On the Discrete Heat Equation Taking Values on a Tree

Carl Mueller[1]
Dept. of Mathematics
University of Rochester
Rochester, NY 14627
E-mail: cmlr@math.rochester.edu

Kijung Lee
Dept. of Mathematics
Yonsei University
Seoul 120-749 Korea
E-mail: kijung@yonsei.ac.kr

1 Introduction

The basic ideas about differential equations taking values on manifolds are well known. Usually one localizes to a coordinate patch, and studies the differential equation using local coordinates. If the manifold structure is smooth, then this procedure yields a new differential equation on Euclidean space, and one can use existing theory.

However, there are two cases which may cause trouble: the manifold may not be smooth, or the equation may have stochastic terms which lead to singularities. We will not say much about the stochastic case, although it was the motivation for this work. The reader can consult Funaki [Fun83], where the author studies the heat equation with additive noise taking values on a manifold. He requires the noise to be very smooth in x, excluding the interesting case of white noise. For the white noise case, the problem is still unsolved.

Turning to the deterministic case, there has been some work about PDE on nonsmooth manifolds. For example, Eells and Fuglede [EF01] deal with harmonic maps between Riemannian polyhedra. For the most part, though, not much is known about this situation, and the reasons are not hard to discover. Traditional PDE theory completely relies on linear techniques. For example, $u+v$ makes no sense if u, v take values in a manifold, and for the same reason we cannot integrate manifold-valued functions or define Sobolev spaces like L_p or W_p^k in the usual way. This is why one sets up coordinate patches which have a linear

[1]Supported by NSF and NSA grants.
Key words and phrases. heat equation, white noise, stochastic partial differential equations.
AMS 1991 *subject classifications* Primary, 60H15; Secondary, 35R60, 35K05.

structure.

As for a physical application, the stochastic heat equation with solutions taking values in $\mathbb{R}^3$ is one of the main models used in polymer science, see Doi and Edwards, [DE88]. Thus, we can think of $u(t, x)$ as the position of an elastic string at time t and at length x along the string. In some situations the string might be confined to a region or surface, so we are led to consider the heat equation taking values on a manifold.

This paper has a modest goal. We consider a very simple nonsmooth manifold, the union of rays emanating from the origin, and consider the heat equation taking values in this space. We find a sequence of discrete approximations, and show sequential compactness using the Arzela-Ascoli theorem. Our main tool is a probabilistic representation of solutions. Unfortunately, we do not prove uniqueness, but our guess is that there is a unique limit for such sequences. It would also be nice to define our equation in the generalized sense, so we could say that our limit point satisfies the equation. These questions are unsolved at the moment.

This paper is organized in the following way. In section 2 we set up our heat equation and present our notation. In section 3 we consider discrete approximations and suggest a tool for sequential compactness. After giving some estimates in section 4, we prove our main theorem in section 5, which states that our sequence of discrete approximations is sequentially compact.

Here are some general references. For the heat equation and other PDE, see Evans [Eva98]. For probability, random walks and martingales, see Durrett [Dur96]. Finally, we list some commonly known notation.

$\mathbb{N}$: the set of natural numbers.

$\mathbb{Z}$: the set of integers.

$\mathbb{R}$: the set of real numbers.

$\mathbb{R}_+$: the set of nonnegative numbers.

EX : the mathematical expectation of a random variable X.

$x \wedge y$: the minimum of $\{x, y\}$.

$\mathbf{1}_A$: indicator or characteristic function on event set A.

$N(\mu, \sigma^2)$: normal distribution with mean μ and variance σ^2.

2 Notation

We wish to consider solutions $u(t,x)$ with $t>0$ and $x\in\mathbb{R}$, to the heat equation

$$\frac{\partial u}{\partial t}=\frac{1}{2}\frac{\partial^2 u}{\partial x^2} \tag{2.1}$$

taking values in a tree which we define below.

For the range of $u(t,x)$, we consider a tree $\mathcal{T}$ embedded in $\mathbb{R}^2$. Let $\mathcal{T}$ be the union of a N rays which meets at one common point $(0,0)$. To put it concretely, we can set

$$\mathcal{T}=\bigcup_{i=1}^{N}\{(r,\theta_i):r\geq 0\}$$

with a finite number N of distinct angles θ_i in polar coordinates.

Note that for such a tree $\mathcal{T}$, we can define a metric $\rho(a,b)$ between two points $a,b\in\mathcal{T}$ as the length of the shortest path (geodesic) in $\mathcal{T}$ from a to b. In particular, the triangle inequality holds:

$$\rho(a,b)\leq\rho(a,c)+\rho(c,b)$$

for any $a,b,c\in\mathcal{T}$. Note that there is a unique such geodesic joining a and b, and a unique midpoint along the geodesic. We denote this midpoint as

$$a\odot b$$

and we use this key notion throughout the paper.

3 The discrete heat equations

Let us consider $u_0:\mathbb{R}\to\mathcal{T}$ which serves as our initial condition.

Given a positive integer n we define a grid T_n on $[0,\infty)$ with spacing $\frac{1}{n}$, and a grid R_n on $\mathbb{R}$ with spacing $\frac{1}{\sqrt{n}}$. i.e.

$$\begin{aligned} T_n &= \left\{t_i=\frac{i}{n}\in[0,\infty):i\in\mathbb{N}\cup\{0\}\right\} \\ R_n &= \left\{x_j=\frac{j}{\sqrt{n}}\in\mathbb{R}:j\in\mathbb{Z}\right\}. \end{aligned}$$

Next, we define a grid $\mathbf{L}_n$ containing every other point of $T_n\times R_n$.

$$\mathbf{L}_n=\left\{\left(\frac{i}{n},\frac{j}{\sqrt{n}}\right)\in T_n\times R_n : i+j \text{ is even}\right\}$$

We define the discrete heat equation on $\mathbf{L}_n$. We are interested in finding the solution $u_n = u_n(t_i, x_j)$, $(t_i, x_j) \in \mathbf{L}_n$, satisfying

$$\begin{aligned} u_n(t_{i+1}, x_j) &= u_n(t_i, x_{j+1}) \odot u_n(t_i, x_{j-1}) \\ u_n(0, x_j) &= u_0(x_j). \end{aligned} \tag{3.1}$$

This is our discrete approximation for a given n. Equation (3.1) is just the analogue of the discrete approximation to the heat equation taking values in $\mathbb{R}$:

$$\begin{aligned} \frac{\partial v}{\partial t} &= \frac{1}{2}\frac{\partial^2 v}{\partial x^2} \\ v(0, x) &= v_0(x). \end{aligned} \tag{3.2}$$

Indeed, for $(t_i, x_j) \in \mathbf{L}_n$ we could approximate (3.2) by the Euler method :

$$\begin{aligned} \frac{\Big[v_n(t_{i+1}, x_j) - v_n(t_i, x_j)\Big]}{\frac{1}{n}} &= \frac{1}{2}\frac{\Big[v_n(t_i, x_{j+1}) - 2v_n(t_i, x_j) + v_n(t_i, x_{j-1})\Big]}{\left(\frac{1}{\sqrt{n}}\right)^2} \\ v_n(0, x_j) &= v_0(x_j). \end{aligned} \tag{3.3}$$

which reduces to (3.1) with v_n in place of u_n.

The equation (3.1) allows us to find the solution $u_n(t_i, x_j)$ by induction on time t_i. To define $u_n(t, x)$ for all $t \geq 0$, $x \in \mathbb{R}$, we use linear interpolation as follows. First fix t_i and define $u(t_i, x)$ for $x \in \mathbb{R}$ by linear interpolation. Next, fix $x \in \mathbb{R}$ and define $u(t, x)$ for $t \geq 0$ by linear interpolation.

We conjecture that u_n, $n = 1, 2, \ldots$ are good approximations to the solution satisfying (2.1) in a natural sense. So, our question is how to show that the approximations are reliable.

Let $\{X_i : i = 0, 1, 2, \ldots\}$ be a simple random walk on the space grid R_n with X_0 at a given grid point. This means that the step size of the walk is $\frac{1}{\sqrt{n}}$, and the time between steps is $\frac{1}{n}$. Using this random walk, we can give a probabilistic solution to (3.3) as follows:

$$v_n(t_i, x_j) = E_{x_j}\left[v_0(X_i)\right] = \sum_{y_k \in R_n} v_0(y_k) P(X_i = y_k), \tag{3.4}$$

where E_{x_j} means the expectation given $X_0 = x_j$. One can use this expression for showing sequential compactness of $\{v_n\}$. Our goal is to carry over this reasoning to the heat equation with values in $\mathcal{T}$. However, $\mathcal{T}$ is not a linear space, so the summation in (3.4) is not well-defined in $\mathcal{T}$ and we suffer from this difficulty. However, if we use coupling, a probabilistic tool whose motivation comes from the random walk, and keep track of the entire history of the random walks, then we can overcome the difficulty.

4 Preliminary estimates

For this section, fix n. We will consider two simple random walks $X_\cdot^{(i)},\ i=1,2$ taking values in R_n. Thus, each random walk has steps of size $\frac{1}{\sqrt{n}}$ and time between steps $\frac{1}{n}$. We assume that $X_0^{(i)}=x_i$ with $(0,x_i)\in\mathbf{L}_n$ for $i=1,2$. We note that with probability 1, $(t_k,X_k^{(i)})\in\mathbf{L}_n$ for $i,j=1,2$ and $k=0,1,\ldots$.

We are interested in two simple random walks which are coupled; we will use the coupling later in Section 5. Recall that a coupling of the random walks is a realization of $X_\cdot^{(i)},\ i=1,2$ on a single probability space such that the marginal distribution of each walk $X_\cdot^{(i)}$ is the same as before.

Suppose the random walks $X_\cdot^{(i)}$ are coupled in such a way that for a fixed k we have $X_k^{(1)}=X_k^{(2)}$ with high probability. Then it seems reasonable that $E_{x_1}[v_0(X_k^{(1)})]$ and $E_{x_2}[v_0(X_k^{(2)})]$ should be close. Since this kind of coupling is easier if x_1 and x_2 are close, we might use coupling to show that solutions $v(t,x)$ of (3.2), which are given by (3.4) are regular in x. This reasoning has been used before, see Cranston [Cra91], for example. Such regularity helps to prove sequential compactness of solutions.

Lemma 4.1. *There is a coupling of the random walks $X_\cdot^{(i)},\ i=1,2$ and a constant $C>0$ not depending on n such that for all $k\in\mathbb{N}$ we have*

$$P\left(X_k^{(1)}\neq X_k^{(2)}\right)\leq\left[C\left(\frac{n}{k}\right)^{1/2}|x_1-x_2|\right]\wedge 1.$$

Proof. 1. Let $h=|x_1-x_2|/2$. Translating if necessary, we may assume that $x_1=h$, $x_2=-h$. Construct $X_\cdot^{(1)}$ on some probability space $(\Omega,\mathcal{F},P)$ in the usual way. Let τ be the first time k that $X_k^{(1)}=0$. Define $X_\cdot^{(2)}$ as follows:

$$X_k^{(2)}=\begin{cases}-X_k^{(1)} & \text{if } k<\tau\\ X_k^{(1)} & \text{if } k\geq\tau\end{cases}$$

so with probability 1, that two walks meet eventually. This coupling is well known, and the reader can easily verify that $X_\cdot^{(2)}$ has the correct marginal distribution.

2. If $x_1 = x_2$, then $X^{(1)}_\cdot$ is identical with $X^{(2)}_\cdot$ and our claim holds trivially. Hence, let us assume $x_1 \neq x_2$. Since we assumed that $(0, x_i) \in \mathbf{L}_n$ for $i = 1, 2$, we note that $|x_1 - x_2| \geq \frac{2}{\sqrt{n}}$, or $h\sqrt{n} \geq 1$. We use this fact to get (4.2) below from the previous line.

Using the reflection principle in the usual way, (see Feller [Fel68], Chapter III) we get

$$\begin{aligned} P\left(X_k^{(1)} \neq X_k^{(2)}\right) &= P(\tau > k) \\ &= 1 - P\left(\inf_{0 \leq i \leq k} X_i^{(1)} \leq 0\right) \\ &= 1 - P\left(X_k^{(1)} \leq 0\right) - P\left(X_k^{(1)} < 0\right) \\ &\leq P\left(\left|X_k^{(1)} - X_0^{(1)}\right| \leq h\right) \end{aligned} \tag{4.1}$$

Heuristically, one can easily finish the proof using the normal approximation, but we will give a rigorous argument using Stirling's approximation. We assume that $k = 2m$ is even, leaving the odd case to the reader. Let S_k be the simple random walk with steps of size 1 starting at $S_0 = 0$. Note that the probability function $p_i = P(S_k = i)$ is even. p_i is zero at every other site i, and where it is nonzero it is radially nonincreasing. Therefore,

$$\begin{aligned} P\left(\left|X_k^{(1)} - X_0^{(1)}\right| \leq h\right) &= P\left(|S_{2m}| \leq h\sqrt{n}\right) \\ &\leq (2h\sqrt{n} + 1)P(S_{2m} = 0) \\ &\leq 3h\sqrt{n}P(S_{2m} = 0). \end{aligned} \tag{4.2}$$

Next, Stirling's formula (see, for instance, (7.3.4) in p.211 of [Chu79]) gives us the following :

$$\begin{aligned} P(S_{2m} = 0) &= \binom{2m}{m} 2^{-2m} \\ &\leq \frac{e^{\frac{1}{12 \cdot 2m}}}{e^{\frac{1}{12(m+1/2)} \cdot 2}} \cdot \frac{\sqrt{2\pi 2m}\, e^{-2m}(2m)^{2m}}{\left(\sqrt{2\pi m}\, e^{-m} m^m\right)^2} \cdot 2^{-2m} \\ &= \exp\left[\frac{1}{24} \frac{\frac{1}{2} - 3m}{m(m + \frac{1}{2})}\right] \cdot (\pi m)^{-1/2} \\ &\leq \sqrt{\frac{2}{m}}, \end{aligned}$$

and so

$$P(S_k = 0) \leq \frac{1}{\sqrt{k}}. \tag{4.3}$$

We can finish the proof of Lemma 4.1 by putting together (4.1), (4.2), (4.3), remembering that $h = |x_1 - x_2|/2$, and choosing $C \geq \frac{3}{2}$. ∎

Now we state a lemma about midpoints in $\mathcal{T}$.

Lemma 4.2. *Let $x_1, x_2, y_2, y_2 \in \mathcal{T}$. Then*

$$\rho\left(x_1 \odot x_2, y_1 \odot y_2\right) \leq \frac{\rho(x_1, y_1)}{2} + \frac{\rho(x_2, y_2)}{2}. \tag{4.4}$$

Proof. There are several cases by the locations of x_1, x_2, y_1, y_2 in rays.

1. If x_1, x_2, y_1, y_2 lie on just 1 or 2 rays, then (4.4) follows from the triangle inequality on $\mathbb{R}_+$ or $\mathbb{R}$.

2. Suppose that x_1, x_2, y_1, y_2 lie on 3 rays and they do not lie on just 1 or 2 rays. Then one ray must contain 2 points, and the other 2 rays contain 1 point each. By symmetry, we can reduce to two situations. Let us denote three rays of $\mathcal{T}$ by $R_i,\ i = 1, 2, 3$.

Case 1. Suppose that x_1, x_2 lie on a common ray R_3, y_1 lies on ray R_1, and y_2 lies on ray R_2. Let $y_1' \in R_2$ be the same distance from the origin as y_1. Then $\rho(x_1, y_1) = \rho(x_1, y_1')$. However, it is easily seen that

$$\rho\left(0, y_1 \odot y_2\right) \leq \rho\left(0, y_1' \odot y_2\right)$$

and that

$$\rho\left(x_1 \odot x_2, y_1 \odot y_2\right) = \rho\left(x_1 \odot x_2, 0\right) + \rho\left(0, y_1 \odot y_2\right)$$

Thus, since x_1, x_2, y_1', y_2 lie on 2 rays, we conclude that

$$\begin{aligned} \rho\left(x_1 \odot x_2, y_1 \odot y_2\right) &= \rho\left(x_1 \odot x_2, 0\right) + \rho\left(0, y_1 \odot y_2\right) \\ &\leq \rho\left(x_1 \odot x_2, 0\right) + \rho\left(0, y_1' \odot y_2\right) \\ &= \rho\left(x_1 \odot x_2, y_1' \odot y_2\right) \\ &\leq \frac{\rho(x_1, y_1')}{2} + \frac{\rho(x_2, y_2)}{2} \\ &= \frac{\rho(x_1, y_1)}{2} + \frac{\rho(x_2, y_2)}{2} \end{aligned}$$

Case 2. Suppose that x_1, y_1 lie on a common ray R_3, x_2 lies on ray R_1, and y_2 lies on ray R_2.

First suppose that either $x_1 \odot x_2$ or $y_1 \odot y_2$ lie on R_3. By changing the labeling, we may assume $x_1 \odot x_2$ lies on R_3. Let $y_2' \in R_1$ be the same distance from the origin as y_2. Replacing

y_2 by y_2', we see that the left side of (4.4) has not changed, but the right side has become smaller, and all 4 points lie on just 2 rays. Thus, (4.4) follows from the triangle inequality on $\mathbb{R}$.

Secondly, suppose that neither $x_1 \odot x_2$ nor $y_1 \odot y_2$ lie on R_3. Thus, $x_1 \odot x_2 \in R_1$ and $y_1 \odot y_2 \in R_2$. Also, suppose without loss of generality that x_1 lies closer to the origin than y_1. One can see that moving y_1 to position x_1 changes $\rho(0, y_1 \odot y_2)$ and $\frac{\rho(x_1,y_1)}{2}$ by the same amount. Hence, we have

$$\begin{aligned}
\rho(x_1 \odot x_2, y_1 \odot y_2) - \frac{\rho(x_1, y_1)}{2} &= \rho(0, x_1 \odot x_2) + \rho(0, y_1 \odot y_2) - \frac{\rho(x_1, y_1)}{2} \\
&= \rho(0, x_1 \odot x_2) + \rho(0, x_1 \odot y_2) \\
&\leq \frac{\rho(0, x_2)}{2} + \frac{\rho(0, y_2)}{2} \\
&= \frac{\rho(x_2, y_2)}{2}.
\end{aligned}$$

3. Finally, let's assume that x_1, x_2, y_1, y_2 lie on 4 rays and none of them shares the same ray with others. In this case, we clearly see

$$\begin{aligned}
\rho(x_1 \odot x_2, y_1 \odot y_2) &\leq \rho(x_1, 0) \wedge \rho(x_2, 0) + \rho(y_1, 0) \wedge \rho(y_2, 0) \\
&\leq \frac{\rho(x_1, 0) + \rho(x_2, 0)}{2} + \frac{\rho(y_1, 0) + \rho(y_2, 0)}{2} \\
&= \frac{\rho(x_1, 0) + \rho(y_1, 0)}{2} + \frac{\rho(x_2, 0) + \rho(y_2, 0)}{2} \\
&= \frac{\rho(x_1, y_1)}{2} + \frac{\rho(x_2, y_2)}{2}.
\end{aligned}$$

This finishes the proof of Lemma 4.2. ∎

5 Sequential compactness

Here we prove our main theorem which indicates that the process of our discrete approximations is stable. In the theorem, our condition on u_0 is certainly not the weakest possible (see Assumption 5.2), but it yields a simple proof.

First, we need some definitions and an assumption on u_0. Since ρ is a metric on $\mathcal{T}$, we can speak of uniform continuity and equicontinuity for functions taking values in $\mathcal{T}$.

Definition 5.1. If a positive function ζ is defined on positive real numbers $(0, \infty)$ and satisfies $\lim_{h \to 0^+} \zeta(h) = 0$, then we write $\zeta \in \mathcal{M}_0$.

Assumption 5.2. We assume that u_0 is uniformly continuous with a modulus function κ. That is, $\kappa \in \mathcal{M}_0$ is monotone increasing and for all $x, y \in \mathbb{R}$ we have

$$\rho(u_0(x), u_0(y)) \leq \kappa(|x-y|). \tag{5.1}$$

We also assume that the growth of κ is no faster than linear:

$$\kappa(h) \quad \leq \quad N_0 + N_1 h \tag{5.2}$$

for some absolute constants $N_0, N_1 > 0$.

We state our main theorem.

Theorem 5.3. *Let $u_0 : \mathbb{R} \to \mathcal{T}$ satisfy Assumption 5.2. Then the collection of solutions $\{u_n : n = 1, 2, \ldots\}$ to (3.1) is sequentially compact in the topology of uniform convergence on compact subsets of* $[0, \infty) \times \mathbb{R}$.

According to the Arzela-Ascoli theorem for the functions taking values on a metric space, need only show that on each compact subset of $K \subset [0, \infty) \times \mathbb{R}$, the family $\{u_n\}$ is equicontinuous. In fact, we will show that for any K of the form $K = [0, a] \times [-b, b]$

$$\rho\Bigg(u_n(t,x), u_n(s,y)\Bigg) \leq \zeta(|(t,x)-(s,y)|), \quad (t,x),(s,y) \in K \tag{5.3}$$

for a function $\zeta \in \mathcal{M}_0$ not depending on n.

The following two lemmas are the main ingredients in the proof of Theorem 5.3.

Lemma 5.4. *For given $K = [0, a] \times [-b, b]$ with $a, b > 0$ there exists a function $\zeta_1 \in \mathcal{M}_0$ which does not depend on n, t and satisfying*

$$\rho\Bigg(u_n(t,x), u_n(t,y)\Bigg) \leq \zeta_1(|x-y|), \quad (t,x),(t,y) \in K$$

Remark 5.5. The reader can check that the proof of Lemma 5.4 even works with K of the form $[0, a] \times \mathbb{R}$.

Proof. 1. Fix $t = t_I \in T_n \bigcap [0, a]$ and $x, y \in R_n \bigcap [-b, b]$. We are using the two random walks $X_\cdot^{(i)}$ defined on a Probability space $(\Omega, \mathcal{F}, P)$ from Lemma 4.1 such that

$$X_0^{(1)} = x, \quad X_0^{(2)} = y.$$

Let

$$M_k = \rho\left(u_n(t_I - t_k, X_k^{(1)}), u_n(t_I - t_k, X_k^{(2)})\right), \quad k = 0, \ldots, I$$

and let $\mathcal{F}_k$ be the σ-field generated by $\{X_m^{(1)}, X_m^{(2)} : m \le k\}$. We claim that $(M_k, \mathcal{F}_k)$ is a submartingale.

Fix an event set A which belongs to the partition of Ω which generates $\mathcal{F}_k$. Naturally, $X_k^{(1)}, X_k^{(2)}$, and, hence, M_k are constants on A. Let $X_k^{(1)} = \alpha, X_k^{(2)} = \beta, M_k = \gamma$ on A.

Case 1: $\alpha = \beta$.

In this case $X_{k+1}^{(1)} - X_k^{(1)} = X_{k+1}^{(2)} - X_k^{(2)}$ on A by the construction of $X_\cdot^{(2)}$. The set A can be divided into A^+ and A^- according to the value of $X_{k+1}^{(1)} - X_k^{(1)}$:

$$A^+ = A \bigcap \{X_{k+1}^{(1)} - X_k^{(1)} = \frac{1}{\sqrt{n}}\} \quad A^- = A \bigcap \{X_{k+1}^{(1)} - X_k^{(1)} = -\frac{1}{\sqrt{n}}\}.$$

It is easy to see that the conditional probabilities $P(A^+|A)$, $P(A^-|A)$ are $\frac{1}{2}$.

By our discrete heat equation (3.1), on A we have

$$\begin{aligned} M_k = \gamma \;\; &= \;\; \rho\Bigg(u_n\left(t_I - t_{k+1}, \alpha + \frac{1}{\sqrt{n}}\right) \odot u_n\left(t_I - t_{k+1}, \alpha - \frac{1}{\sqrt{n}}\right), \\ &\qquad u_n\left(t_I - t_{k+1}, \beta + \frac{1}{\sqrt{n}}\right) \odot u_n\left(t_I - t_{k+1}, \beta - \frac{1}{\sqrt{n}}\right)\Bigg) \end{aligned} \tag{5.4}$$

for $k = 0, \ldots, I-1$ and by Lemma 4.2,

$$\begin{aligned} \gamma \;\; &\le \;\; \frac{1}{2}\rho\Bigg(u_n\left(t_I - t_{k+1}, \alpha + \frac{1}{\sqrt{n}}\right),\; u_n\left(t_I - t_{k+1}, \beta + \frac{1}{\sqrt{n}}\right)\Bigg) \\ &\quad + \frac{1}{2}\rho\Bigg(u_n\left(t_I - t_{k+1}, \alpha - \frac{1}{\sqrt{n}}\right),\; u_n\left(t_I - t_{k+1}, \beta - \frac{1}{\sqrt{n}}\right)\Bigg). \end{aligned}$$

This implies

$$\begin{aligned} M_k \;\; &\le \;\; \frac{E\left[M_{k+1} I_{A^+}\right]}{P(A)} + \frac{E\left[M_{k+1} I_{A^-}\right]}{P(A)} \\ &= \;\; E\left[M_k | \mathcal{F}_k\right] \end{aligned}$$

on A.

Case 2: $\alpha \neq \beta$.

This case is similar to the previous case. The differences are the followings. This time we have $X_{k+1}^{(1)} - X_k^{(1)} = -(X_{k+1}^{(2)} - X_k^{(2)})$ on A and we use

$$\begin{aligned} \gamma \;\; &\le \;\; \frac{1}{2}\rho\Bigg(u_n\left(t_I - t_{k+1}, \alpha + \frac{1}{\sqrt{n}}\right),\; u_n\left(t_I - t_{k+1}, \beta - \frac{1}{\sqrt{n}}\right)\Bigg) \\ &\quad + \frac{1}{2}\rho\Bigg(u_n\left(t_I - t_{k+1}, \alpha - \frac{1}{\sqrt{n}}\right),\; u_n\left(t_I - t_{k+1}, \beta + \frac{1}{\sqrt{n}}\right)\Bigg) \end{aligned}$$

which again follows (5.4) and Lemma 4.2 with commutativity of $\odot$. The rest for reaching the inequality (5.10) is essentially the same.

Since we chose arbitrary A in the partition, $(M_k, \mathcal{F}_k)$ is, indeed, a submartingale. It follows that

$$M_0 \leq E\left[M_I\right].$$

2. Therefore, for $t = t_I \in T_n \bigcap [0,a]$ and $x, y \in R_n \bigcap [-b,b]$,

$$\begin{aligned}
\rho\Big(u_n(t,x), u_n(t,y)\Big) &\leq E\left[\rho\Big(u_0(X_I^{(1)}), u_0(X_I^{(2)})\Big)\right] \\
&\leq \rho\Big(u_0(x), u_0(y)\Big) + E\left[\rho\Big(u_0(X_I^{(1)}), u_0(x))\Big)\mathbf{1}_{\{X_I^{(1)} \neq X_I^{(2)}\}}\right] \\
&\quad + E\left[\rho\Big(u_0(y), u_0(X_I^{(2)})\Big)\mathbf{1}_{\{X_I^{(1)} \neq X_I^{(2)}\}}\right] \\
&= (I) + (II) + (III)
\end{aligned}$$

by the triangle inequality. Let $h = |x-y|$. First note that by (5.1),

$$(I) \leq \kappa(h).$$

Next, since (II) and (III) are similar, we will only estimate (II). Again by (5.1),

$$(II) \leq E\left[\kappa(|X_I^{(1)} - x|)\mathbf{1}_{\{X_I^{(1)} \neq X_I^{(2)}\}}\right] =: (IIb).$$

Recall that our coupling estimate in Lemma 4.1 implies that

$$P\left(X_I^{(1)} \neq X_I^{(2)}\right) \leq \left[Ct^{-1/2}|x-y|\right] \wedge 1. \tag{5.5}$$

Roughly speaking, if $t = t_I$ is small, then $\kappa(|X_I^{(1)} - x|)$ will be small. If t is not small then we can use (5.5). Choose a constant $\gamma \in (0,1)$.

Case 1: $t \leq h^{2-2\gamma} = |x-y|^{2-2\gamma}$.

We have

$$\begin{aligned}
(IIb) &= E\left[\kappa(|X_I^{(1)} - x|)\mathbf{1}_{\{X_I^{(1)} \neq X_I^{(2)}\}}\right] \\
&\leq E\left[\kappa(|X_I^{(1)} - x|)\right] \\
&\leq \xi(h)
\end{aligned}$$

with $\xi \in \mathcal{M}_0$. To show that the final line holds with ξ not depending on n, we prove the following : For any constant $\alpha \in (0, \frac{1}{2})$

$$E\left[\kappa(|X_I^{(1)} - x|)\right] \leq \kappa(t^{\frac{1}{2}-\alpha}) + N_0 t^{2\alpha} + N_1 t^{1/2} =: \delta(t) \tag{5.6}$$

holds where the constants N_0, N_1 come from the linear growth condition of κ. We recall that the walk $X^{(1)}_\cdot$ is designed after the choice of n and x. We note that $X^{(1)}_\cdot - x$ doesn't depend on x and $\delta(\cdot)$ itself in (5.6) works for all n. We observe

$$\begin{aligned} E\left[\kappa(|X_I^{(1)} - x|)\right] &= \sum_{-tn\le j\le tn} \kappa\left(\frac{|j|}{\sqrt{n}}\right) P\left(X_I^{(1)} - x = \frac{j}{\sqrt{n}}\right) \\ &= \sum_{|j|<t^{1/2-\alpha}\sqrt{n}} \kappa\left(\frac{|j|}{\sqrt{n}}\right) P\left(\frac{S_{tn}}{\sqrt{tn}} = \frac{j}{\sqrt{tn}}\right) \\ &\quad + \sum_{|j|\ge t^{1/2-\alpha}\sqrt{n}} \kappa\left(\frac{|j|}{\sqrt{n}}\right) P\left(\frac{S_{tn}}{\sqrt{tn}} = \frac{j}{\sqrt{tn}}\right) \\ &= \sigma_1 + \sigma_2, \end{aligned}$$

where $S_\cdot$ is a simple random walk with the step size 1. Obviously, $\sigma_1 \le \kappa(t^{\frac{1}{2}-\alpha})$. Meanwhile, we have

$$\begin{aligned} \sigma_2 &\le N_0 P\left(\left|\frac{S_{tn}}{\sqrt{tn}}\right| \ge \frac{1}{t^\alpha}\right) + N_1\sqrt{t}E\left|\frac{S_{tn}}{\sqrt{tn}}\right| \\ &\le N_0 t^{2\alpha}\cdot E\left|\frac{S_{tn}}{\sqrt{tn}}\right|^2 + N_1\sqrt{t}\cdot\left(E\left|\frac{S_{tn}}{\sqrt{tn}}\right|^2\right)^{1/2} \\ &\le N_0 t^{2\alpha} + N_1\sqrt{t} \end{aligned}$$

since the variance of $\frac{S_{tn}}{\sqrt{tn}}$ is 1. Hence, (5.6) holds. We set $\xi(h) = \delta(h^{2-2\gamma})$ in the situation $t \le h^{2-2\gamma}$.

Case 2: $t \ge h^{2-2\gamma} = |x-y|^{2-2\gamma}$.

Using (5.2), we find

$$\begin{aligned} (IIb) &\le E\left[\left(N_0 + N_1\left|X_t^{(1)} - x\right|\right)\mathbf{1}_{\{X_t^{(1)}\ne X_t^{(2)}\}}\right] \\ &= N_0 P\left(X_t^{(1)} \ne X_t^{(2)}\right) + N_1 E\left[\left|X_t^{(1)} - x\right|\mathbf{1}_{\{X_t^{(1)}\ne X_t^{(2)}\}}\right] \end{aligned} \tag{5.7}$$

Now (5.5) states that

$$P\left(X_t^{(1)} \ne X_t^{(2)}\right) \le Ct^{-1/2}|x-y|. \tag{5.8}$$

Using the Cauchy-Schwarz inequality and redefining C if necessary, we get

$$\begin{aligned} E\left[\left|X_t^{(1)} - x\right|\mathbf{1}_{\{X_t^{(1)}\ne X_t^{(2)}\}}\right] &\le \left(E\left|X_t^{(1)} - x\right|^2\right)^{1/2} P\left(X_t^{(1)} \ne X_t^{(2)}\right)^{1/2} \\ &= Ct^{1/2}\left(t^{-1/2}|x-y|\right)^{1/2}. \end{aligned} \tag{5.9}$$

Putting together (5.7), (5.8), and (5.9), we conclude that

$$\begin{aligned}(IIb) &\leq C\left[t^{-1/2}|x-y| + \left(t^{-1/2}|x-y|\right)^{1/2}\right] \\ &\leq C\left(h^{\gamma} + h^{\gamma/2}\right)\end{aligned}$$

where C may depend on a, the bound of K in time direction, but it does not depend on n, x, y.

3. We note that in our continuity argument it is enough to consider u_n on $(t,x) \in (T_n \times R_n) \bigcap K$ since we construct u_n outside of $T_n \times R_n$ by linear interpolation. We conclude that there exists a function ζ_1 satisfying our claim. ∎

Lemma 5.6. *There exists a function $\zeta_2 \in \mathcal{M}_0$ which does not depend on n, x and satisfies*

$$\rho\Big(u_n(t,x), u_n(s,x))\Big) \leq \zeta_2(|t-s|), \quad t, s \geq 0,\ x \in \mathbb{R}.$$

Proof. 1. Translating if necessary, we may assume that $s = 0, x = 0$.

2. Fix $t_I \in T_n$ and consider a random walk $X.$ with $X_0 = 0$ as in the beginning of the section 4. We denote the probability space by $(\Omega, \mathcal{F}, P)$. Let

$$Z_k = \rho(u_n(t - t_k, X_k), u(0,0)), \quad k = 0, \ldots, I$$

and let $\mathcal{F}_k$ be the σ-field generated by $\{X_m : m \leq k\}$. We claim that $(Z_k, \mathcal{F}_k)$ is a sub-martingale. As in the discussion of the submartingale M_k, fix an event set A which belongs to the partition of Ω which generates $\mathcal{F}_k$. Let $X_k = \alpha, Z_k = \gamma$ on A and define

$$A^+ = A\bigcap\{X_{k+1} - X_k = \frac{1}{\sqrt{n}}\} \quad A^- = A\bigcap\{X_{k+1} - X_k = -\frac{1}{\sqrt{n}}\},$$

where we note that the conditional probabilities $P(A^+|A)$, $P(A^-|A)$ are $\frac{1}{2}$. On A we have

$$\begin{aligned}Z_k = \gamma &= \rho\Bigg(u_n\left(t_I - t_{k+1}, \alpha + \frac{1}{\sqrt{n}}\right) \odot u_n\left(t_I - t_{k+1}, \alpha - \frac{1}{\sqrt{n}}\right), \\ &\qquad u_n(0,0) \odot u_n(0,0)\Bigg)\end{aligned}$$

for $k = 0, \ldots, I-1$ and by Lemma 4.2,

$$\begin{aligned}\gamma &\leq \frac{1}{2}\rho\Bigg(u_n\left(t_I - t_{k+1}, \alpha + \frac{1}{\sqrt{n}}\right),\ u_n(0,0)\Bigg) \\ &\quad + \frac{1}{2}\rho\Bigg(u_n\left(t_I - t_{k+1}, \alpha - \frac{1}{\sqrt{n}}\right),\ u_n(0,0)\Bigg)\end{aligned}$$

which implies

$$\begin{aligned} Z_k &\leq \frac{E\left[Z_{k+1} I_{A^+}\right]}{P(A)} + \frac{E\left[Z_{k+1} I_{A^-}\right]}{P(A)} \\ &= E\left[Z_k | \mathcal{F}_k\right] \end{aligned}$$

on A. Since we chose arbitrary A in the partition, $(Z_k, \mathcal{F}_k)$ is a submartingale and we have

$$Z_0 \leq E\left[Z_I\right].$$

3. Hence, for $t = t_I \in T_n$ we get

$$\begin{aligned} \rho\Big(u_n(t,0), u_n(0,0)\Big) &\leq E\left[\rho\Big(u_0(X_I), u_0(0)\Big)\right] \\ &\leq E\left[\kappa(X_I)\right] \\ &\leq \delta(t), \end{aligned}$$

with $\delta(\cdot)$ in the proof of Lemma 5.4. The choice of ζ_2 is easy and Lemma is proved. ∎

We are ready to prove Theorem 5.3.

Proof of Theorem 5.3. We continue with the comment following the statement of Theorem 5.3. Observing

$$\rho\Big(u_n(t,x), u_n(s,y)\Big) \leq \rho\Big(u_n(t,x), u_n(t,y)\Big) + \rho\Big(u_n(t,y), u_n(s,y)\Big)$$

and using Lemmas 5.4 and 5.6, we have (5.3) with $\zeta = \zeta_1 + \zeta_2$. Theorem 5.3 follows. ∎

Acknowledgement The beginning of this paper came from discussions authors had near the nice espresso machine in the 8th floor of Hylan building, University of Rochester during the second author's stay.

References

[Cra91] Michael Cranston. Gradient estimates on manifolds using coupling. *J. Funct. Anal.*, 99(1):110–124, 1991.

[Chu79] Kai Lai Chung. *Elementary probability theory with stochastic processes*. Third edition. Undergraduate Texts in Mathematics. Springer-Verlag, New York-Heidelberg, 1979.

[DE88] M. Doi and S.F. Edwards. *The theory of polymer dynamics*, volume 73 of *The international series of monographs in physics*. Oxford University Press, Oxford, 1988.

[Dur96] Richard Durrett. *Probability: theory and examples*. Duxbury Press, Belmont, CA, second edition, 1996.

[EF01] J. Eells and B. Fuglede. *Harmonic maps between Riemannian polyhedra*, volume 142 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2001. With a preface by M. Gromov.

[Eva98] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[Fel68] William Feller. *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley & Sons Inc., New York, 1968.

[Fun83] Tadahisa Funaki. Random motion of strings and related stochastic evolution equations. *Nagoya Math. J.*, 89:129–193, 1983.