}p_{2i,2j-1}(t) - (\lambda_{(2i+1,2j-1),(2i,2j)} \\
& +\lambda_{(2i+1,2j-1),(2i+1,2j)} + \lambda_{(2i+1,2j-1),(2i,2j-1)} + \lambda_{(2i+1,2j-1),(2i+1,2j-2)} \\
& +\lambda_{(2i+1,2j-1),(2i+2,2j-2)})p_{2i+1,2j-1}(t) \\
& +\lambda_{(2i+1,2j-2),(2i+1,2j-1)}p_{2i+1,2j-2}(t) \\
& +\lambda_{(2i+2,2j-2),(2i+1,2j-1)}p_{2i+2,2j-2}(t), \qquad (4.4) \\
\frac{\mathrm{d}p_{2i,2j}(t)}{\mathrm{d}t} &= \lambda_{(2i+1,2j),(2i,2j)}p_{2i+1,2j}(t) + \lambda_{(2i-1,2j),(2i,2j)}p_{2i-1,2j}(t) \\
& +\lambda_{(2i+1,2j-1),(2i,2j)}p_{2i+1,2j-1}(t) - (\lambda_{(2i,2j),(2i+1,2j)} + \lambda_{(2i,2j),(2i-1,2j)})
\end{aligned}
$$

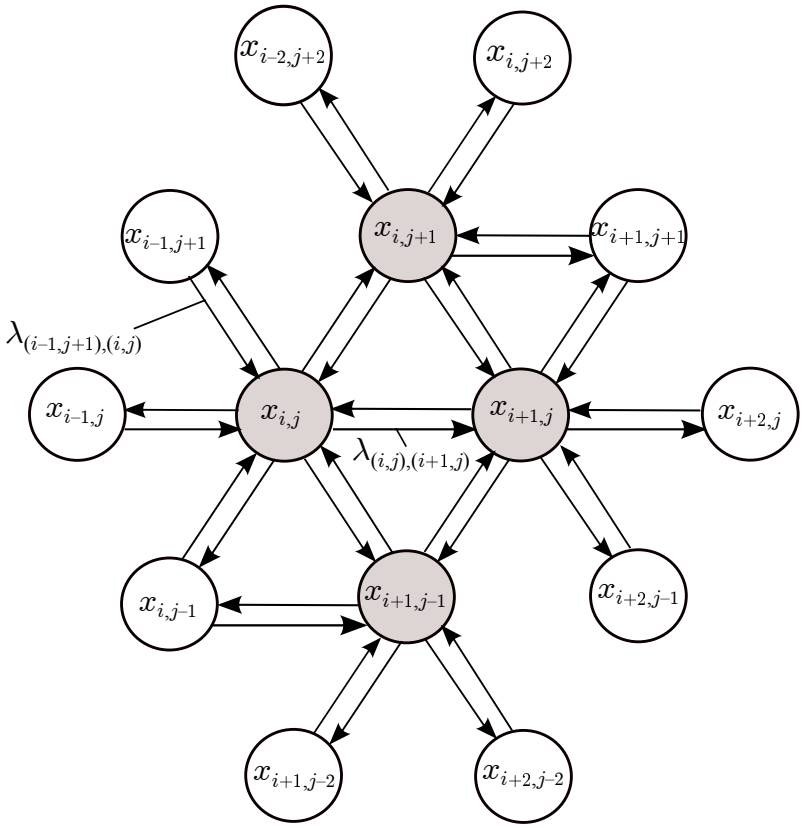

Figure 4.3: State-transition diagram of the continuous-time Markov chain on the example triangular lattice with missing edges

$$
\begin{aligned}
&+\lambda_{(2i,2j),(2i+1,2j-1)} + \lambda_{(2i,2j),(2i-1,2j+1)} + \lambda_{(2i,2j),(2i,2j+1)} \\
&+\lambda_{(2i,2j),(2i,2j-1)})p_{2i,2j}(t) + \lambda_{(2i-1,2j+1),(2i,2j)}p_{2i-1,2j+1}(t) \\
&+\lambda_{(2i,2j+1),(2i,2j)}p_{2i,2j+1}(t) + \lambda_{(2i,2j-1),(2i,2j)}p_{2i,2j-1}(t). \qquad (4.5)
\end{aligned}
$$

4.3 Mean and Covariance of the Continuous-Time Markov Chain

In Chapter 3 we already derived the theoretical mean and covariance matrix for the random variable of the two-dimensional Gaussian distribution, i.e., mean 0 and covariance $\frac{1}{2}tI_2$. We can therefore directly continue with the proof that the random variable of the Markov chain can be characterized in such a way that the mean and the covariance matrix of the random variable of the Markov chain are equal to the theoretical mean and covariance matrix of the two-dimensional Gaussian distribution. The proof works slightly different and is also a bit longer.

Theorem 4.1. Let $w(t)$ be a random variable of a continuous-time Markov chain whose state-transition diagram is given by Figure 4.3. Furthermore, let $w(0) = x_{0,0} (= 0, 0)$ and let

$$\begin{aligned}
\lambda_{(2i,2j),(m,n)} &= \lambda_{(2i+1,2j),(k,l)} \triangleq \lambda^{(1)}, \\
\lambda_{(2i,2j+1),(2i+1,2j+1)} &= \lambda_{(2i,2j+1),(2i,2j)} = \lambda_{(2i,2j+1),(2i-1,2j+2)} \\
&= \lambda_{(2i+1,2j-1),(2i+1,2j)} = \lambda_{(2i+1,2j-1),(2i,2j-1)} \\
&= \lambda_{(2i+1,2j-1),(2i+2,2j-2)} \triangleq \lambda^{(2)}, \\
\lambda_{(2i,2j+1),(2i+1,2j)} &= \lambda_{(2i,2j+1),(2i,2j+2)} = \lambda_{(2i+1,2j-1),(2i+1,2j-2)} \\
&\triangleq \lambda^{(3)}, \qquad (4.6)
\end{aligned}$$

where $\lambda_{(2i,2j),(m,n)}$ and $\lambda_{(2i+1,2j),(k,l)}$ denote all transition rates from $x_{2i,2j}$ and $x_{2i+1,2j}$ and where

$$\lambda^{(1)} \triangleq \frac{1}{6a^2}, \qquad (4.7)$$

$$\lambda^{(2)} \triangleq \frac{1}{3a^2}, \qquad (4.8)$$

$$\lambda^{(3)} \triangleq 0. \qquad (4.9)$$

Then the random variable $w(t)$ has a covariance matrix $\frac{1}{2}tI_2$ and mean 0 for all time $t > 0$.

Proof. The covariance of $w(t)$ at time t can be calculated using (1.10) and the well known equation for the covariance [43]

$$\begin{aligned}
\mathbb{E}[w(t)w^{\mathrm{T}}(t)] - \mathbb{E}[w(t)]\mathbb{E}[w^{\mathrm{T}}(t)] &= \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{i,j}x_{i,j}^{\mathrm{T}}\mathbb{P}[w(t) = x_{i,j}|w(0) = x_{0,0}] \\
&\quad -\Big(\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{i,j}\mathbb{P}[w(t) = x_{i,j}|w(0) = x_{0,0}] \\
&\quad \cdot \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{i,j}^{\mathrm{T}}\mathbb{P}[w(t) = x_{i,j}|w(0) = x_{0,0}]\Big) \qquad (4.10)
\end{aligned}$$

To find the covariance, we look for the expectation of $w(t)$ first. Since we have four

different stencils, we write the expectation as

$$\begin{aligned}\mathbb{E}[w(t)] &= \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}\Big(x_{2i,2j}\mathbb{P}[w(t)=x_{2i,2j}|w(0)=x_{0,0}]\\ &\quad +x_{2i,2j+1}\mathbb{P}[w(t)=x_{2i,2j+1}|w(0)=x_{0,0}]\\ &\quad +x_{2i+1,2j-1}\mathbb{P}[w(t)=x_{2i+1,2j-1}|w(0)=x_{0,0}]\\ &\quad +x_{2i+1,2j}\mathbb{P}[w(t)=x_{2i+1,2j}|w(0)=x_{0,0}]\Big)\\ &= \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}\Big(x_{i,2j}p_{i,2j}+x_{2i,2j+1}p_{2i,2j+1}\\ &\quad +x_{2i+1,2j-1}p_{2i+1,2j-1}+x_{2i+1,2j}p_{2i+1,2j}\Big). \end{aligned} \tag{4.11}$$

Taking the time derivative of (4.11) gives us

$$\begin{aligned}\frac{\mathrm{d}\mathbb{E}[w(t)]}{\mathrm{d}t} &= \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}\Big(x_{i,2j}\frac{\mathrm{d}p_{i,2j}}{\mathrm{d}t}+x_{2i,2j+1}\frac{\mathrm{d}p_{2i,2j+1}}{\mathrm{d}t}\\ &\quad +x_{2i+1,2j-1}\frac{\mathrm{d}p_{2i+1,2j-1}}{\mathrm{d}t}+x_{2i+1,2j}\frac{\mathrm{d}p_{2i+1,2j}}{\mathrm{d}t}\Big). \end{aligned} \tag{4.12}$$

Using (4.2)–(4.5), we write (4.12) as

$$\begin{aligned}\frac{\mathrm{d}\mathbb{E}[w(t)]}{\mathrm{d}t} &= \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}x_{2i+1,2j}\Big(\lambda_{(2i,2j),(2i+1,2j)}p_{2i,2j}(t)\\ &\quad +\lambda_{(2i,2j+1),(2i+1,2j)}p_{2i,2j+1}(t)+\lambda_{(2i+1,2j-1),(2i+1,2j)}p_{2i+1,2j-1}(t)\\ &\quad -(\lambda_{(2i+1,2j),(2i,2j)}+\lambda_{(2i+1,2j),(2i,2j+1)}+\lambda_{(2i+1,2j),(2i+1,2j-1)}\\ &\quad +\lambda_{(2i+1,2j),(2i+1,2j+1)}+\lambda_{(2i+1,2j),(2i+2,2j)}\\ &\quad +\lambda_{(2i+1,2j),(2i+2,2j-1)})p_{2i+1,2j}(t)+\lambda_{(2i+1,2j+1),(2i+1,2j)}p_{2i+1,2j+1}(t)\\ &\quad +\lambda_{(2i+2,2j),(2i+1,2j)}p_{2i+2,2j}(t)+\lambda_{(2i+2,2j-1),(2i+1,2j)}p_{2i+2,2j-1}(t)\Big)\\ &\quad +\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}x_{2i,2j+1}\Big(\lambda_{(2i,2j),(2i,2j+1)}p_{2i,2j}(t)\\ &\quad +\lambda_{(2i+1,2j),(2i,2j+1)}p_{2i+1,2j}(t)+\lambda_{(2i+1,2j+1),(2i,2j+1)}p_{2i+1,2j+1}(t)\\ &\quad -(\lambda_{(2i,2j+1),(2i,2j)}+\lambda_{(2i,2j+1),(2i+1,2j)}+\lambda_{(2i,2j+1),(2i+1,2j+1)}\\ &\quad +\lambda_{(2i,2j+1),(2i,2j+2)}+\lambda_{(2i,2j+1),(2i-1,2j+2)})p_{2i,2j+1}(t)\\ &\quad +\lambda_{(2i,2j+2),(2i,2j+1)}p_{2i,2j+2}(t)+\lambda_{(2i-1,2j+2),(2i,2j+1)}p_{2i-1,2j+2}(t)\Big)\end{aligned}$$

$$
\begin{aligned}
&+\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{2i+1,2j-1}\Big(\lambda_{(2i,2j),(2i+1,2j-1)}p_{2i,2j}(t)\\
&+\lambda_{(2i+1,2j),(2i+1,2j-1)}p_{2i+1,2j}(t)+\lambda_{(2i,2j-1),(2i+1,2j-1)}p_{2i,2j-1}(t)\\
&-(\lambda_{(2i+1,2j-1),(2i,2j)}+\lambda_{(2i+1,2j-1),(2i+1,2j)}+\lambda_{(2i+1,2j-1),(2i,2j-1)}\\
&+\lambda_{(2i+1,2j-1),(2i+1,2j-2)}+\lambda_{(2i+1,2j-1),(2i+2,2j-2)})p_{2i+1,2j-1}(t)\\
&+\lambda_{(2i+1,2j-2),(2i+1,2j-1)}p_{2i+1,2j-2}(t)+\lambda_{(2i+2,2j-2),(2i+1,2j-1)}p_{2i+2,2j-2}(t)\Big)\\
&+\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{2i,2j}\Big(\lambda_{(2i+1,2j),(2i,2j)}p_{2i+1,2j}(t)\\
&+\lambda_{(2i-1,2j),(2i,2j)}p_{2i-1,2j}(t)+\lambda_{(2i+1,2j-1),(2i,2j)}p_{2i+1,2j-1}(t)\\
&-(\lambda_{(2i,2j),(2i+1,2j)}+\lambda_{(2i,2j),(2i-1,2j)}+\lambda_{(2i,2j),(2i+1,2j-1)}+\lambda_{(2i,2j),(2i-1,2j+1)}\\
&+\lambda_{(2i,2j),(2i,2j+1)}+\lambda_{(2i,2j),(2i,2j-1)})p_{2i,2j}(t)\\
&+\lambda_{(2i-1,2j+1),(2i,2j)}p_{2i-1,2j+1}(t)+\lambda_{(2i,2j+1),(2i,2j)}p_{2i,2j+1}(t)\\
&+\lambda_{(2i,2j-1),(2i,2j)}p_{2i,2j-1}(t)\Big). \qquad (4.13)
\end{aligned}
$$

Now, consider the first term of (4.13). We rewrite the first term of (4.13) as

$$
\begin{aligned}
\frac{\mathrm{d}\mathbb{E}[w(t)]}{\mathrm{d}t} &= \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{2i+1,2j}\lambda_{(2i,2j),(2i+1,2j)}p_{2i,2j}(t)\\
&+\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{2i+1,2j}\lambda_{(2i,2j+1),(2i+1,2j)}p_{2i,2j+1}(t)\\
&+\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{2i+1,2j}\lambda_{(2i+1,2j-1),(2i+1,2j)}p_{2i+1,2j-1}(t)\\
&-\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{2i+1,2j}(\lambda_{(2i+1,2j),(2i,2j)}+\lambda_{(2i+1,2j),(2i,2j+1)}\\
&+\lambda_{(2i+1,2j),(2i+1,2j-1)}+\lambda_{(2i+1,2j),(2i+1,2j+1)}\\
&+\lambda_{(2i+1,2j),(2i+2,2j)}+\lambda_{(2i+1,2j),(2i+2,2j-1)})p_{2i+1,2j}(t)\\
&+\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{2i+1,2j}\lambda_{(2i+1,2j+1),(2i+1,2j)}p_{2i+1,2j+1}(t)\\
&+\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{2i+1,2j}\lambda_{(2i+2,2j),(2i+1,2j)}p_{2i+2,2j}(t)\\
&+\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} x_{2i+1,2j}\lambda_{(2i+2,2j-1),(2i+1,2j)}p_{2i+2,2j-1}(t). \qquad (4.14)
\end{aligned}
$$

Due to the infinite sums, we can shift the indices of each of the seven terms of (4.14). We

decrease j by one in the fifth term, decrease i by one in the sixth term and decrease i by one and increase j by one in the seventh term. We then get

$$\begin{aligned}
\frac{\mathrm{d}\mathbb{E}[w(t)]}{\mathrm{d}t} &= \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} p_{2i,2j}(t)\Big(x_{2i+1,2j}\lambda_{(2i,2j),(2i+1,2j)} + x_{2i-1,2j}\lambda_{(2i,2j),(2i-1,2j)}\Big) \\
&+p_{2i,2j+1}(t)\Big(x_{2i+1,2j}\lambda_{(2i,2j+1),(2i+1,2j)} + x_{2i-1,2j+2}\lambda_{(2i,2j+1),(2i-1,2j+2)}\Big) \\
&+p_{2i+1,2j-1}(t)\Big(x_{2i+1,2j}\lambda_{(2i+1,2j-1),(2i+1,2j)} \\
&+x_{2i+1,2j-2}\lambda_{(2i+1,2j-1),(2i+1,2j-2)}\Big) - x_{2i+1,2j}(\lambda_{(2i+1,2j),(2i,2j)} \\
&+\lambda_{(2i+1,2j),(2i,2j+1)} + \lambda_{(2i+1,2j),(2i+1,2j-1)} \\
&+\lambda_{(2i+1,2j),(2i+1,2j+1)} + \lambda_{(2i+1,2j),(2i+2,2j)} \\
&+\lambda_{(2i+1,2j),(2i+2,2j-1)})p_{2i+1,2j}(t). \qquad (4.15)
\end{aligned}$$

Here we group the probabilities of being at the centers of each of the four types of nodes, i.e., $p_{2i,2j}(t)$, $p_{2i,2j+1}(t)$, $p_{2i+1,2j-1}(t)$, and $p_{2i+1,2j}$. By using the shifting technique applied in (4.14) and (4.15) on the remaining terms of (4.13) and then substituting the transition rates from (4.6) we get

$$\begin{aligned}
\frac{\mathrm{d}\mathbb{E}[w(t)]}{\mathrm{d}t} &= \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} p_{2i+1,2j}(t)\lambda^{(1)}\Big((x_{2i,2j+1} - x_{2i+1,2j}) + (x_{2i+2,2j-1} - x_{2i+1,2j}) \\
&+(x_{2i+1,2j-1} - x_{2i+1,2j}) + (x_{2i+1,2j+1} - x_{2i+1,2j}) \\
&+(x_{2i,2j} - x_{2i+1,2j}) + (x_{2i+2,2j} - x_{2i+1,2j})\Big) \\
&+ \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} p_{2i,2j}(t)\lambda^{(1)}\Big((x_{2i+1,2j} - x_{2i,2j}) + (x_{2i-1,2j} - x_{2i,2j}) \\
&+(x_{2i,2j+1} - x_{2i,2j}) + (x_{2i,2j-1} - x_{2i,2j}) + (x_{2i+1,2j-1} - x_{2i,2j}) \\
&+(x_{2i-1,2j+1} - x_{2i,2j})\Big) \\
&+ \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} p_{2i,2j+1}(t)\Big((x_{2i+1,2j} - x_{2i,2j+1})\lambda^{(3)} \\
&+(x_{2i-1,2j+2} - x_{2i,2j+1})\lambda^{(2)} + (x_{2i+1,2j+1} - x_{2i,2j+1})\lambda^{(2)} \\
&+(x_{2i,2j} - x_{2i,2j+1})\lambda^{(2)} + (x_{2i,2j+2} - x_{2i,2j+1})\lambda^{(3)}\Big) \\
&+ \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} p_{2i+1,2j-1}(t)\Big((x_{2i+1,2j} - x_{2i+1,2j-1})\lambda^{(2)} \\
&+(x_{2i+1,2j-2} - x_{2i+1,2j-1})\lambda^{(3)} + (x_{2i,2j-1} - x_{2i+1,2j-1})\lambda^{(2)}
\end{aligned}$$

$$+(x_{2i+2,2j-2} - x_{2i+1,2j-1})\lambda^{(2)} + (x_{2i,2j} - x_{2i+1,2j-1})\lambda^{(3)}\Big). \tag{4.16}$$

Inserting the coordinates from the example lattice in Figure 4.1 and substituting (4.7)–(4.9) leads to

$$\begin{aligned}\frac{\mathrm{d}\mathbb{E}[w(t)]}{\mathrm{d}t} &= \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}\Big(\begin{bmatrix}0\\0\end{bmatrix}p_{2i,2j}(t) + \begin{bmatrix}0\\0\end{bmatrix}p_{2i,2j+1}(t) \\ &\quad + \begin{bmatrix}0\\0\end{bmatrix}p_{2i+1,2j-1}(t) + \begin{bmatrix}0\\0\end{bmatrix}p_{2i+1,2j}(t)\Big) \\ &= 0. \end{aligned} \tag{4.17}$$

Integrating (4.17) gives us a mean of 0, since $w(0) = (0, 0)$. This result simplifies the covariance to $\mathbb{E}[w(t)w^{\mathrm{T}}(t)]$. To find the covariance, we use the same approach we used for the mean. In fact, we only need to substitute $x_{i,j}x_{i,j}^{\mathrm{T}}$ for $x_{i,j}$ in (4.16), insert the coordinates for the example lattice and substitute (4.7)–(4.9). This gives us

$$\begin{aligned}\frac{\mathrm{d}\mathbb{E}[w(t)w^{\mathrm{T}}(t)]}{\mathrm{d}t} &= \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}\Big(p_{2i,2j}(t)\begin{bmatrix}\frac{1}{2} & 0\\ 0 & \frac{1}{2}\end{bmatrix} + p_{2i,2j+1}(t)\begin{bmatrix}\frac{1}{2} & 0\\ 0 & \frac{1}{2}\end{bmatrix} \\ &\quad + p_{2i+1,2j-1}(t)\begin{bmatrix}\frac{1}{2} & 0\\ 0 & \frac{1}{2}\end{bmatrix} + p_{2i+1,2j}(t)\begin{bmatrix}\frac{1}{2} & 0\\ 0 & \frac{1}{2}\end{bmatrix}\Big). \end{aligned} \tag{4.18}$$

Integration of (4.18) gives us the covariance

$$\mathbb{E}[w(t)w^{\mathrm{T}}(t)] = \begin{bmatrix}\frac{1}{2} & 0\\ 0 & \frac{1}{2}\end{bmatrix}t. \tag{4.19}$$

□

As a remark, note that $\lambda^{(3)}$ is set to 0. The edges associated with $\lambda^{(3)}$ are considered "superfluous" by the framework. We discuss this in more detail in the discussion.

Table 4.1: Mean and covariance for n Markov chains after time t

	n=100,000 t=50	n=100,000 t=500	n=1000 t=500
Mean	[-0.0078 0.0243]	[0.0556 -0.0253]	[0.1713 -0.135]
Covariance	diag[25.63,25.64]	diag[250.47,250.05]	diag[249.84,248.81]

4.4 Simulation and Discussion

In this section, we demonstrate the utility of our framework on the lattice shown in Figure 4.1. We set a to one and consider the random walk of a number of continuous-time Markov chains, whose infinitesimal generator is characterized by $\lambda^{(1)} = \frac{1}{6}$, $\lambda^{(2)} = \frac{1}{3}$ and $\lambda^{(3)} = 0$, with origin $x_{0,0} = (0, 0)$. The mean and covariance of the distribution of particles can be seen in Table 4.1. As in previous results, a higher number of employed particles results in less deviation of the mean and covariance from mean 0 and covariance $\frac{1}{2}tI_2$. The histogram of the positions of particles is shown in Figure 4.4, for $100,000$ particles at $t = 500$. We compare the histogram with a multivariate Gaussian distribution, in Figure 4.5, with mean 0 and covariance $\frac{1}{2}tI_2$. In terms of shape, height and width we get a good match between the histogram and the Gaussian bell curve. Generally, as in Chapter 3, the results show a very good approximation of the diffusion of heat.

With the subdivision of a lattice into smallest units and the classification of nodes within that unit into types as shown in Figure 4.2, we can extend the use of our framework to different classes of lattices. In the derivation of the mean and covariance, (4.16) and (4.18), each type of node has a mean of 0 and covariance of $\frac{1}{2}tI_2$ so that the overall mean and covariance is 0 and $\frac{1}{2}tI_2$. That is the reason why type two nodes, (nodes with no edge to the left, see Figure 4.2), have the transition rates $\lambda_{(2i,2j+1),(2i+1,2j)}$ and $\lambda_{(2i,2j+1),(2i,2j+2)}$ set to 0. They are not needed or "superfluous", so to say. The three remaining outgoing transition rates allow for a mean of 0 for this type of node, since particles can jump equally in each direction. The angle between each pair of those three nodes is 120 degrees. If we would want the two transition rates we set to 0 to be positive, we cannot find the desired mean and covariance without making assumptions about the mean and covariance

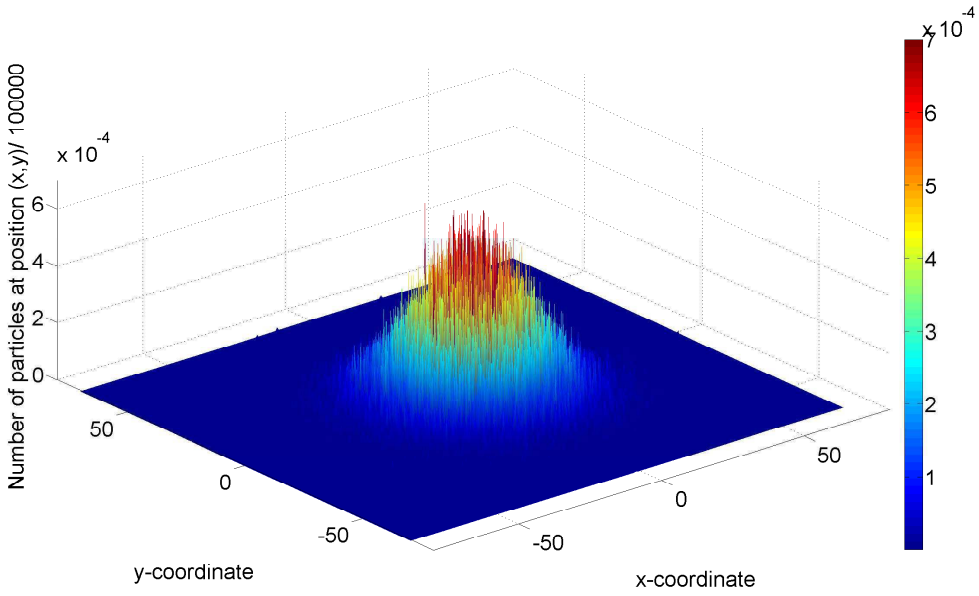

Figure 4.4: Distribution of 100,000 Markov chains at t=500

of each type of node within the smallest unit, other than being 0 and $\frac{1}{2}tI_2$. In other words, looking at type two nodes, if we want particles to jump to each of its five neighbors, we cannot find the desired mean and covariance, because an edge going to the left is "missing". Therefore, our framework gives us transition rates such that our requirements are satisfied. Note that although particles cannot jump from state $x_{2i,2j+1}$ to state $x_{2i+1,2j}$, jumps from state $x_{2i+1,2j}$ to state $x_{2i,2j+1}$ are possible.

These interesting results create the question; how we can build the configuration of neighbors for one type of node such that it allows for the usage of our framework? What are limitations and is a further extension feasible? The main statement is that, if we can retrieve the theoretical mean and covariance of the Gaussian distribution on each type of node, then we can retrieve the theoretical mean and covariance for the random variable of the Markov chain on the whole lattice. This may lead to even more irregular grids in two dimensions, with limitations as outlined above in this section. Applications might be grids in three dimensions which have been projected onto the plane. Chapter 7 looks into this topic in more detail.

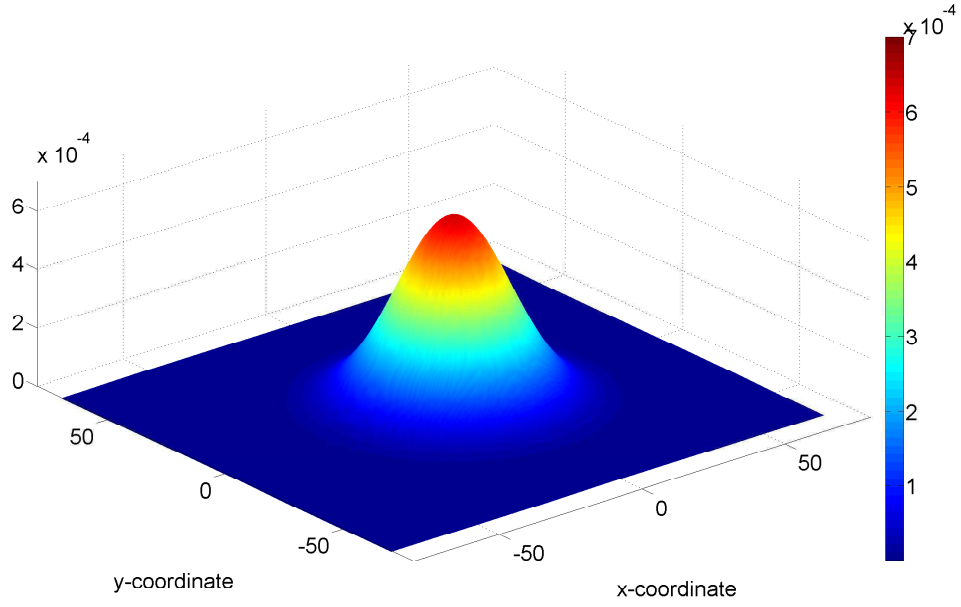

Figure 4.5: Multivariate Gaussian at t=500

Chapter 5

Heat Diffusion Modelling on Closed Curves

5.1 Introduction

In the last chapters we considered manifolds without boundary or curvature. In this chapter we consider the modelling of the diffusion of heat on the closed curve. To describe the diffusion of heat on the closed curve we have to deal with circular boundary conditions. This is visible in the definition of the Markov-chain-driven random walk which takes place on a finite set of states. Secondly, a closed curve has non-zero curvature so that the distance between two nodes on the closed curve is not necessarily the Euclidean distance. Instead, we use the geodesic distance as mentioned in Chapter 1.

5.2 Markov-Chain-Driven Random Walk on the Closed Curve

For the closed curve we deal with a finite set of states, hence the definition of the continuous-time Markov chain in Chapter 2 has to be slightly changed. Let $w(t)$ be the random variable of a continuous-time, time-homogeneous Markov chain on a finite set $X = \{x_0, x_1, x_2, \ldots, x_{N-2}, x_{N-1}\}$, where N is the number of nodes. We express the

conditional probability of the time-homogeneous, continuous-time Markov chain as

$$\mathbb{P}[w(s+t) = x_i | w(s) = x_j] = \mathbb{P}[w(t) = x_i | w(0) = x_j], \quad i, j = 0, \ldots, N-1, \tag{5.1}$$

for all $s \in \mathbb{R}_+$. We also define

$$p_{j,i}(t) \triangleq \mathbb{P}[w(t) = x_i | w(0) = x_j], \tag{5.2}$$

which are the elements of $P(t)$, where $P(t)$ is the transition probability matrix of the Markov chain with size $\mathbb{R}^{N \times N}$. We then assume nodes $\{s_0, s_1, s_2, \ldots, s_{N-2}, s_{N-1}\}$ are unequally distributed along the circumference of the closed curve as shown in the example, the circle with radius r, in Figure 5.1. The values of nodes s_i are the geodesic distance from s_0 in clock-wise direction. The constant N indicates the number of nodes on the closed line and equals 21 in Figure 5.1. On the lower half circle, the distance to the next node is $\frac{2\pi}{36}r$ and the distance from the previous node is $\frac{4\pi}{36}r$ if i is even. If i is odd then the distance to the next node is $\frac{4\pi}{36}r$ and the distance from the previous node is $\frac{2\pi}{36}r$. On the upper half circle, the distance between each pair of nodes is $\frac{4\pi}{36}r$.

Each particle is initially placed at node $s_0 = 0$, and then jumps between neighboring nodes along the geodesics. The state-transition diagram describing the successive random jumps of particles at random times is given in Figure 5.2. The transition rate for a jump from node s_i to node s_{i+1} is denoted as $\lambda_{i,i+1}$ and the transition rate from state s_i to state s_{i-1} is denoted as $\lambda_{i,i-1}$. Note that the behavior of the random walk is captured by a continuous-time Markov chain with infinitesimal generator Q with elements $\lambda_{i,j}$. Furthermore, note that the time derivative of (1.10) is given by

$$\frac{\mathrm{d}p_i(t)}{\mathrm{d}t} = (e^{Qt}Q)_{(0,i)}, \quad i = 0, \ldots, N-1. \tag{5.3}$$

From any given node, particles can jump only to the two neighboring nodes. The transition rates to the other nodes are therefore 0. Due to this structure of Q and the Kolmogorov forward equation, we can rewrite (5.3) as a combination of transition rates and transition

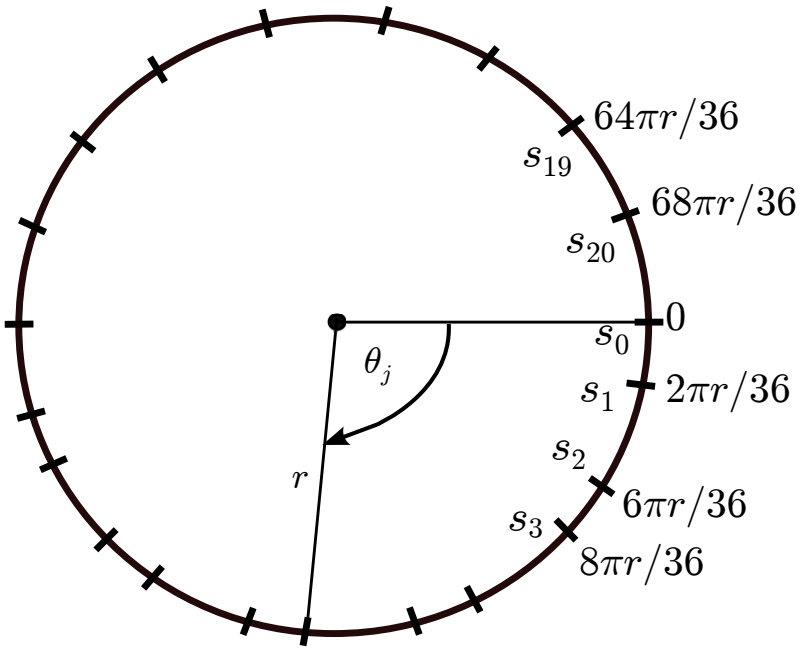

Figure 5.1: Example of 21 non-uniformly placed nodes on the circle

probabilities to each neighboring node

$$\frac{\mathrm{d}p_i(t)}{\mathrm{d}t} = \lambda_{i-1,i}p_{i-1}(t) - (\lambda_{i,i-1} + \lambda_{i,i+1})p_i(t) + \lambda_{i+1,i}p_{i+1}(t), \tag{5.4}$$

where the subscript N is understood as 0 for the circular boundary condition given by Figure 5.2. Note that on the infinite straight line the infinitesimal generator Q has a tridiagonal form. Since the particles in our case are on the closed curve it is possible to jump from s_{N-1} to s_0 and s_0 to s_{N-1}, resulting in a form of Q which incorporates the circular boundary conditions. The infinitesimal generator Q can also be seen as the matrix representation of the graph which consists of nodes which have been placed on the closed curve. In this case Q is called the Laplacian matrix or discrete Laplace-Beltrami operator [19], [44]. We want to know how given non-uniformly placed nodes imposed on a manifold determine the graph Laplacian and therefore the Markov chain such that the continuous Laplace-Beltrami operator is approximated, thus approximating the diffusion of heat on the closed curve [17].

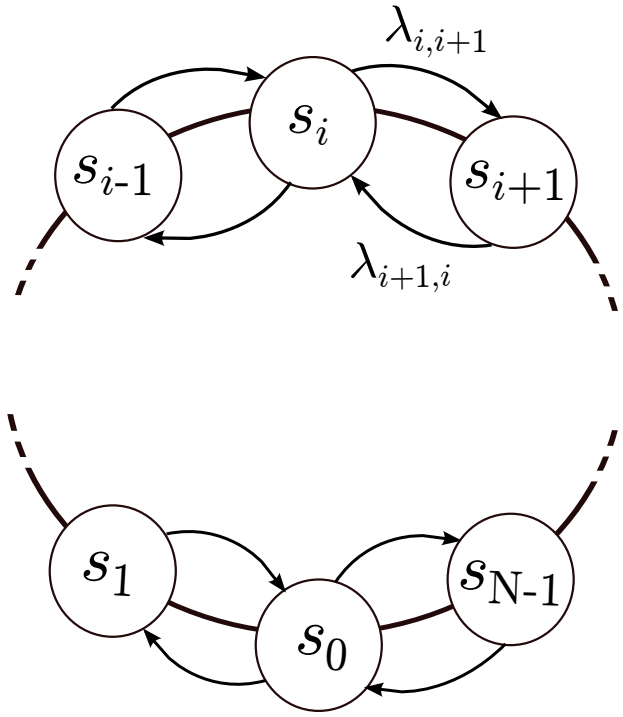

Figure 5.2: State-transition diagram of the continuous-time Markov chain on the example circle

5.3 Statistics on Closed Curves and the Heat Equation

The solution to the heat equation on the interval $s \in [0, L)$ with boundary conditions $u(0,t) = u(L,t)$ and with initial condition $u(s,0) = \delta(s)$ is given as [45]

$$u(s,t) = \frac{1}{L} + \frac{2}{L}\sum_{n=1}^{\infty} e^{-(\frac{2n\pi}{L})^2 \zeta t} \cos\Big(\frac{2\pi n}{L} s\Big), \tag{5.5}$$

where $u(s,t)$ denotes the temperature at position s and time t, ζ denotes the diffusion parameter, and L denotes the length of the closed curve. To relate (5.5) to a probability distribution we use the concept of wrapped distributions [46], [47]. In this approach, a distribution on the line is taken and wrapped around a closed curve where the length of the closed curve is determined by the metric g in (1.3). Note that the regular mean and variance as defined in previous chapters are not applicable to the closed curve. We consider the characteristic function and trigonometric moments instead, and show how the probability density function and the trigonometric moments of a random variable on a closed curve can be found and how the trigonometric moments lead to our main result. Furthermore, we show how the probability density function and the trigonometric

moments of the random variable on the closed curve relate to (5.5).

To this end, we define the random variable on the closed curve as $w_{\rm c}(t) \in [0, L)$. We denote the probability distribution of $w_{\rm c}(t)$ as $p_{\rm c}(s,t)$. Note that $p_{\rm c}(s,t)$ is given by (5.5), following the idea that the solution to the heat equation is also a Gaussian function, such as in previous chapters. Now, since every function in a bounded interval can be written as a Fourier series we can express $p_{\rm c}(s,t)$ as

$$p_{\rm c}(s,t) = \frac{1}{L} + \frac{2}{L}\sum_{k=1}^{\infty}\left(\alpha(k,t)\cos\left(\frac{2\pi k}{L}s\right) + \beta(k,t)\sin\left(\frac{2\pi k}{L}s\right)\right). \tag{5.6}$$

The functions α and β are the Fourier coefficients or trigonometric moments of the random variable $w_{\rm c}(t)$, which is distributed according to the wrapped Gaussian distribution [48–50]. The functions α and β are therefore given by (5.5) as

$$\begin{aligned}\alpha(k,t) &= \mathbb{E}[\cos k w_{\rm c}(t)] \\ &= e^{-(\frac{2\pi k}{L})^2\zeta t},\end{aligned} \tag{5.7}$$

and

$$\begin{aligned}\beta(k,t) &= \mathbb{E}[\sin k w_{\rm c}(t)] \\ &= 0.\end{aligned} \tag{5.8}$$

Note that α and β also determine the characteristic function of $w_{\rm c}(t)$ given by [49, 50]

$$\begin{aligned}\phi_{\rm c}(k,t) &= \mathbb{E}[e^{ikw_{\rm c}(t)}] = \int_0^L e^{iks}p_{\rm c}(s,t)\mathrm{d}s = \alpha(k,t) + i\beta(k,t) \\ &= e^{-(\frac{2\pi k}{L})^2\zeta t},\end{aligned} \tag{5.9}$$

where

$$\alpha(k,t) = \mathbb{E}[\cos k w_{\rm c}(t)], \tag{5.10}$$

$$\beta(k,t) = \mathbb{E}[\sin k w_{\rm c}(t)]. \tag{5.11}$$

The characteristic function is called the kth trigonometric moment of the random variable

$w_{\mathrm{c}}(t)$. The first trigonometric moment is usually used to check the similarity between distributions and is found by setting $k = 1$ such that

$$\phi_{\mathrm{c}}(1, t) = e^{-(\frac{2\pi}{L})^2 \zeta t}. \tag{5.12}$$

Equation (5.12) is the theoretical first trigonometric moment and it is used later in this chapter to show that the first trigonometric moment of the random variable of the continuous-time Markov chain is identical when the infinitesimal generator of the continuous-time Markov chain is characterized in a certain way.

As a remark, when considering the special case of the unit circle, we can define the mean resultant $R_{\mathrm{c}} \in [0, 1]$ and use it as an additional criterion for the validation of our results. The mean resultant determines the spread of a distribution on the circle and is an analogue to the variance on the infinite line [51]. To compute the mean resultant from a distribution of particles on the circle, the position of each particle on the circle is taken as a two-dimensional vector with the origin being the center of the circle. Using the Euler's formula $e^{i\theta} = \cos\theta + i \sin\theta$, where $i = \sqrt{-1}$, we can express the mean resultant vector of a set of particle positions on the circle as [52]

$$R_{\mathrm{c}} = \sum_j e^{i\theta_j}, \tag{5.13}$$

where $\theta_j \in [0, 2\pi)$. R_{c} is the sum of all vectors from the center of the circle to each particle position on the circle. The mean direction μ_{c} is the angle θ of (5.13). To find the circular variance σ_{c}^2 we simply subtract the magnitude of R_{c} from one so that

$$\sigma_{\mathrm{c}}^2 = 1 - |R_{\mathrm{c}}|. \tag{5.14}$$

When all particles are concentrated at $\theta = 0$ the circular variance corresponds to one. A uniform distribution of particles on the circle corresponds to a circular variance of 0. This notion is analogue to the diffusion of heat on the circle, where we have a point source of heat particles at time 0 and a diffusion of particles as time continues until an equilibrium is reached.

Now consider the Gaussian distribution on the infinite line given by

$$p_x = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{5.15}$$

where μ is the mean and σ^2 the variance. According to [48, 50, 51], we can wrap this distribution around the unit circle and retrieve a distribution that satisfies the heat equation on the unit circle given by

$$p(\theta, t) = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{k=1}^{\infty} e^{-\frac{k^2\sigma^2}{2}} \cos(\theta - \mu), \tag{5.16}$$

where the variance is set to $2\zeta t$. The variance σ^2 and μ are then related to σ_c^2 and μ_c by

$$\sigma_c^2 = 1 - e^{\frac{\sigma^2}{2}}, \tag{5.17}$$

and

$$\mu_c = \mu \pmod{2\pi}. \tag{5.18}$$

5.4 Heat Diffusion on a Circle

In this section we examine the special case of the heat equation on the circle with radius r. The derivation in this section serves as an example on how one can find the metric and Laplace-Beltrami operator defined on a manifold. The heat equation on the circle was previously defined in (1.1). Also, see references [53], [54] for a treatment of the heat equation on the circle. We also previously defined the Laplace-Beltrami operator in (1.2). We deal with the circle in $\mathbb{R}^2$, therefore the local coordinate is $x^1 = \theta$ with a parametrization of the circle with radius r given by

$$y_1 = r\cos\theta, \tag{5.19}$$

$$y_2 = r\sin\theta, \tag{5.20}$$

where $(y_1, y_2) \in \mathbb{R}^2$ are called the ambient coordinates. The metric g is then calculated as [55]

$$\begin{aligned} g &= \left\langle \left[\frac{\partial y_1}{\partial \theta} \frac{\partial y_2}{\partial \theta} \right], \left[\frac{\partial y_1}{\partial \theta} \frac{\partial y_2}{\partial \theta} \right] \right\rangle \\ &= r^2. \end{aligned} \tag{5.21}$$

Using (5.21) in (1.2) leads to the Laplace-Beltrami operator for the circle given by [56], [53]

$$\Delta_{S^1} = \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2}. \tag{5.22}$$

Note that θ on the circle is in the interval $[0, 2\pi)$ and $u(0, t) = u(2\pi, t)$. This equality expresses the circular boundary conditions for the heat equation. Now recall the infinitesimal generator Q. As mentioned, Q characterizes the random walk, governed by the Markov chain, by using functions of the distances between nodes on the circle, defined by the metric. In addition, Q incorporates the circular boundary conditions similar to (5.22) by allowing a transition from state s_{N-1} to state s_0 and from state s_0 to state s_{N-1}.

To solve the heat equation the Fourier series expansion of the solution to the heat equation on the infinite line is used [54]. We can think of this as wrapping the solution to the heat equation on the infinite line around the circle such that [53]

$$u(r\theta, t) = \frac{1}{2\pi r} + \frac{1}{\pi r} \sum_{n=1}^{\infty} e^{-n^2 \frac{\zeta}{r^2} t} \cos n\theta. \tag{5.23}$$

Note that we can obtain (5.23) by substituting $2\pi r$ for L and θr for s in (5.5). The corresponding first moment of the random variable on the circle is then

$$\phi_{\mathrm{c}}(1, t) = e^{-\frac{1}{r^2} \zeta t}. \tag{5.24}$$

As mentioned in the introduction, the parameter ζ describes how fast heat diffuses, in other words, how far particles diffused for a given time. In case of circular boundary conditions the diffusion parameter determines how fast an equilibrium distribution is reached, or in other words, how long it takes until the temperature on the circle is everywhere the same.

5.5 Characteristic Function of the Continuous-Time Markov Chain

In this section we show how we can characterize the continuous-time Markov chain on the closed curve, such that the first trigonometric moment of the random variable of the Markov chain is guaranteed to be $e^{-(\frac{2\pi}{L})^2\zeta t}$, so that the probability density function of the random variable of the Markov chain satisfies the heat equation on the closed curve. We show this by choosing the transition rates of the continuous-time Markov chain as a function of node positions on the closed curve. Without loss of generality, we set the diffusion parameter ζ to $\frac{1}{2}$. For the statement of the following result, let $\overline{s}_j \triangleq s_j \frac{2\pi}{L}$.

Theorem 5.1. Let $w_{\mathrm{MC}}(t)$ be a random variable of a continuous-time Markov chain on a closed curve whose state-transition diagram is given by Figure 5.2. Furthermore, let $w_{\mathrm{MC}}(0) = 0$ and let

$$\lambda_{j,j+1} \triangleq \frac{\sin(\overline{s}_{j-1} - \overline{s}_j)}{2\gamma}\Big(\frac{2\pi}{L}\Big)^2, \tag{5.25}$$

$$\lambda_{j,j-1} \triangleq \frac{\sin(\overline{s}_j - \overline{s}_{j+1})}{2\gamma}\Big(\frac{2\pi}{L}\Big)^2, \tag{5.26}$$

where

$$\gamma = \sin(\overline{s}_{j-1} - \overline{s}_j) - \sin(\overline{s}_{j-1} - \overline{s}_{j+1}) + \sin(\overline{s}_j - \overline{s}_{j+1}).$$

Then for any time $t > 0$, the random variable $w_{\mathrm{MC}}(t)$ has the first moment of $e^{-(\frac{2\pi}{L})^2 \frac{t}{2}}$.

Proof The first trigonometric moment of the random variable $w_{\mathrm{MC}}(t)$ is given by

$$\begin{aligned} \phi_{\mathrm{MC}}(1,t) &= \mathbb{E}[e^{iw_{\mathrm{MC}}(t)}] \\ &= \sum_{j=0}^{N-1} e^{i\overline{s}_i} p_i(t). \end{aligned} \tag{5.27}$$

The random variable $w_{\mathrm{MC}}(t)$ takes only values $\overline{s}_i$. Note that $p_i(t)$ is the transition probability of being in state $\overline{s}_i$ at time t and not the probability density function. In order to

find the transition rates we take the time derivative of the first trigonometric moment

$$\frac{\mathrm{d}\phi_{\mathrm{MC}}(1,t)}{\mathrm{d}t} = \sum_{i=0}^{N-1} e^{i\overline{s}_i}\frac{\mathrm{d}p_i(t)}{\mathrm{d}t}. \tag{5.28}$$

Using (5.4) we can rewrite the expression (5.28) as

$$\begin{aligned}\frac{\mathrm{d}\phi_{\mathrm{MC}}(1,t)}{\mathrm{d}t} &= \sum_{i=0}^{N-1} e^{i\overline{s}_i}\Big(\lambda_{i-1,i}p_{i-1}(t) - (\lambda_{i,i-1}+\lambda_{i,i+1})p_i(t) + \lambda_{i+1,i}p_{i+1}(t)\Big)\\ &= \sum_{i=0}^{N-1} e^{i\overline{s}_i}\lambda_{i-1,i}p_{i-1}(t) - \sum_{i=0}^{N-1} e^{i\overline{s}_i}(\lambda_{i,i-1}+\lambda_{i,i+1})p_i(t)\\ &\quad + \sum_{i=0}^{N-1} e^{i\overline{s}_i}\lambda_{i+1,i}p_{i+1}(t).\end{aligned} \tag{5.29}$$

We now use the substitution $\bar{i} = i-1$ in the first term of (5.29), $\bar{i} = i$ in the second term and $\bar{i} = i+1$ in the third term to get

$$\begin{aligned}\frac{\mathrm{d}\phi_{\mathrm{MC}}(1,t)}{\mathrm{d}t} &= \sum_{\bar{i}=-1}^{N-2} e^{i\overline{s}_{\bar{i}+1}}\lambda_{\bar{i},\bar{i}+1}p_{\bar{i}}(t) - \sum_{\bar{i}=0}^{N-1} e^{i\overline{s}_{\bar{i}}}(\lambda_{\bar{i},\bar{i}-1}+\lambda_{\bar{i},\bar{i}+1})p_{\bar{i}}(t)\\ &\quad + \sum_{\bar{i}=1}^{N} e^{i\overline{s}_{\bar{i}-1}}\lambda_{\bar{i},\bar{i}-1}p_{\bar{i}}(t).\end{aligned} \tag{5.30}$$

Since the index -1 in the first term of (5.30) is actually the index $N-1$ and since the index N in the third term of (5.30) is the index 0, we rename $\bar{i}$ to i and write

$$\frac{\mathrm{d}\phi_{\mathrm{MC}}(1,t)}{\mathrm{d}t} = \sum_{i=0}^{N-1} p_i(t)\Big(e^{i\overline{s}_{i+1}}\lambda_{i,i+1} - e^{i\overline{s}_i}(\lambda_{i,i-1}+\lambda_{i,i+1}) + e^{i\overline{s}_{i-1}}\lambda_{i,i-1}\Big). \tag{5.31}$$

Using the Euler's formula we write (5.31) as a combination of sine and cosine function and after substituting the transition rates in (5.25) and (5.26) we can rewrite (5.31) as

$$\frac{\mathrm{d}\phi_{\mathrm{MC}}(1,t)}{\mathrm{d}t} = -\frac{1}{2}\Big(\frac{2\pi}{L}\Big)^2\phi_{\mathrm{MC}}(1,t). \tag{5.32}$$

We can then solve this differential equation with the condition $\phi_{\mathrm{MC}}(1,0) = 1$ and obtain the first trigonometric moment of the random variable of the continuous-time Markov

chain as

$$\phi_{\mathrm{MC}(1,t)} = e^{-(\frac{2\pi}{L})^2 \frac{t}{2}} \tag{5.33}$$

In other words, (5.25) and (5.26) represent the transition rates that lead to a characterization of the continuous-time Markov chain such that the first trigonometric moment of $e^{-(\frac{2\pi}{L})^2 \frac{t}{2}}$ is guaranteed. □

Note that the difference between each pair of nodes $\overline{s}_i$ has to be smaller than π in order to obtain positive transition rates. We can substitute the node positions on the closed curve and substitute these distances into (5.25) and (5.26), which yield the transition rates for a particle to jump from any node on the closed curve to its neighbors. Note that to determine the transition rate to one neighbor not only requires the position of that neighbor but also the position of the second neighbor. It is also important to note that the particles jump along the geodesic distances and not the Euclidean distances, leading to more accurate results on some geometries. In the Chapter 2 we presented the transition rates which guarantee that the distribution of the continuous-time Markov chain on the infinite straight line has the same mean and variance as the Gaussian distribution on the infinite straight line. The transition rates were given by

$$\lambda^{\ell}_{i,i+1} \triangleq \frac{1}{(x_{i+1} - x_i)(x_{i+1} - x_{i-1})}, \tag{5.34}$$

$$\lambda^{\ell}_{i,i-1} \triangleq \frac{1}{(x_i - x_{i-1})(x_{i+1} - x_{i-1})}. \tag{5.35}$$

The variables x_i denote the Euclidean distance from the origin. The metric g is always 1, i.e., no curvature. Furthermore, these transition rates do not incorporate any boundary conditions. The transition rates in (5.25) and (5.26) on the other hand incorporate the circular boundary conditions by using the sine of the distance between neighbors, and the interval L.

5.6 Simulation and Discussion

The first example in this section demonstrates the utility of our framework on the grid given in Figure 5.1, where the radius r is set to 1. This grid consists of 21 nodes, non-

Table 5.1: Mean direction in radians and variance for 10,000 Markov chains after time t, using Gaussian distribution or simulated distribution by using geodesic or Euclidean distances between nodes

		Gaussian	**Geodesic**	**Euclidean**
t**=1**	Mean	0	0.0935	0.041
	Variance	0.397	0.389	0.43
t**=3**	Mean	0	0.013	0.087
	Variance	0.771	0.779	0.789
t**=8**	Mean	0	0.081	0.101
	Variance	0.9773	0.9765	0.9883

equally placed on the circle such that

$$s_i = \frac{3(i-1)\pi}{36} + \frac{2\pi}{36}, \quad i \quad \text{odd}, \tag{5.36}$$

$$s_i = \frac{3(i-1)\pi}{36} + \frac{3\pi}{36}, \quad i \quad \text{even}, \tag{5.37}$$

on the lower half circle and

$$s_i = \frac{3(i-1)\pi}{36} + \frac{3\pi}{36}, \tag{5.38}$$

on the upper half circle. Thus, we consider random walks which obey the Markov chain with infinitesimal generator characterized by the transition rates in (5.25) and (5.26) by $\lambda_{i,i+1} = 10.91$ and $\lambda_{i,i-1} = 5.54$ for even i and $\lambda_{i,i+1} = 10.91$ and $\lambda_{i,i-1} = 5.54$ for odd i on the lower half circle. On the upper half circle the infinitesimal generator is characterized by the transition rates $\lambda_{i,i-1} = \lambda_{i,i+1} = 4.15$ for all i. In the example s_{12} and s_0 are considered to be on the lower half circle. In the simulation 10,000 Markov chain are employed where each Markov chain starts at the origin $w_{\mathrm{MC}}(0) = 0$. The rose histograms with bin size 45 degrees in Figures 5.3–5.5 show the circular distribution of particle positions after $t = 1$, $t = 3$ and $t = 8$. As t increases we can see how particles diffuse symmetrically from the point source of heat particles at position 0 until an almost uniform distribution is reached at $t = 8$. We now use the previously defined mean resultant R_{c} to find the circular mean and the circular variance. The circular mean and circular variance serve as another criterion for the validation of our results. Column one of Table 5.1 shows the theoretical circular mean

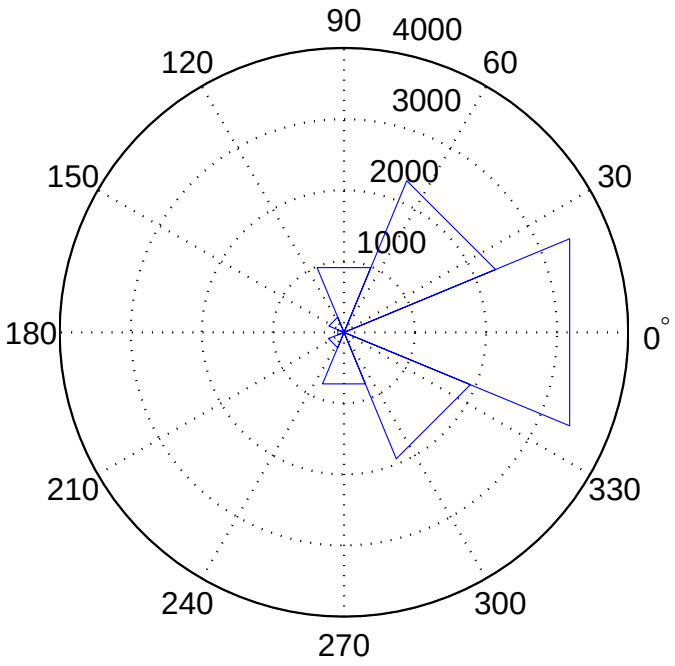

Figure 5.3: Distribution of 10,000 Markov chains at t=1

and the theoretical circular variance obtained from (5.18) and (5.17), where the Gaussian distribution on the infinite straight line has a of mean 0 and a variance of t. This column represents the theoretical values for different times. The second column shows the values for the circular mean and the circular variance at time t, for a distribution which was obtained by simulating $10,000$ Markov chains with our framework, where the transition rates were obtained by using the geodesic distance between nodes on the circle. Column three shows the circular mean and the circular variance at time t for a distribution which was obtained by simulating $10,000$ Markov chains with our framework, where the distances between nodes on the circle were not the geodesic distances but Euclidean distances, i.e., straight lines. The deviation between values from the first and second shows a good approximation of our framework for heat diffusion on the circle. Comparing columns one and three shows that, when using Euclidean distances between nodes, we get also good results, with a slightly bigger deviation from the theoretical values than column two.

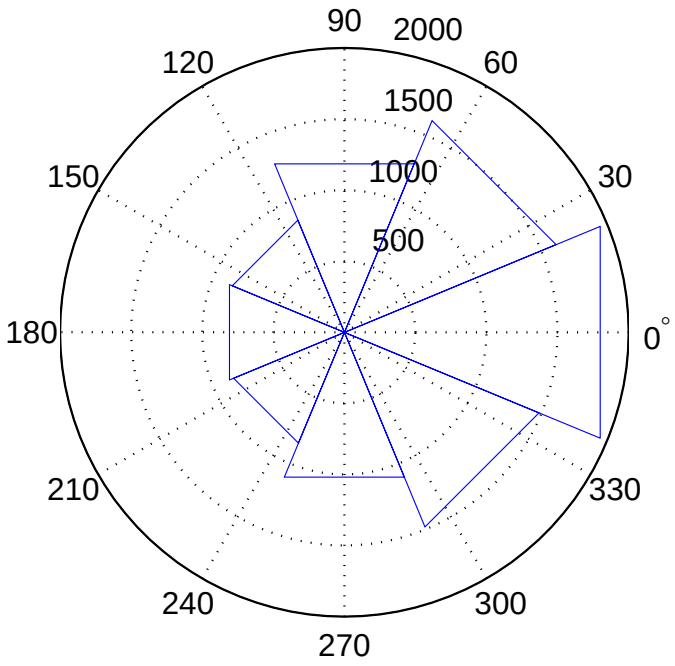

Figure 5.4: Distribution of 10,000 Markov chains at t=3

The second example we chose to discuss is the ellipse. We parametrize the ellipse as

$$y_1 = a\cos\theta, \tag{5.39}$$

$$y_2 = b\sin\theta, \tag{5.40}$$

where a is the length of the major axis and b is the length of the minor axis. To calculate the metric g for the ellipse we use (5.21) and obtain

$$g = a^2\cos^2\theta + b^2\sin^2\theta. \tag{5.41}$$

Using (5.41) in (1.2) yields the Laplace-Beltrami operator for the ellipse

$$\nabla^2 = \frac{1}{a^2\sin^2\theta + b^2\cos^2\theta}\frac{\partial^2}{\partial\theta^2}. \tag{5.42}$$

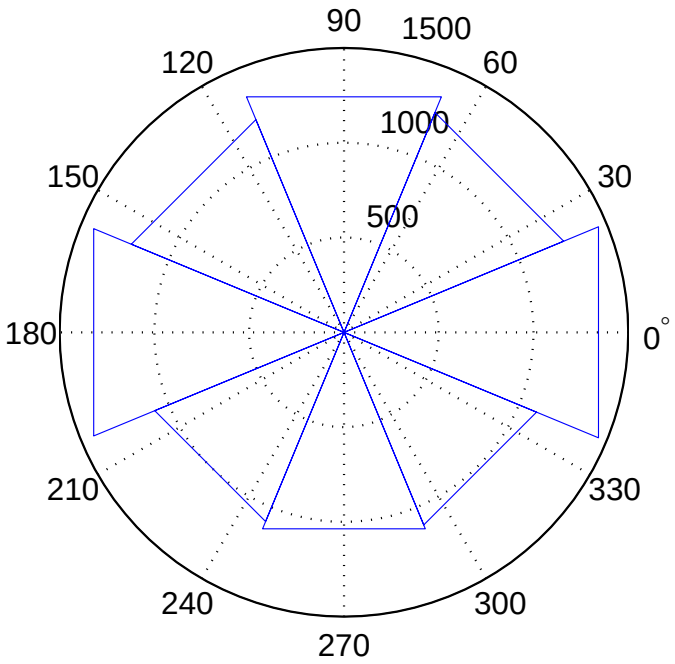

Figure 5.5: Distribution of 10,000 Markov chains at t=8

We are given the nodes Figure 5.6, where the nodes are distributed as

$$s_i = i\frac{\pi}{12}, \quad i = 0, 1, \ldots, 12, \tag{5.43}$$

$$s_i = i\frac{\pi}{24}, \quad i = 13, 14, \ldots, 35, \tag{5.44}$$

In the example we set a to 1, b to $\frac{1}{2}$ and distribute the nodes by dividing the lower half of the ellipse into 12 segments and the upper half into 24 segments. Even though the angles are equally distributed on each half of the ellipse, the distances between nodes are irregular. Overall we have 36 nodes resulting in 36 different transition rates.

Figures 5.7–5.9 show the histograms of the particle position at $t = 1$, $t = 3$ and $t = 6$, respectively, where we divide the circumference into 6 bins and count the particles in each bin at the three time instants. We can see that as time goes on, the distribution approaches an equilibrium distribution where the particles are equally distributed around the ellipse. In addition, the Gaussian distributions, indicating the distribution of heat on the circumference of the ellipse at a given time, were included to show that we receive a

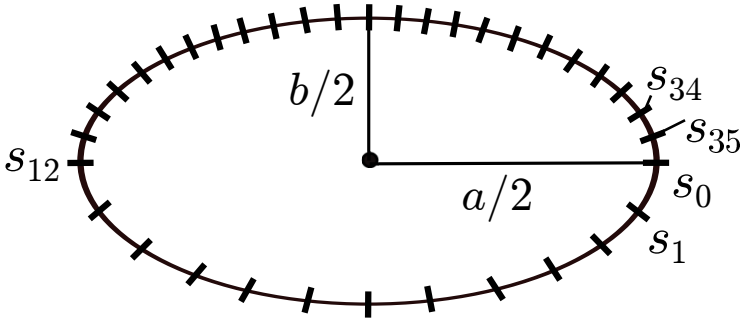

Figure 5.6: Example of 36 non-uniformly placed nodes on the ellipse

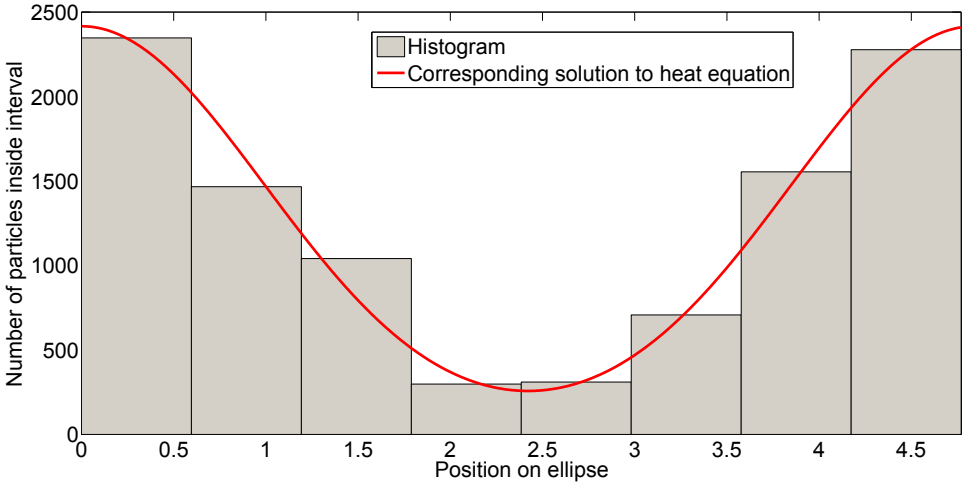

Figure 5.7: Distribution of 10,000 Markov chains at t=1 with corresponding solution to the heat equation on the ellipse

good approximation of the heat distribution on the ellipse.

Note that if we wanted to know a more detailed distribution of particles, i.e., smaller bin size in the histograms, we would increase the number of nodes accordingly. This might be desirable if more information is needed in a line segment with high curvature. In case we only want to know the amount of particles in a big segment of the ellipse, we can use less nodes. Compared to frameworks which only offer the use of a regular grid, our framework offer a real advantage when numerically approximating the heat distribution on a closed curve. Furthermore, note that the diffusion of heat can be approximated on any closed curve with different parameterizations. If the curvature, i.e., the metric, is not known, we can still utilize straight line approximations if the position of the nodes is known. Lastly we can conduct the simulations by utilizing parallel computing techniques, resulting in less computation time and decreased computational load.

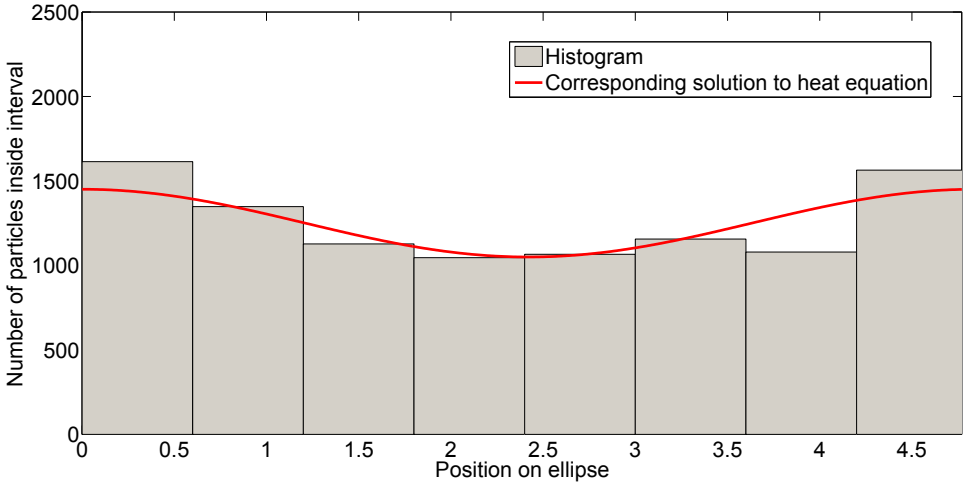

Figure 5.8: Distribution of 10,000 Markov chains at t=3 with corresponding solution to the heat equation on the ellipse

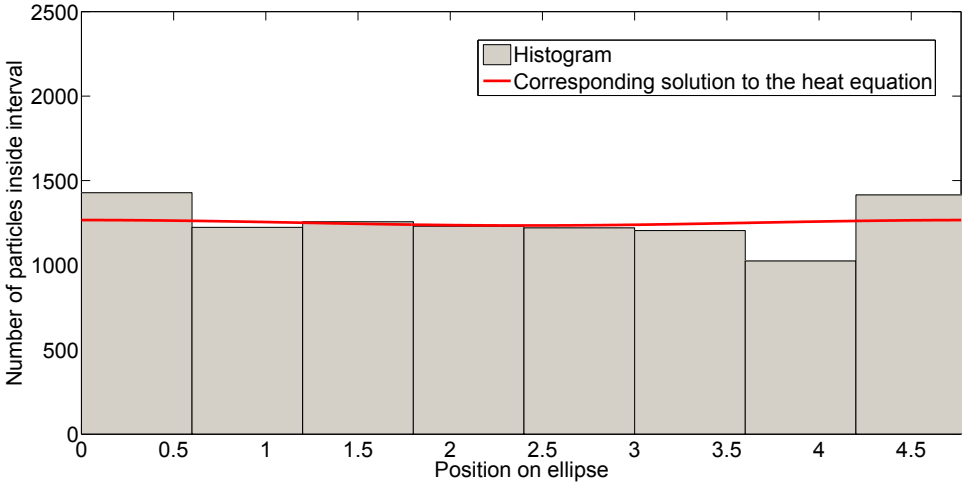

Figure 5.9: Distribution of 10,000 Markov chains at t=6 with corresponding solution to the heat equation on the ellipse

Chapter 6

Empirical Study on the Sphere

6.1 Introduction

In this chapter we discuss the extension of our framework to the S^2-sphere. The solution to the heat equation on the S^2-sphere is known, and we can also describe the probability that a Brownian particle is at a certain position on the S^2-sphere at a give time. Similarly to the closed curve, one would assume that it is possible to wrap the Gaussian distribution on the plane around the S^2-sphere. Unfortunately we can not wrap the two-dimensional Gaussian distribution around the S^2-sphere. Another problem with this approach is the rigorous definition of a grid on the S^2-sphere. As shown in Chapter 5, we can wrap nodes on the infinite straight line around the closed curve. On the other hand, we cannot simply wrap the a lattice on the plane around the S^2-sphere in the same manner. The reason is as follows. Assume that we choose a node on the plane and choose another node whose position is given by adding a vector with magnitude $2\pi r$ to the position of the first node. If we obtain many of these nodes by adding k vectors with the same direction and with magnitude $2\pi r$ to the position of the first node, where $k \in \mathbb{I}$, then after wrapping the planar grid around the sphere, the nodes obtained on the plane should correspond to the same node on the sphere, just as on the circle. However this is not the case. This can be illustrated by trying to wrap a piece of paper around a globe. It is not possible to do this without cutting holes into the paper. Circular statistics provides the tools to define a probability distribution on the S^2-sphere [46, 47]. The problem is that these probability distributions are not Gaussian functions and have no similarity to the solution to the heat

equation, as we have seen in previous chapters.

The method we use to still approximate the diffusion of heat on the sphere is the projection of nodes onto the plane. By using the smallest unit approach we can categorize each node on the sphere as a type and then project each type of node onto the plane and compute the transition rates. Using these transition rates we can then conduct simulation and compare the result by comparing the mean and variance of the distribution of particles and the theoretical mean and variance computed from the probability density function of a Brownian particle, which is also the solution to the heat equation on the sphere.

6.2 Statistics and the Diffusion of Heat on the S^2-Sphere

The S^2-sphere is commonly parametrized with the latitude and longitude, $\theta \in [0, \pi)$ and $\varphi \in [0, 2\pi)$, and the radius $r \in \mathbb{R}$. The solution to the heat equation with a point source of heat at the north pole, $(\theta = 0)$, on the sphere is well known and expressed as [57]

$$u(\theta, t) = \frac{1}{4\pi} \sum_{l=1}^{\infty} (2l+1) P_l(\cos\theta) e^{-l(l+1)\frac{1}{2}\zeta t}, \tag{6.1}$$

where P_l is Legendre polynomial of order l. Note that there is no ϕ in (6.1). The reason is the symmetric diffusion of heat from the north pole, which is why the temperature at a given time t does not change on the same latitude θ. The probability density function of a Brownian particle on the S^2-sphere is given as [58]

$$p_{\mathrm{m}}(\theta) = \frac{\sin\theta}{2} \sum_{l=1}^{\infty} (2l+1) P_l(\cos\theta) e^{-l(l+1)\frac{1}{2}\sigma^2}. \tag{6.2}$$

The parameter σ^2 denotes the variance of the single variable θ. In many sources (6.2) is considered to be the analogue to the Gaussian distribution on the plane [59, 60]. That is the reason why we use (6.2) to find the theoretical mean and variance of θ. As in previous chapters we can then use the variance and mean of the variable θ in (6.2) as a means to compare (6.2) to the distribution obtained from simulating continuous-time Markov chains on the sphere, where we compute the sample mean of the angle θ of particle positions. To

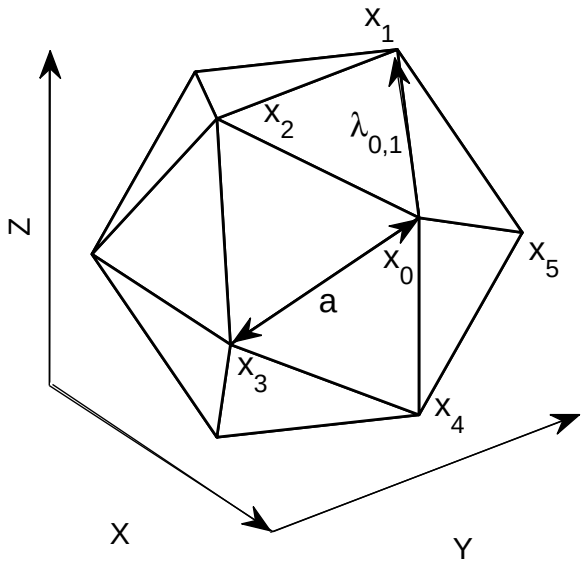

Figure 6.1: Example grid of an icosahedron with 12 nodes

find the theoretical mean and variance of θ we use the well known expressions for mean

$$\mu_\theta = \int_0^\pi P(\theta, t)\theta \mathrm{d}\theta, \tag{6.3}$$

and the variance

$$\sigma_\theta^2 = \int_0^\pi P(\theta, t)(\theta - \mu)^2 \mathrm{d}\theta. \tag{6.4}$$

6.3 Heat Diffusion Modelling on the Unit S^2-Sphere

In this section we present an empirical way for finding the transition rates for a Markov chain-driven random walk on the S^2-sphere. The grid we are considering is an icosahedron given in Figure 6.1. The icosahedron has 12 nodes and 20 faces and it is regular in the sense that all faces are equilateral triangles. The nodes of the icosahedron lie on the unit sphere that encloses the icosahedron and are evenly spaced. We can then take a node and project all neighboring nodes onto the tangent plane of the node in question. There are numerous projection schemes available [61]. The projection scheme we use is similar to the equal-area projection, (see Figure 6.2). For the icosahedron we simply assume that

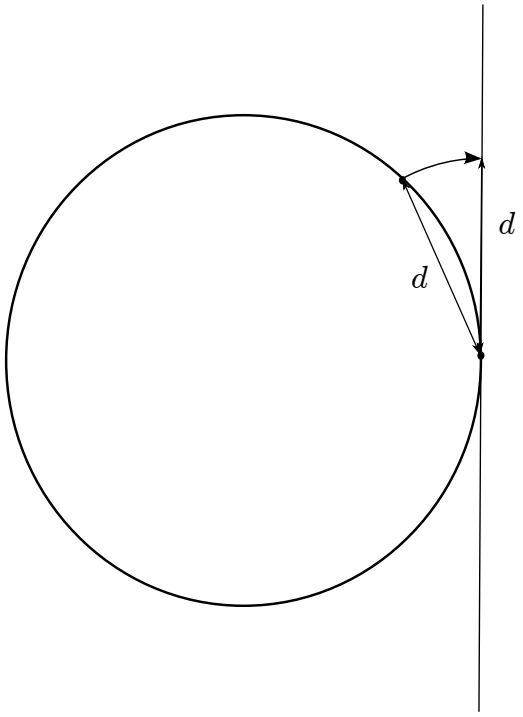

Figure 6.2: Equal-area projection

the angle between each pair of adjacent edges is 72 degree and that the length of each edge corresponds to the length a on the sphere. Note that this is a naive approach to project a grid on the sphere onto a grid on the plane. We can not actually create a regular lattice on the plane, where each node has five neighbors and where all edges have same length. However, with the assumptions made above, we can approximate the actual node and the connected edges on the sphere. We essentially use an unstructured lattice on the plane. For the example of the icosahedron we have only one type of node, given in Figure 6.3. With this configuration we can calculate the transition rate as

$$\lambda_{\mathrm{s}} = \frac{1}{5a^2}. \tag{6.5}$$

The variable a is the length of each edge on the icosahedron, which is why we have only one transition rate.

We then simulate $10,000$ Markov chains on the unit sphere, using the icosahedron grid. To determine the sample mean and sample variance of the angle θ of particle positions with

$$\mu_{\mathrm{s}} = \frac{1}{N}\sum_{i=1}^{N}\theta_i, \tag{6.6}$$

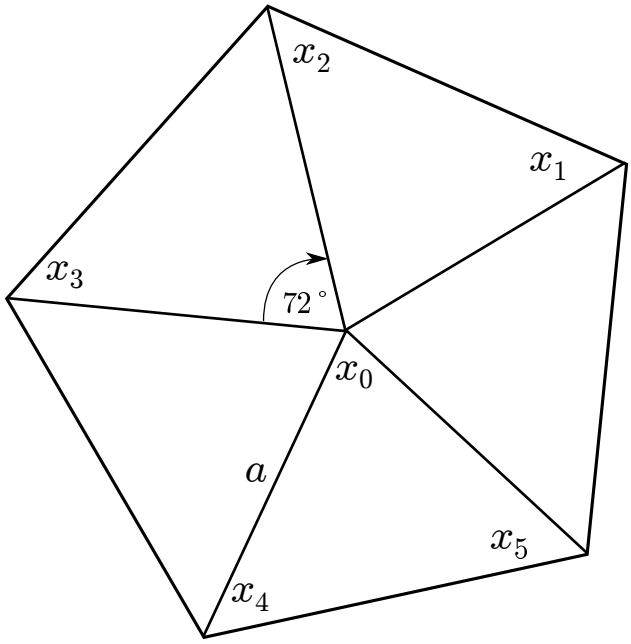

Figure 6.3: Icosahedron stencil projected onto the plane

for the sample mean and

$$\sigma_{\mathrm{s}}^2 = \frac{1}{N}\sum_{i=1}^{N}(\theta_i - \mu_{\mathrm{s}})^2, \tag{6.7}$$

for the sample variance.

Figures 6.4 and 6.5 show the theoretical mean and variance compared to the sample mean and sample variance as a function of t. We can see a good approximation of the theoretical values by the values obtained from simulations.

One conjecture in previous chapters was that we obtain more accurate results if more grid nodes are used. Assume we have a grid on the sphere with overall 168 nodes as shown in Figure 6.6. This type of grid is commonly created by taking an icosahedron and then subdividing each triangle into four new equal triangles by placing new nodes on each edge. The grid in Figure 6.6 is called spherical geodesic grid and has some nodes with six neighbors and some nodes with five neighbors. Note that the number of nodes with five neighbors is always 12, regardless of how many times we subdivided each triangle.

Overall the grid has five types of nodes. With the procedure explained previously in this chapter and the ideas in Chapter 4 we can again find the transition rates for this particular grid. We then simulate $10,000$ particles on the grid given in Figure 6.6 and compute the sample mean and sample variance of the angle θ of particle positions.

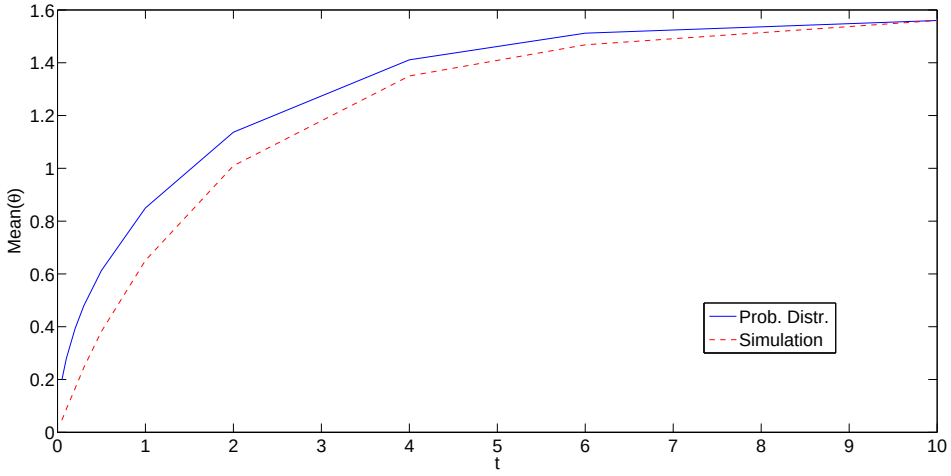

Figure 6.4: Sample mean of θ of $10,000$ particles on the icosahedron

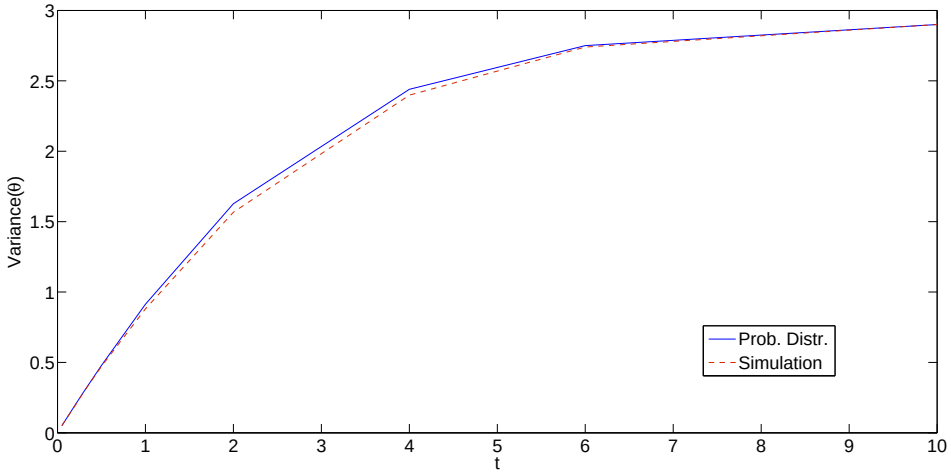

Figure 6.5: Sample variance of θ of $10,000$ particles on the icosahedron

Figures 6.7 and 6.8 show the comparison between the theoretical mean and variance of the angle θ and the sample mean and sample variance from the simulated distribution. This time we see a slightly better match between the two curves due to the higher amount of particles. As the number of nodes increases, the difference between a node and the edges connected to it on the sphere and the projected node with the connected edges on the plane becomes smaller, thus decreasing the error between theoretical mean and variance and simulated mean and variance. Note that for grids on the sphere with more than 12 nodes, the usage of non-equally placed nodes is necessary.

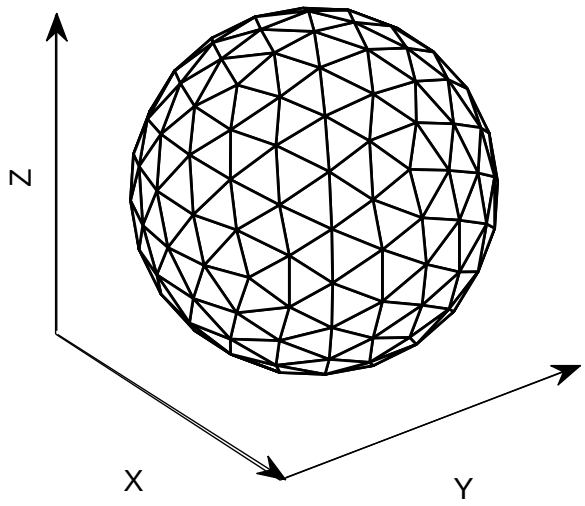

Figure 6.6: Example of a spherical geodesic grid with 168 nodes

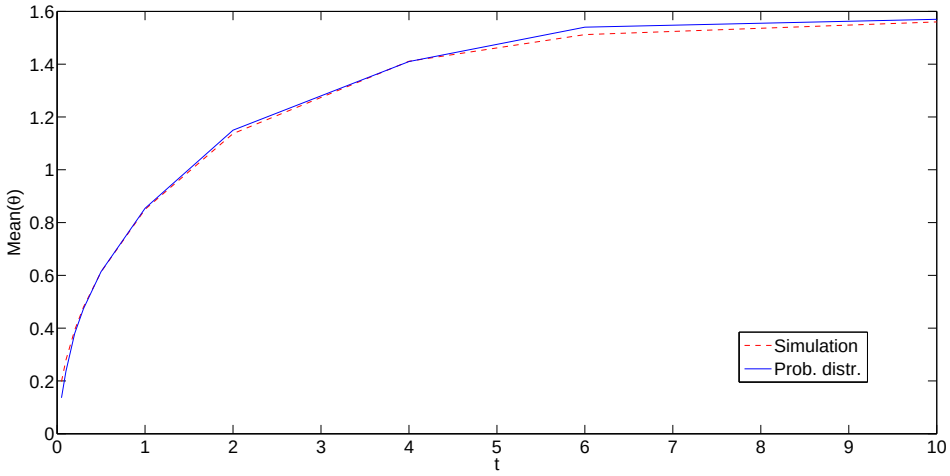

Figure 6.7: Sample mean of θ of $10,000$ particles on a spherical geodesic grid with 168 nodes

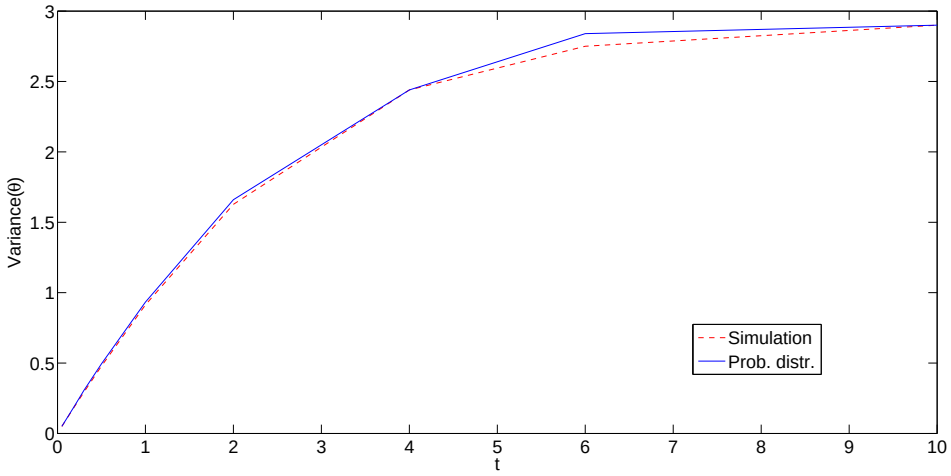

Figure 6.8: Sample variance of θ of $10,000$ particles on a spherical geodesic grid with 168 nodes

Chapter 7

Conclusion and Future Directions

In this thesis we developed a framework for the approximation of the diffusion of heat on several manifolds. When considering the microscopic random movement of heat particles, we showed in stochastic simulations how a multitude of particles, modelled with continuous-time Markov chains with a certain characterization of their infinitesimal generator, diffuses. Naturally, the position of particles is different each time we run the simulation, including some outlying particles whose position is farer away from the main group of particles. The mean distribution of those particle positions delivers a good approximation of the distribution of heat. This has been shown by comparing the mean and variance of the distribution of particles with the mean and variance of the Gaussian distribution. The comparison between the mean distribution of particle positions and the analytical heat distribution takes place on a macroscopic level. This comparison can be viewed as the comparison between the solution to the Kolmogorov forward equation and the analytical heat distribution. Recall that elements of the infinitesimal generator utilized in the Kolmogorov forward equation are the transition rates which describe the transition probability per time unit. Keeping this relation in mind, we provided the proofs which explain how a particular choice of transition rates leads to guaranteed statistical properties of the random variable of the Markov chain on several manifolds. Specifically, we interpreted the solution to the heat equation as a Gaussian distribution and obtained the theoretical mean and variance from this Gaussian distribution, which satisfies the heat equation. The proofs show that the random variable of the continuous-time Markov chain has indeed the same mean and variance if we choose the transition rates as a function of

the distances and angles between nodes. The important point is that the nodes do not need to be placed uniformly, i.e., different distances and angles between each pair of nodes on the manifold. Furthermore, even though we use the continuous Gaussian distribution to find the theoretical mean and variance, we could still show that the random variable of the random walk on a discrete state space has the same statistical properties.

In the Chapter 2 we considered the infinite straight line as an example manifold. The transition rates were given as

$$\lambda_{i,i+1} \triangleq \frac{1}{(x_{i+1} - x_i)(x_{i+1} - x_{i-1})}, \tag{7.1}$$

$$\lambda_{i,i-1} \triangleq \frac{1}{(x_i - x_{i-1})(x_{i+1} - x_{i-1})}, \tag{7.2}$$

and are simply functions of the distances from a given node to the next node and the distance from a given node to the previous node. There are no limitations for the arrangement of nodes, we can always find the transition rates.

In Chapter 3, we extended the framework from Chapter 2 for the plane. For the plane we presented the transition rates for a certain class of lattices. It turned out that there are limitation for the placement of lattice nodes on a manifold. The transition rates were given as

$$\begin{aligned} \lambda_{(i,j),(i+1,j)} &= \lambda_{(i,j),(i-1,j)} \triangleq \lambda^{(1)}, \\ \lambda_{(i,j),(i,j-1)} &= \lambda_{(i,j),(i-1,j+1)} = \lambda_{(i,j),(i+1,j-1)} \\ &= \lambda_{(i,j),(i,j+1)} \triangleq \lambda^{(2)}, \end{aligned} \tag{7.3}$$

where

$$\lambda^{(1)} \triangleq \frac{2b^2 - a^2}{2a^2(4b^2 - a^2)}, \tag{7.4}$$

$$\lambda^{(2)} \triangleq \frac{1}{2(4b^2 - a^2)}. \tag{7.5}$$

The limitation $\sqrt{2}b > a$ translates to an angle of at least 45 degrees between edges with length a and edges with length b (see Figure 7.1). The limitation illustrated in Figure 7.1 means that the lattice can be contracted and stretched only to a certain degree.

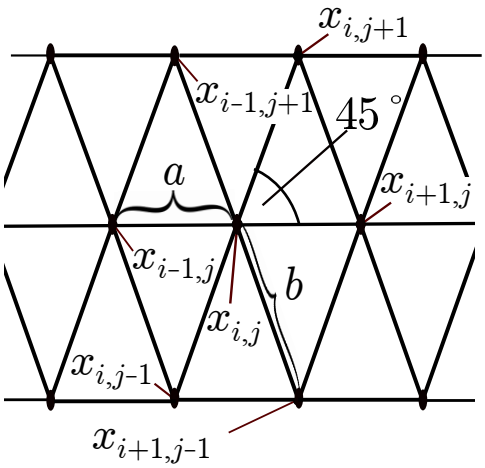

Figure 7.1: Limitation on the triangular lattice from Chapter 3

This limitation can be related to physical heat diffusion which we view as the random movement of heat particles to each direction. If there is no possibility for a particle to jump to a specific direction due to the lack of connected nodes in that direction, a symmetric diffusion of heat cannot happen, thus the theoretical mean and covariance of particle positions cannot be achieved. Figure 7.2 illustrates a type of lattice where this is the case. For the node $x_{i,j}$ particle have no possibility to jump to the right, thus it is not possible to find transition rates for this type of node.

Chapter 4 alleviates this issue somewhat with the introduction of the smallest unit approach. We showed how our framework can be used on lattices with missing edges. Interestingly, the framework will set the transition rates of "superfluous" edges to 0, to ensure a symmetric diffusion in all direction, or in other words, to ensure that the random variable of the continuous-time Markov chain has the same mean and covariance matrix as the Gaussian distribution.

In the Chapter 5 we introduced circular boundary conditions and curved spaces by considering heat diffusion modelling on closed curves. We gave an explanation of the wrapped Gaussian distribution and circular statistics. These two tools are necessary since the previously used notions of mean and variance are not applicable on closed curves. Instead, we use the characteristic function of the random variable of the wrapped Gaussian distribution, where the characteristic function also describes the trigonometric moments of the random variable, to compare it to the characteristic function of the random variable

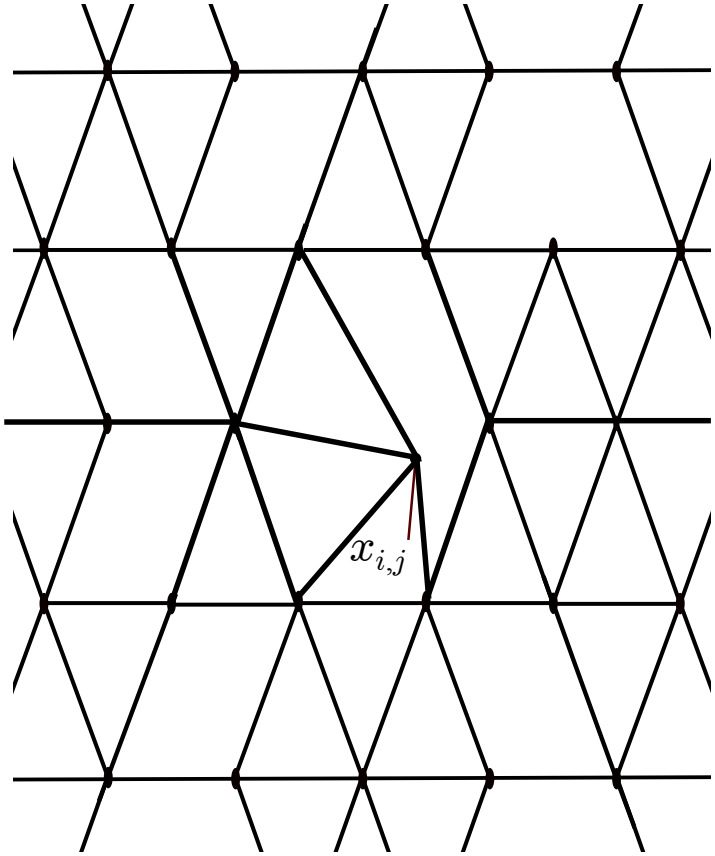

Figure 7.2: Example of an irregular lattice where the framework is not usable

of the continuous-time Markov chain. With these tools we prove that the theoretical first trigonometric moment and the first trigonometric moment of the Markov chain are equal if the transition rates are chosen as

$$\lambda_{i,i+1} \triangleq \frac{\sin(x_{i-1} - x_i)}{2\gamma}\left(\frac{2\pi}{L}\right)^2, \tag{7.6}$$

$$\lambda_{i,i-1} \triangleq \frac{\sin(x_i - x_{i+1})}{2\gamma}\left(\frac{2\pi}{L}\right)^2, \tag{7.7}$$

where

$$\gamma = \sin(x_{i-1} - x_i) - \sin(x_{i-1} - x_{i+1}) + \sin(x_i - x_{i+1}).$$

This time we take the sine of differences between particle positions, scaled by the length of the curve. The usage of trigonometric functions in the transition rates is due to the circular boundary condition. We also mentioned the similarity between the transition rates obtained in Chapter 5 and the transition rates obtained in the Chapter 2. With this similarity in mind we could take a naive approach and use the transition rates for the infinite straight line and apply those transition rates to the closed curves by "unwrapping" the closed curve onto the infinite straight line and then conduct the random walk of particles. In that case we would not obtain accurate results but could still find a reasonable approximation of the diffusion of heat.

In Chapter 6 we conducted an empirical study on the heat diffusion approximation on the S^2-sphere. We discussed a naive approach of projecting a grid on the sphere into the plane such that we were able to calculate transition rates which characterize the random walk on the S^2-sphere. The obtained results show a good approximation when considering the variance and mean of the latitude. The main problem with regards to the derivation of theoretical results is the definition of a grid on the infinite plane that can be used to describe a grid on the sphere by wrapping it around.

This thesis gives the foundation for a technique for the approximation of the diffusion of heat with particular emphasis on the random movement of heat particles on a microscopic level. We discussed diffusion equations and particularly the heat equation in various contexts. We gave examples of other techniques for the modelling the diffusion of heat, deterministic and stochastic, and compared them to our framework. One motivation for this research is the investigation of the usage of non-uniform grids for the characterization of a stochastic process for the modelling of a specific natural phenomenon, the heat diffusion. We are interested how the metric and the curvature in particular influences the transition rates. The result we obtained for the closed curve show that we are able to model the diffusion of heat on any closed curve if the metric is known. Longer distance means lower transition rates and shorter distance means higher transition rates. The reason for the good results on closed curves is that our framework has no limitations regarding the placements of nodes on the infinite straight line which can then, relatively easy, be wrapped around the closed curve.

For two dimensions we encountered certain limitations as explained. In addition to the

length of an edge, the angles between edges play an important role when determining the transition rates. Ultimately we want to be able to use our framework on any Riemannian manifold. Current limitations are 1) the limitation of lattice structures on the plane and 2) the ability to wrap this lattice around a two-dimensional manifold, i.e., the boundaries or surfaces of solid objects in three-dimensional Euclidean space, in a clearly defined manner. In Chapter 6 we described a naive approach to model the diffusion of heat on the S^2-sphere. In this approach we considered the projection of the sphere to the plane, and let particles do the random walk on the plane. Even though we obtained good results, we cannot guarantee good results due to the lack of theoretical results. This is clearly the most promising direction for future research. Previous techniques such as wrapped Gaussian distributions as well as the comparison between statistical properties of the distribution of heat and the distribution of particles did not lead to any usable results. It is therefore necessary to find a new technique that allows for a theoretically correct derivation of transition rates and that guarantees certain statistical properties for the distribution of particles.

Another topic is the modelling of the diffusion of heat on manifolds with limited information regarding the curvature. If we are, for instance, only given an upper bound for the metric of some segments for the closed curve, how well are we able to model the diffusion of heat? This is an interesting question since in real application, there are cases where not the whole manifold is known. Regarding the computational complexity, the actual implementation of the computation of random walks for distributed computing networks is a very interesting topic since our framework basically offers itself for parallel distributed computing. In conjunction with the smart placement of nodes due to the usage of non-uniform placed nodes, the modelling of the diffusion of heat on grids with hundreds of thousands of nodes.

Bibliography

[1] S. Chapra and R. Canale, *Numerical Methods for Engineers*. McGraw-Hill, NY, 1998.

[2] F. Black and M. Scholes, "The pricing of options and corporate liabilities," *Journal of Political Economy*, vol. 81, pp. 637–654, 1973.

[3] H. Risken, *The Fokker-Planck Equation*. Springer, 1989.

[4] C. W. Gardiner, *Handbook of Stochastic Methods*. Springer-Verlag, Berlin, 2004.

[5] J. H. Lienhard, *A Heat Transfer Textbook*. Phlogiston Press, 2008.

[6] W. Stoecker, *Design of Thermal Systems*. McGraw-Hill, NY, 1989.

[7] E. Oran and J. Boris, *Numerical Simulation of Reactive Flow*. Elsevier, Amsterdam, 1987.

[8] W. Gong, "Are stochastic dynamics the foundation of intelligence?," in *Proc. IEEE Conf. Dec. Contr. and Eur. Contr. Conf.*, (Orlando, FL), 2011.

[9] H. Carslaw and J. C. Jaeger, *Conduction of Heat in Solids*. Oxford University Press, 1959.

[10] J. Crank, "A theoretical investigation of the influence of molecular relaxation and internal stress on diffusion in polymers," *Journal of Polymer Science*, vol. 11, no. 2, pp. 151–168, 1953.

[11] M. Briani, *Numerical methods for option pricing in jump-diffusion markets*. PhD thesis, 2003.

[12] C. Yeang and M. Szummer, "Markov random walk representation with continuous distributions," in *Proc. Uncertainty in Artificial Intelligence*, pp. 600–607, 2003.

[13] R. I. Kondor and J. Lafferty, "Diffusion kernels on graphs and other discrete input spaces," in *Proc. of the 19th ICML*, pp. 315–322, Morgan Kaufmann, 2002.

[14] J. Lafferty and G. Lebanon, "Diffusion kernels on statistical manifolds," *Journal of Machine Learning Research*, vol. 6, pp. 129–163, 2005.

[15] E. Gine and V. Koltchinskii, "Empirical graph Laplacian approximation of Laplace–Beltrami operators: Large sample results," *High Dimensional Probability. Institute of Mathematical Statistics Lecture Notes Monograph Series*, vol. 51, pp. 238–259, 2006.

[16] N. Belkin, J. Sun, and Y. Wang, "Discrete Laplace operator on meshed surfaces," *Proc. SOCG*, pp. 278–287, 2008.

[17] K. Sinha and M. Belkin, "Towards a theoretical foundation for Laplacian-based manifold methods," *J. Comput. Syst. Sci.*, vol. 74, pp. 1289–1308, 2008.

[18] D. Ting, L. Huang, and M. Jordan, "An analysis of the convergence of the graph laplacians," *Arxiv preprint math.ML/1101.5435v1*, 2011.

[19] W. Zeng, R. Guo, and X. Gu, "Discrete heat kernel determines discrete Riemannian metric," *Graphical Models*, vol. 74, pp. 121–129, 2012.

[20] G. Lawler, *Random Walk and the Heat Equation.* American Mathematical Society, 2010.

[21] W. Gilks, *Markov Chain Monte Carlo.* Encyclopedia of Biostatistics, 2005.

[22] J. Howell, "The monte carlo method in radiative heat transfer," *J. Heat Transfer*, vol. 120, pp. 547–560, 1998.

[23] R. Erban and S. J. Chapman, "Stochastic modelling of reactiondiffusion processes: algorithms for bimolecular reactions," *Phys. Biol.*, vol. 6, 2009.

[24] M. Cerrato, C. Lo, and K. Skindilias, "Adaptive continuous time markov chain approximation model to general jump-diffusion," in *working paper*, 2011.

[25] L. Komzsik, *Trends Parallel,. Distributed, Grid and Cloud Computing.* Saxe-Coburg Publications, UK, 2011.

[26] F. Guias, "A stochastic numerical method for diffusion equations and applications to spatially inhomogeneous coagulation processes," *Monte Carlo and Quasi-Monte Carlo Methods*, pp. 147–161, 2006.

[27] J. Reddy and D. Gartling, *The Finite Element Method in Heat Transfer and Fluid Dynamics*. CRC Press, Boca Raton, 2010.

[28] D. Wolf-Gladrow, "A lattice Boltzman equation for diffusion," *Journal of Statistical Physics*, vol. 79, pp. 1023–1032, 1995.

[29] H. Huang, X. Lu, and M. Sukop, "Numerical study of lattice boltzmann methods for a convection–diffusion equation coupled with navier–stokes equations," *Journal of Physics A: Mathematical and Theoretical*, vol. 44, no. 5, p. 055001, 2011.

[30] S. J. Ruuth and B. Merriman, "A simple embedding method for solving partial differential equations on surfaces," *J. Comput. Phys.*, vol. 227, no. 3, pp. 1943–1961, 2008.

[31] M. Bertalmio, L. Cheng, S. Osher, and G. Sapiro, "Variational problems and partial differential equations on implicit surfaces," *J. Comput. Phys.*, vol. 174, no. 2, pp. 759 – 780, 2001.

[32] C. B. Macdonald, J. Brandman, and S. J. Ruuth, "Solving eigenvalue problems on curved surfaces using the closest point method," *J. Comput. Phys.*, vol. 230, no. 22, pp. 7944–7956, 2011.

[33] E. A. Codling, M. J. Plank, and S. Benhamou, "Random walk models in biology," *J. R. Soc. Interface*, vol. 5, pp. 813–834, 2008.

[34] R. Metzler and J. Klafter, "The restaurant at the end of the random walk: recent developments in the description of anomalous transport by fractional dynamics," *J. Phys. A: Math. Gen.*, vol. 37, pp. 161–208, 2004.

[35] L. Kleinrock, *Queueing Systems, Volume 1: Theory*. Wiley-Interscience, 1975.

[36] D. Mozyrska and Z. Bartosiewicz, "Observability of a class of linear dynamic infinite systems on time scales," *Proc. Estonian Acad. Sci. Phys. Math.*, vol. 56, pp. 347–358, 2007.

[37] D. Mozyrska and Z. Bartosiewicz, "Dualities for linear control differential systems with infinite matrices," *Control and Cybernetics*, vol. 35, pp. 887–904, 2006.

[38] N. van Kampen, "Markov processes and their diffusion approximations," in *Reprinted from: Thermodynamics & Kinetics of Biological Processes*, pp. 181–195, Walter de Gruyter & Co., New York, 1982.

[39] Z. Schuss, "Markov processes and their diffusion approximations," in *Theory and Applications of Stochastic Processes*, vol. 170 of *Applied Mathematical Sciences*, pp. 207–256, Springer New York, 2010.

[40] M. Bossy and N. Champagnat, "Markov processes and parabolic partial differential equations," in *Encyclopedia of Quantitative Finance*, pp. 1142–1159, John Wiley & Sons Ltd. Chichester, UK, 2010.

[41] V. Horak and P. Gruber, "Parallel numerical solution of 2-D heat equation," *Parallel Numerics*, pp. 47–56, 2005.

[42] M. R. Driels and S. Y. S., "Determining the number of iterations for monte carlo simulations of weapon effectiveness," 2004.

[43] P. Moerters and P. Yuval, *Brownian Motion*. Cambridge University Press, 2010.

[44] A. I. Bobenko and B. A. Springborn, "A discrete LaplaceBeltrami operator for simplicial surfaces," *Discrete Comput. Geom.*, vol. 38, pp. 740–756, 2007.

[45] P. Neta, *Partial differential equations. MA3132. Lecture notes*. Department of Mathematics, Naval Postgraduate School, 2002.

[46] K. V. Mardia and P. E. Jupp, "A unified view of the theory of directional statistics, 1975-1988," *International Statistical Review*, 1989.

[47] N. I. Fisher, *Statistical Analysis of Circular Data*. Cambridge University Press, 1993.

[48] K. V. Mardia and P. E. Jupp, *Directional Statistics*. John Wiley & Sons, 2000.

[49] Z. Bisgaard, T. M. an Sasvari, *Characteristic Functions and Moment Sequences*. Nova Science, 2000.

[50] G. Jona-Lasinio, A. Gelfand, and M. Jona-Lasinio, "Spatial analysis of wave direction data using wrapped Gaussian processes," *The Annals of Applied Statistics*, vol. 6, pp. 1478–1498, 2012.

[51] A. C. Codling, *Biased Random Walks in Biology*. PhD thesis, The University of Leeds, 2003.

[52] P. Berens, "Circstat: A Matlab toolbox for circular statistics," *Journal of Statistical Software*, vol. 31, no. 10, pp. 1–21, 2009.

[53] R. Castaneda-Priego, P. Castro-Villarreal, S. Estrada-Jimenez, and J. M. Mendez-Alcaraz, "Brownian motion of free particles on curved surfaces," *arXiv/1211.5799*, 2012.

[54] G. S. Chirikijan, *Stochastic models, Information Theory, and Lie Groups, Volume 1*. Springer, 2009.

[55] S. Waner, *Introduction to Differential Geometry & General Relativity*. Department of Mathematics, Hofstra University, 2005.

[56] J. Gallier, *Notes on spherical harmonics and linear representations of Lie groups*. Lecture Notes, http://www.cis.upenn.edu/ cis610/, 2009.

[57] V. Tulovsky and L. Papiez, "Formula for the fundamental solution of the heat equation on the sphere," *Applied Mathematics Letters*, vol. 14, pp. 881–884, 2001.

[58] A. Ghosh, J. Samuel, and S. S., "A "*g*aussian″ for diffusion on the sphere," *EPL*, vol. 98, no. 3, 2012.

[59] G. S. Watson, "Distributions on the circle and sphere," *Journal of Applied Probability*, vol. 19, pp. 265–280, 1982.

[60] P. H. Roberts and H. D. Ursell, "Random walk on a sphere and on a Riemannian manifold," *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, vol. 252, pp. 317–356, 1960.

[61] J. P. Snyder, *Map Projections: A Working Manual*. Geological Survey (U.S.), 1987.