A Skew-product decomposition of diffusions on a manifold equipped with a group action, A Lorentz model with variable density in a conservative force field, and Reconstruction of a manifold from the intrinsic metric of an associated Markov chain

by

Eric Stephen Wayman

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Mathematics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Steven N. Evans, Chair
Professor Andrew E. Lim
Professor Fraydoun Rezakhanlou

Fall 2014

A Skew-product decomposition of diffusions on a manifold equipped with a group action, A Lorentz model with variable density in a conservative force field, and Reconstruction of a manifold from the intrinsic metric of an associated Markov chain

Abstract

A Skew-product decomposition of diffusions on a manifold equipped with a group action, A Lorentz model with variable density in a conservative force field, and Reconstruction of a manifold from the intrinsic metric of an associated Markov chain

by

Eric Stephen Wayman

Doctor of Philosophy in Mathematics

University of California, Berkeley

Professor Steven N. Evans, Chair

My thesis consists of three different projects.

1) We consider a 2×2 matrix-valued process $(x_t)_{t \geq 0}$ that is obtained by taking a matrix-valued process with entries that are independent one-dimensional standard Brownian motions and time-changing it in a natural way so that the determinant is nonzero for all $t \geq 0$. The QR factorization decomposes $(x_t)_{t \geq 0}$ into a "radial" part $(T_t)_{t \geq 0}$ that is an autonomous diffusion on the set of upper triangular matrices with positive determinant and an "angular" process $(U_{R_t})_{t \geq 0}$, where U is a Brownian motion on the group $SO(2)$ of 2×2 orthogonal matrices with determinant one and the time-change $(R_t)_{t \geq 0}$ is adapted to the filtration generated by $(T_t)_{t \geq 0}$. In this project we show that, unlike classical skew-products such as the celebrated skew-product decomposition of planar Brownian motion into its radial and angular parts, the Brownian motion $(U_t)_{t \geq 0}$ on $SO(2)$ is not independent of the radial part $(T_t)_{t \geq 0}$. We observe that our process fits into the framework of a theorem from [Lia09] on the existence of a skew-product decomposition of a general continuous Markov process on a smooth manifold whose distribution is equivariant under the action of a Lie group. Our result is a counterexample to the main result of [Lia09], but the conclusion of that result holds after a slight strengthening of the hypotheses. These results appear in [EHW14].

2) In Chapter 2, which is based on [HRW14], we study the diffusion limit of a transport process that models the trajectory in $\mathbb{R}^2$ of a particle under the influence of a conservative, spherically symmetric force field $\mathcal{U}$. The particle travels along the trajectory determined by its initial conditions and $\mathcal{U}$ until, according to a Poisson process with variable intensity on this trajectory, it reflects in a uniform direction. We show that under a proper rescaling of time, energy and the density of obstacles, the trajectory converges to a diffusion whose generator can be found explicitly. This generalizes [BR14], where the force field was taken to be constant, to a large class of force fields.

3) A Dirichlet form on a Hilbert space naturally induces a metric on its domain in terms of the energy measure of the form. This metric, which is known as the Carathéodory or intrinsic metric, is studied extensively in [Dav93] where it is used to establish estimates for the heat kernel of a discrete Laplacian operator on a weighted graph. We study the Carathéodory metric associated with the generator of a continuous time Markov chain on a graph of points sampled independently from a distribution on an embedded manifold. Under a proper rescaling of the edge weights, the generator of the Markov chain converges to a weighted Laplacian on the manifold as the number of points goes to infinity. In this third project we conjecture that a rescaling of the Carathéodory distances between any two fixed points on the graph converges to the geodesic distance on the manifold as the number of points on the graph goes to infinity. We prove that the geodesic distances form a limiting lower bound for the Carathéodory distances, and provide some heuristic arguments to indicate why they may be limiting upper bounds as well. However, the upper bound limit remains an open question for future study.

To my parents

Contents

Acknowledgments

I would like to thank my advisor, Steven N. Evans, for his guidance, support and advice. Without his help, this work would not have been possible. I would also like to thank Alexandru Hening and Douglas Rizzolo for the countless discussions about stochastic processes. Lastly, I would like to thank my family and friends for their encouragement and support throughout the years.

Chapter 1

A Skew-Product Decomposition Counterexample

1.1 Introduction

The archetypal skew-product decomposition of a Markov process is the decomposition of a Brownian motion in the plane $(B_t)_{t\geq 0}$ into its radial and angular part

$$B_t = |B_t| \exp(i\theta_t).$$

Here the radial part $(|B_t|)_{t\geq 0}$ is a two-dimensional Bessel process and $\theta_t = y_{\tau_t}$, where $(y_t)_{t\geq 0}$ is a one-dimensional Brownian motion that is independent of the radial part $(|B_t|)_{t\geq 0}$ and τ is a time-change that is adapted to the filtration generated by the process $|B|$. Specifically, $\tau_t = \int_0^t \frac{1}{|B_s|^2} ds$. See Corollary 18.7 from [Kal02] for more details.

The most obvious generalization of this result is obtained in [Gal63]. The process considered is any time-homogeneous diffusion $(x_t)_{t\geq 0}$ with state space $\mathbb{R}^3$ that satisfies the additional assumptions that almost surely every path does not pass through the origin at positive times and that $(x_t)_{t\geq 0}$ is isotropic in the sense that the law of $(x_t)_{t\geq 0}$ is equivariant under the group of orthogonal transformations; that is, if we consider a point $(r, \theta) \in \mathbb{R}^3$ in spherical coordinates, where $r \in \mathbb{R}_+$ is the radial coordinate and θ is a point on the unit sphere S^2, and if we take $k \in O(3)$, the orthogonal group on $\mathbb{R}^3$, then

$$P_{(r,k\theta)}(kA) = P_{(r,\theta)}(A)$$

for any Borel set A in path space $C(\mathbb{R}_+, \mathbb{R}^3)$. Here $P_x(A)$ is the probability a path started at x belongs to the Borel set A [Gal63, (2.2)]. Theorem 1.2 of [Gal63] states that we can decompose $(x_t)_{t\geq 0}$ as $x_t = r_t\theta_t$ where the radial motion $(r_t)_{t\geq 0}$ is a time-homogeneous Markov process on $\mathbb{R}_+$ and the angular process $(\theta_t)_{t\geq 0}$ can be written as $\theta_t = B_{\tau_t}$, with $(B_t)_{t\geq 0}$ a spherical Brownian motion independent of the radial part and with the time-change $(\tau_t)_{t\geq 0}$ adapted to the filtration generated by the radial part.

More generally, one can consider a group G acting on $\mathbb{R}^n$ and $(x_t)_{t\geq 0}$ a Markov process on $\mathbb{R}^n$ such that the distribution of $(x_t)_{t\geq 0}$ satisfies the equivariance condition

$$P_{gx}(gA) = P_x(A)$$

for any Borel set A in path space. The existence of a skew-product decomposition for this setting is explored in [Chy08] when $(x_t)_{t\geq 0}$ is a Dunkl process and G is the group of distance preserving transformations of $\mathbb{R}^n$.

The paper [PR88] investigates the skew-product decomposition of a Brownian motion on a C^∞ Riemannian manifold (M, g) which can be written as a product of a radial manifold R and an angular manifold Θ, both of which are assumed to be smooth and connected. Provided the Riemannian metric respects the product structure of the manifold in a suitable manner, Theorem 4 of [PR88] establishes the existence of a skew-product decomposition such that the radial motion is a Brownian motion with drift on R and the angular motion is a time-change of a Brownian motion on Θ that is independent of the radial motion.

A related skew-product decomposition is obtained in [Lia09] for a general continuous Markov process $(x_t)_{t\geq 0}$ with state space a smooth manifold X and distribution that is equivariant under the smooth action of a Lie group K on X. Here the decomposition of $(x_t)_{t\geq 0}$ is into a radial part $(y_t)_{t\geq 0}$ that is a Markov process on the submanifold Y which is transversal to the orbits of K and an angular part $(z_t)_{t\geq 0}$ that is a process on a general K-orbit which can be identified with the homogeneous space K/M, where M is the isotropy subgroup of K that is assumed to be the same for all elements $x \in X$. Theorem 4 of [Lia09] asserts that under suitable conditions the process $(x_t)_{t\geq 0}$ has the same distribution as $(B(a_t)y_t)_{t\geq 0}$, where the radial part $(y_t)_{t\geq 0}$ is a diffusion on Y, $(B_t)_{t\geq 0}$ is a Brownian motion on K/M that is independent of $(x_t)_{t\geq 0}$, and $(a_t)_{t\geq 0}$ a time-change that is adapted to the filtration generated by $(y_t)_{t\geq 0}$.

Analogous skew-product decompositions of superprocesses have been studied in [Per92, EM91, Hir00]. The continuous Dawson-Watanabe (DW) superprocess is a rescaling limit of a system of branching Markov processes while the Fleming-Viot (FV) superprocess is a rescaling limit of the empirical distribution of a system of particles undergoing Markovian motion and multinomial resampling. It is shown in [EM91] that a FV process is a DW process conditioned to have total mass one. More generally, it is demonstrated in [Per92] that the distribution of the DW process conditioned on the path of its total mass process is equal to the distribution of a time-change of a FV process that has a suitable underlying time-inhomogeneous Markov motion. The latter result is extended to measure-valued processes that may have jumps in [Hir00].

A sampling of other results involving skew-products can be found in [Tay92, LCO09, ELJL10, BN06].

This paper was motivated by our desire to understand better the structural features that give rise to skew-product decompositions of Markov processes. In attempting to do so, we read the paper [Lia09] but were unable to follow some of the details of the proof of the main result, Theorem 4. We subsequently came across a natural and quite simple counterexample to that result which we believe is rather illuminating and which we describe here.

We construct a diffusion $(x_t)_{t\geq 0}$ with state space the manifold of 2×2 matrices that have a positive determinant. This diffusion can be represented via the well-known QR decomposition as the product of an autonomously Markov "radial" process $(T_t)_{t\geq 0}$ on the manifold of 2×2 upper-triangular matrices with positive diagonal entries and a time-changed "angular" process $(U_{R_t})_{t\geq 0}$, where $(U_t)_{t\geq 0}$ is a Brownian motion on the group $SO(2)$ of 2×2 orthogonal matrices with determinant one and the time-change $(R_t)_{t\geq 0}$ is adapted to the filtration of the radial process. However, unlike in the skew-product decompositions described above, the processes $(U_t)_{t\geq 0}$ and $(T_t)_{t\geq 0}$ are **not** independent.

Our process $(x_t)_{t\geq 0}$ satisfies the assumptions of [Lia09, Theorem 4] which asserts that the processes $(U_t)_{t\geq 0}$ and $(T_t)_{t\geq 0}$ are independent. This apparent contradiction appears because the assumption from [Lia09] that K/M is irreducible is not strong enough to ensure the nonexistence of a nonzero M-invariant tangent vector in the case when, as in our construction, K/M has dimension 1. It is the nonexistence of such a tangent vector that is used in the proof in [Lia09] to deduce the independence of the processes $(U_t)_{t\geq 0}$ and $(T_t)_{t\geq 0}$.

1.2 Construction of the counterexample

Recall the well-known QR decomposition which says that any square matrix can be written as the product of an orthogonal matrix and an upper triangular matrix, and that this decomposition is unique for invertible matrices if we require the diagonal entries in the upper triangular matrix to be positive (see, for example, [HJ13]). This decomposition is essentially a special case of the Iwasawa decomposition for semisimple Lie groups.

In the 2×2 case, uniqueness also holds for invertible matrices if we require the orthogonal matrix to have determinant one and there are simple explicit formulae for the factors. Indeed, if

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{1.2.1}$$

and $\det A = ad - bc \neq 0$, then $A = \tilde{Q}\tilde{R}$, where

$$\tilde{Q} = \frac{1}{\sqrt{a^2+c^2}} \begin{pmatrix} a & -c \\ c & a \end{pmatrix} \in SO(2) \tag{1.2.2}$$

and

$$\tilde{R} = \begin{pmatrix} \sqrt{a^2+c^2} & \frac{ab+cd}{\sqrt{a^2+c^2}} \\ 0 & \frac{ad-bc}{\sqrt{a^2+c^2}} \end{pmatrix}. \tag{1.2.3}$$

In this setting, we consider a 2×2 matrix of independent Brownian motions and time-change it to produce a Markov process with the property that if the determinant is positive at time 0, then it stays positive at all times. This ensures that uniqueness of the QR-factorization holds at all times and also that the time-changed process falls into the setting of [Lia09].

Following the notation of [Lia09], we consider the following set-up.

1. Let X be the manifold of 2×2 matrices over $\mathbb{R}$ with strictly positive determinant equipped with the topology it inherits as an open subset of $\mathbb{R}^{2\times 2} \cong \mathbb{R}^4$.

2. Let K be the Lie group $SO(2)$ of 2×2 orthogonal matrices with determinant 1. This group acts on X by $A \mapsto Q^{-1}A$ for $Q \in K$ and $A \in X$.

3. The quotient of X with respect to the action of K can, via the QR decomposition, be identified with the set Y of upper triangular 2×2 matrices with strictly positive diagonal entries.

4. The isotropy subgroup of K for an element $x \in X$ is, as usual, the subgroup $\{k \in K : kx = x\}$. Since every element of X is an invertible matrix, this subgroup is always the trivial group consisting of just the identity. In particular, this subgroup is the same for every y in the interior of Y, as required in [Lia09, pg 168]. We denote this subgroup by M.

5. Let $(x_t)_{t\geq 0}$ be the X-valued process that satisfies the stochastic differential equation (SDE)

$$dx_t = \begin{pmatrix} dx_t^{1,1} & dx_t^{1,2} \\ dx_t^{2,1} & dx_t^{2,2} \end{pmatrix} = \begin{pmatrix} f(x_t)\, dA_t^{1,1} & f(x_t)\, dA_t^{1,2} \\ f(x_t)\, dA_t^{2,1} & f(x_t)\, dA_t^{2,2} \end{pmatrix}, \quad x_0 \in X, \tag{1.2.4}$$

   where $A_t^{1,1}$, $A_t^{1,2}$, $A_t^{2,1}$, and $A_t^{2,2}$ are independent standard one-dimensional Brownian motions, and $f(x) := \frac{\det(x)}{\operatorname{tr}(x'x)+1}$ with det and tr denoting the determinant and the trace. We establish below that (1.2.4) has a unique strong solution and that this solution does indeed take values in X.

It follows from the QR decomposition that $x_t = Q_t T_t$, where, in the terminology of [Lia09], the "angular part" Q_t belongs to K and the "radial part" T_t belongs to Y. We will show that $(T_t)_{t\geq 0}$ is an autonomous diffusion on Y and that $Q_t = U_{R_t}$, where $(U_t)_{t\geq 0}$ is a Brownian motion on K and $(R_t)_{t\geq 0}$ is an increasing process adapted to the filtration generated by $(T_t)_{t\geq 0}$. However, we will establish that **it is not possible** to take the Brownian motion $(U_t)_{t\geq 0}$ to be independent of the process $(T_t)_{t\geq 0}$. This will contradict the claim of Theorem 4 of [Lia09] once we have also checked that the conditions of that result hold.

Note that if we consider f as a function on the space $\mathbb{R}^{2\times 2} \cong \mathbb{R}^4$ of all 2×2 matrices, then it has bounded partial derivatives, and hence it is globally Lipschitz continuous. Consequently, if we allow the initial condition in (1.2.4) to be an arbitrary element of $\mathbb{R}^{2\times 2}$, then the resulting

SDE has a unique strong solution (see, for example, [RW00, Ch 5, Thm 11.2]). Moreover, the resulting process is a Feller process on $\mathbb{R}^{2\times 2}$ (see, for example, [RW00, Ch 5, Thm 22.5]).

We now check that $(x_t)_{t\geq 0}$ actually takes values in X. That is, we show that if x_0 has positive determinant, then x_t also has positive determinant for all $t \geq 0$. It follows from Itô's Lemma that

$$[\det(x_\cdot)]_t = \int_0^t \operatorname{tr}(x_s' x_s) f^2(x_s)\, ds,$$

$$[\operatorname{tr}(x_\cdot' x_\cdot)]_t = \int_0^t 4\operatorname{tr}(x_s' x_s) f^2(x_s)\, ds,$$

and

$$[\det(x_\cdot), \operatorname{tr}(x_\cdot' x_\cdot)] = \int_0^t 4\det(x_s) f^2(x_s)\, ds.$$

Thus, $((\det(x_t), \operatorname{tr}(x_t' x_t)))_{t\geq 0}$ is a Markov process and there exist independent standard one-dimensional Brownian motions $(B_t^1)_{t\geq 0}$ and $(B_t^2)_{t\geq 0}$ such that

$$d\ \det(x_t) = \sqrt{\operatorname{tr}(x_t' x_t)} f(x_t)\, dB_t^1$$

and

$$\begin{aligned} d\operatorname{tr}(x_t' x_t) = &\frac{4\det(x_t) f(x_t)}{\sqrt{\operatorname{tr}(x_t' x_t)}}\, dB_t^1 + \sqrt{\frac{4\operatorname{tr}^2(x_t' x_t) - 16\det(x_t)^2}{\operatorname{tr}(x_t' x_t)}} f(x_t)\, dB_t^2 \\ &+ 4 f^2(x_t)\, dt. \end{aligned}$$

When we substitute for f, the above equations transform into

$$d\ \det(x_t) = \frac{\det(x_t)\sqrt{\operatorname{tr}(x_t' x_t)}}{\operatorname{tr}(x_t' x_t) + 1}\, dB_t^1$$

and

$$\begin{aligned} d\operatorname{tr}(x_t' x_t) = &\frac{4(\det(x_t))^2}{\sqrt{\operatorname{tr}(x_t' x_t)}(\operatorname{tr}(x' x) + 1)}\, dB_t^1 + \sqrt{\frac{4\operatorname{tr}^2(x_t' x_t) - 16\det(x_t)^2}{\operatorname{tr}(x_t' x_t)}} \frac{\det(x_t)}{\operatorname{tr}(x_t' x_t) + 1}\, dB_t^2 \\ &+ 4\left(\frac{\det(x_t)}{\operatorname{tr}(x_t' x_t) + 1}\right)^2 dt. \end{aligned}$$

In particular, the process $(\det(x_t))_{t\geq 0}$ is the stochastic exponential of the local martingale $(M_t)_{t\geq 0}$, where

$$M_t = \int_0^t \frac{\sqrt{\operatorname{tr}(x_s' x_s)}}{\operatorname{tr}(x_s' x_s) + 1}\, dB_s^1.$$

Since $x_0 \in X$, we have $\det(x_0) > 0$, and hence

$$\det(x_t) = \det(x_0) \exp\left(M_t - M_0 - \frac{1}{2}[M]_t\right)$$

is strictly positive for all $t \geq 0$. This shows that $(x_t)_{t\geq 0}$ takes values in X.

We now check that $(x_t)_{t\geq 0}$ satisfies all the assumptions of [Lia09, Theorem 4]. These are as follows:

1. The process $(x_t)_{t\geq 0}$ is a Feller process with continuous sample paths.

2. The distribution of $(x_t)_{t\geq 0}$ is equivariant under the action of K. That is, for $k \in K$ the distribution of $(kx_t)_{t\geq 0}$ when $x_0 = x_*$ is the same as the distribution of $(x_t)_{t\geq 0}$ when $x_0 = kx_*$ [Lia09, (2)].

3. The set Y is a submanifold of X that is transversal to the action of K [Lia09, (3)].

4. For any $y \in Y^0$ (that is, the relative interior of Y – which in this case is just Y itself) T_yX, the tangent space of X at y, is the direct sum of tangent spaces $T_y(Ky) \bigoplus T_yY$ [Lia09, (5)].

5. The homogeneous space K/M is irreducible; that is, the action of M on $T_o(K/M)$ (the tangent space at the coset o containing the identity) has no nontrivial invariant subspace [Lia09, pg 177].

The verifications of (1)–(5) proceed as follows:

1. We have already observed that solutions of (1.2.4) with initial conditions in $\mathbb{R}^{2\times 2}$ form a Feller process and that this process stays in the open set X if it starts in X, and so $(x_t)_{t\geq 0}$ is a Feller process on X.

2. Suppose that $(x_t)_{t\geq 0}$ is a solution of (1.2.4) with $x_0 = x_*$ and $(\hat{x}_t)_{t\geq 0}$ is a solution of (1.2.4) with $\hat{x}_0 = kx_*$ for some $k \in K$. We have to show that if we set $\tilde{x}_t = k^{-1}\hat{x}_t$, then $(\tilde{x}_t)_{t\geq 0}$ has the same distribution as $(x_t)_{t\geq 0}$. Note that $\det \tilde{x}_t = \det \hat{x}_t$ and $\tilde{x}_t'\tilde{x}_t = \hat{x}_t'\hat{x}_t$, so that $f(\tilde{x}_t) = f(\hat{x}_t)$. Thus,

$$d\tilde{x}_t = f(\tilde{x}_t)k^{-1}\begin{pmatrix} dA_t^{1,1} & dA_t^{1,2} \\ dA_t^{2,1} & dA_t^{2,2} \end{pmatrix}, \quad \tilde{x}_0 = x_*.$$

Now the columns of the matrix

$$\begin{pmatrix} A_t^{1,1} & A_t^{1,2} \\ A_t^{2,1} & A_t^{2,2} \end{pmatrix}$$

are independent standard two-dimensional Brownian motions, and so the same is true of the columns of the matrix

$$k^{-1}\begin{pmatrix} A_t^{1,1} & A_t^{1,2} \\ A_t^{2,1} & A_t^{2,2} \end{pmatrix}$$

by the equivariance of standard two-dimensional Brownian motion under the action of $SO(2)$. Hence,

$$k^{-1}\begin{pmatrix} A_t^{1,1} & A_t^{1,2} \\ A_t^{2,1} & A_t^{2,2} \end{pmatrix} = \begin{pmatrix} \alpha_t^{1,1} & \alpha_t^{1,2} \\ \alpha_t^{2,1} & \alpha_t^{2,2} \end{pmatrix},$$

where $(\alpha_t^{1,1})_{t\geq 0}$, $(\alpha_t^{1,2})_{t\geq 0}$, $(\alpha_t^{2,1})_{t\geq 0}$, and $(\alpha_t^{2,2})_{t\geq 0}$ are independent standard Brownian motions. Since,

$$d\tilde{x}_t = f(\tilde{x}_t)\begin{pmatrix} d\alpha_t^{1,1} & d\alpha_t^{1,2} \\ d\alpha_t^{2,1} & d\alpha_t^{2,2} \end{pmatrix}, \quad \tilde{x_0} = x_0,$$

the existence and uniqueness of strong solutions to (1.2.4) establishes that the distributions of $(x_t)_{t\geq 0}$ and $(\tilde{x}_t)_{t\geq 0}$ are equal.

3. It follows from the existence of the QR decomposition for invertible matrices that X is the union of the orbits Ky for $y \in Y$, and it follows from the uniqueness of the decomposition for such matrices that the orbit Ky intersects Y only at y.

4. Since the tangent space of $K = SO(2)$ at the identity is the vector space of 2×2 skew-symmetric matrices and the tangent space of Y at the identity is the vector space of 2×2 upper-triangular matrices, we have to show that if W is a fixed invertible upper-triangular 2×2 matrix and M is a fixed 2×2 matrix, then

$$M = SW + V$$

for a unique skew-symmetric 2×2 matrix S and unique upper-triangular 2×2 matrix V. Let

$$M := \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \quad \text{and} \quad W := \begin{pmatrix} w_{11} & w_{12} \\ 0 & w_{22} \end{pmatrix}.$$

It is immediate that

$$S = \begin{pmatrix} 0 & -\frac{m_{21}}{w_{11}} \\ \frac{m_{21}}{w_{11}} & 0 \end{pmatrix}$$

and

$$V = \begin{pmatrix} m_{11} & \frac{m_{12}w_{11}+m_{21}w_{22}}{w_{11}} \\ 0 & \frac{m_{22}w_{11}-m_{21}w_{12}}{w_{11}} \end{pmatrix}.$$

5. We have already noted that the tangent space of K at the identity is the vector space of skew-symmetric 2×2 matrices. This vector space is one-dimensional and so this condition holds trivially.

We have now shown that $(x_t)_{t\geq 0}$ satisfies all the hypotheses of [Lia09, Theorem 4]. However, we have the following result.

Proposition 1.2.1. *In the decomposition $x_t = Q_t T_t$ the Y-valued process $(T_t)_{t\geq 0}$ is Markov and the K-valued process $(Q_t)_{t\geq 0}$ may be written as $Q_t = U_{R_t}$, where $(U_t)_{t\geq 0}$ is a K-valued Brownian motion and $(R_t)_{t\geq 0}$ is an increasing continuous process such that $R_0 = 0$ and $R_t - R_s$ is $\sigma\{T_u : s \leq u \leq t\}$-measurable for $0 \leq s < t < \infty$. However, there is no such representation in which $(T_t)_{t\geq 0}$ and $(U_t)_{t\geq 0}$ are independent.*

Proof. For all $t \geq 0$ we have $x_t = Q_t T_t$, where

$$Q_t = \frac{1}{\sqrt{(x_t^{11})^2 + (x_t^{21})^2}} \begin{pmatrix} x_t^{11} & -x_t^{21} \\ x_t^{21} & x_t^{11} \end{pmatrix} \in K$$

and

$$T_t = \begin{pmatrix} \sqrt{(x_t^{11})^2 + (x_t^{21})^2} & \frac{x_t^{11}x_t^{12} + x_t^{21}x_t^{22}}{\sqrt{(x_t^{11})^2 + (x_t^{21})^2}} \\ 0 & \frac{\det(x_t)}{\sqrt{(x_t^{11})^2 + (x_t^{21})^2}} \end{pmatrix} \in Y.$$

Note that $\det(x_t) = \det(T_t)$ and $\operatorname{tr}(x_t' x_t) = \operatorname{tr}(T_t' T_t)$, and so $f(x_t) = f(T_t)$. Note also that the complex-valued process $(x_t^{11} + i x_t^{21})_{t\geq 0}$ is an isotropic complex local martingale in the sense of [Kal02, Ch 18], that is

$$[x^{11}] = [x^{21}]$$

and

$$[x^{22}, x^{21}] = 0.$$

In our case

$$d[x^{11}]_t = d[x^{21}]_t = f^2(T_t)\, dt.$$

By [Kal02, Thm 18.5], $(\log(x_t^{11} + i x_t^{21})_{t\geq 0}$ is a well-defined isotropic complex local martingale that can be written as

$$\log(x_t^{11} + i x_t^{21}) = \log\left(T_t^{11}\right) + i\theta_t,$$

where

$$d[\theta]_t = d[\log(T^{11})]_t = \frac{1}{(T_t^{11})^2} d[x^{11}]_t = \left(\frac{f(T_t)}{T_t^{11}}\right)^2 dt.$$

By the classical result of Dambis, Dubins and Schwarz (see, for example, [Kal02, Thm 18.4]), there exists a standard complex Brownian motion $(\tilde{B}_t + iB_t)_{t\geq 0}$ such that $\log(x_t^{11} + i x_t^{21}) = \tilde{B}_{R_t} + iB_{R_t}$, where

$$R_t = \int_0^t \left(\frac{f(T_s)}{T_s^{11}}\right)^2 ds, \quad t \geq 0.$$

So, $\theta_t = B_{R_t}$ and $\log(T_t^{11}) = \tilde{B}_{R_t}$. Hence,

$$\frac{x_t^{11} + i x_t^{21}}{\sqrt{(x_t^{11})^2 + (x_t^{21})^2}} = (\cos(\theta_t) + i \sin(\theta_t))$$

and

$$Q_t = \begin{pmatrix} \cos(B_{R_t}) & -\sin(B_{R_t}) \\ \sin(B_{R_t}) & \cos(B_{R_t}) \end{pmatrix}.$$

Consequently, $Q_t = U_{R_t}$, where

$$U_t = \begin{pmatrix} \cos(B_t) & -\sin(B_t) \\ \sin(B_t) & \cos(B_t) \end{pmatrix},$$

and $(B_t)_{t\geq 0}$ is a standard one-dimensional Brownian motion.

Note that $(U_t)_{t\geq 0}$ is certainly a Brownian motion on $K = SO(2)$, and so we have uniquely identified the K-valued Brownian motion $(U_t)_{t\geq 0}$ and the increasing process $(R_t)_{t\geq 0}$ that appear in the claimed decomposition of $(x_t)_{t\geq 0}$.

To complete the proof, it suffices to suppose that $(U_t)_{t\geq 0}$ is independent of $(T_t)_{t\geq 0}$ and obtain a contradiction. An application of Itô's Lemma shows that the entries of $(U_t)_{t\geq 0}$ satisfy the system of SDEs

$$\begin{aligned}
dU_t^{1,1} &= -U_t^{2,1}\, dB_t - \frac{1}{2} U_t^{1,1}\, dt \\
dU_t^{2,1} &= U_t^{1,1}\, dB_t - \frac{1}{2} U_t^{2,1}\, dt \\
dU_t^{1,2} &= -U_t^{1,1}\, dB_t + \frac{1}{2} U_t^{2,1}\, dt = -dU_t^{2,1} \\
dU_t^{2,2} &= -U_t^{2,1}\, dB_t - \frac{1}{2} U_t^{1,1}\, dt = dU_t^{1,1}.
\end{aligned}$$

We apply Proposition 1.2.2 below to each of the four SDEs in the system describing $(U_t)_{t\geq 0}$, with, in the notation of that result, (ζ_t, H_t, K_t) being the respective triples $(U_t^{1,1}, U_t^{2,1}, U_t^{1,1})$, $(U_t^{2,1}, U_t^{1,1}, U_t^{2,1})$, $(U_t^{1,2}, U_t^{1,1}, U_t^{2,1})$, and $(U_t^{2,2}, U_t^{2,1}, U_t^{1,1})$. In each of the four applications, we let

- $(\mathcal{F}_t)_{t\geq 0}$ be the filtration generated by $(U_t)_{t\geq 0}$,
- $(\mathcal{G}_t)_{t\geq 0}$ be the filtration generated by $(T_t)_{t\geq 0}$,
- $\beta_t = B_t$,
- $\rho_t = R_t$,
- $J_t = \left(\frac{f(T_t)}{T_t^{11}}\right)^2$,

- $\gamma_t = W_t = \int_0^t \sqrt{\frac{1}{R'_s}} dB_{R_s}$.

Let $\mathcal{H}_t = \mathcal{F}_{\rho_t} \vee \mathcal{G}_t$, $t \geq 0$, as in the Proposition 1.2.2. It follows by the assumed independence of $(U_t)_{t\geq 0}$ and $(T_t)_{t\geq 0}$, part (iii) of Proposition 1.2.2, and equation (1.2.5) that the entries of the time-changed process $Q_t = U_{R_t}$ satisfy the system of SDEs

$$\begin{aligned}
dQ_t^{1,1} &= -Q_t^{2,1}\sqrt{R'_t}\,dW_t - \frac{1}{2}Q_t^{1,1}R'_t\,dt = -Q_t^{2,1}\frac{f(T_t)}{T_t^{11}}\,dW_t - \frac{1}{2}Q_t^{1,1}\left(\frac{f(T_t)}{T_t^{11}}\right)^2 dt \\
dQ_t^{2,1} &= Q_t^{1,1}\sqrt{R'_t}\,dW_t - \frac{1}{2}Q_t^{2,1}R'_t\,dt = Q_t^{1,1}\frac{f(T_t)}{T_t^{11}}\,dW_t - \frac{1}{2}Q_t^{2,1}\left(\frac{f(T_t)}{T_t^{11}}\right)^2 dt \\
dQ_t^{1,2} &= -dQ_t^{2,1} = Q_t^{1,1}\sqrt{R'_t}\,dW_t - \frac{1}{2}Q_t^{2,1}R'_t\,dt = Q_t^{1,1}\frac{f(T_t)}{T_t^{11}}\,dW_t - \frac{1}{2}Q_t^{2,1}\left(\frac{f(T_t)}{T_t^{11}}\right)^2 dt \\
dQ_t^{2,2} &= dQ_t^{1,1} = -Q_t^{2,1}\sqrt{R'_t}\,dW_t - \frac{1}{2}Q_t^{1,1}R'_t = -Q_t^{2,1}\frac{f(T_t)}{T_t^{11}}\,dW_t - \frac{1}{2}Q_t^{1,1}\left(\frac{f(T_t)}{T_t^{11}}\right)^2 dt.
\end{aligned}$$

Set

$$\begin{aligned}
dw_t^1 &= \frac{x_t^{11}}{\sqrt{(x_t^{11})^2 + (x_t^{21})^2}}\,dA_t^{11} + \frac{x_t^{21}}{\sqrt{(x_t^{11})^2 + (x_t^{21})^2}}\,dA_t^{21} \\
dw_t^2 &= \frac{-x_t^{21}}{\sqrt{(x_t^{11})^2 + (x_t^{21})^2}}\,dA_t^{11} + \frac{x_t^{11}}{\sqrt{(x_t^{11})^2 + (x_t^{21})^2}}\,dA_t^{21} \\
dw_t^3 &= \frac{x_t^{11}}{\sqrt{(x_t^{11})^2 + (x_t^{21})^2}}\,dA_t^{12} + \frac{x_t^{21}}{\sqrt{(x_t^{11})^2 + (x_t^{21})^2}}\,dA_t^{22} \\
dw_t^4 &= \frac{-x_t^{21}}{\sqrt{(x_t^{11})^2 + (x_t^{21})^2}}\,dA_t^{12} + \frac{x_t^{11}}{\sqrt{(x_t^{11})^2 + (x_t^{21})^2}}\,dA_t^{22}.
\end{aligned}$$

The processes $(w_t^i)_{t\geq 0}$ are local martingales with $[w_t^i, w_t^j]_t = \delta_{ij}t$, and thus they are independent standard Brownian motions. An application of Itô's Lemma shows that $(T_t)_{t\geq 0}$ is a diffusion satisfying the following system of SDEs.

$$\begin{aligned}
dT_t^{11} &= f(T_t)\,dw_t^1 + \frac{f^2(T_t)}{T_t^{11}}\,dt \\
dT_t^{12} &= \frac{T_t^{22}f(T_t)}{T_t^{11}}\,dw_t^2 + f(T_t)dw_t^3 - \frac{T_t^{12}f^2(T_t)}{2(T_t^{11})^2}\,dt \\
dT_t^{22} &= \frac{T_t^{12}f(T_t)}{T_t^{11}}\,dw_t^2 + f(T_t)dw_t^4 - \frac{T_t^{22}f^2(T_t)}{2(T_t^{11})^2}\,dt.
\end{aligned}$$

The assumed independence of the processes $(U_t)_{t\geq 0}$ and $(T_t)_{t\geq 0}$ and part (iv) of Proposition 1.2.2 give that $[Q^{i,j}, T^{k,l}] \equiv 0$ for all i, j, k and l. It follows from Itô's Lemma that

$$d(Q_tT_t)^{1,1} \;=\; dN_t + \frac{Q_t^{1,1} f^2(T_t)}{T_t^{1,1}} \left(1 - \frac{1}{2T_t^{1,1}}\right) dt,$$

where $(N_t)_{t\geq 0}$ is a continuous local martingale for the filtration $(\mathcal{H}_t)_{t\geq 0}$. This, however, is not possible because $(Q_tT_t)^{1,1} = x_t^{1,1}$ and the process $(x_t^{1,1})_{t\geq 0}$ is a continuous local martingale for the filtration $(\mathcal{H}_t)_{t\geq 0}$.

□

We required the following proposition that collects together some simple facts about time-changes.

Proposition 1.2.2. *Consider two filtrations $(\mathcal{F}_t)_{t\geq 0}$ and $(\mathcal{G}_t)_{t\geq 0}$ on an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Set $\mathcal{F}_\infty = \bigvee_{t\geq 0} \mathcal{F}_t$ and $\mathcal{G}_\infty = \bigvee_{t\geq 0} \mathcal{G}_t$. Assume that the sub-$\sigma$-fields $\mathcal{F}_\infty$ and $\mathcal{G}_\infty$ are independent. Suppose that*

$$\zeta_t = \zeta_0 + \int_0^t H_s \, d\beta_s + \int_0^t K_s \, ds,$$

where ζ_0 is $\mathcal{F}_0$-measurable, the integrands $(H_t)_{t\geq 0}$ and $(K_t)_{t\geq 0}$ are $(\mathcal{F}_t)_{t\geq 0}$-adapted, and $(\beta_t)_{t\geq 0}$ is an $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion. Suppose further that $\rho_t = \int_0^t J_s \, ds$, where $(J_t)_{t\geq 0}$ is a nonnegative, $(\mathcal{G}_t)_{t\geq 0}$-adapted process such that ρ_t is finite for all $t \geq 0$ almost surely. For $t \geq 0$ put

$$\mathcal{F}_{\rho_t} = \sigma\{L_{s\wedge\rho_t} : s \geq 0 \text{ and } L \text{ is } (\mathcal{F}_t)_{t\geq 0}\text{-optional}\}.$$

Set $\mathcal{H}_t = \mathcal{F}_{\rho_t} \vee \mathcal{G}_t$, $t \geq 0$. Then the following hold.

(i) The process $(\beta_{\rho_t})_{t\geq 0}$ is a continuous local martingale for the filtration $(\mathcal{H}_t)_{t\geq 0}$ with quadratic variation $[\beta_{\rho_\cdot}]_t = \rho_t$.

(ii) The process $(\gamma_t)_{t\geq 0}$, where

$$\gamma_t = \int_0^t \sqrt{\frac{1}{J_s}} \, d\beta_{\rho_s},$$

is a Brownian motion for the filtration $(\mathcal{H}_t)_{t\geq 0}$.

(iii) If $\xi_t = \zeta_{\rho_t}$, $t \geq 0$, then

$$\xi_t = \xi_0 + \int_0^t H_{\rho_s} \sqrt{J_s} \, d\gamma_s + \int_0^t K_{\rho_s} J_s \, ds.$$

(iv) If $(\eta_t)_{t\geq 0}$ is a continuous local martingale for the filtration $(\mathcal{G}_t)_{t\geq 0}$, then $(\eta_t)_{t\geq 0}$ is also a continuous local martingale for the filtration $(\mathcal{H}_t)_{t\geq 0}$ and $[\eta, \gamma] \equiv 0$.

Remark 1.2.3. The apparent counterexample to [Lia09, Theorem 4] arises because K/M is one-dimensional and hence trivially irreducible. When K/M has dimension greater than 1, irreducibility implies the nonexistence of a nonzero M-invariant tangent vector and it is this latter property that is actually used in the proof of [Lia09, Theorem 4]. However, in our setting M is the trivial group $\{\mathrm{Id}\}$ and every nonzero 2×2 skew-symmetric matrix is M-invariant, even though K/M is irreducible.

Chapter 2

A Lorentz model with variable density in a conservative force field

2.1 Introduction

We study a Lorentz gas-type model of a (tracer) particle moving in a conservative force field with a large number of scatterers distributed randomly in space. We suppose the particle has mass $m > 0$ and moves in $\mathbb{R}^2$ under the influence of a force field with spherically symmetric potential energy $\mathcal{U}(r)$. The particle has fixed total energy E and evolves along a path determined by $\mathcal{U}$ until at random 'reflection' times the particle undergoes jumps in velocity. The reflections leave the speed of the particle unchanged but assign the particle a new outgoing direction according to a uniform distribution on the unit circle S^1. We look at scaling limits (diffusion approximations) of the trajectory of the particle and show how in the limit we get a diffusion whose generator we can find as a function of the parameters of the model.

Our work can be seen as a significant generalization of [BR14] where the authors study a particle moving in a constant gravitational field with a large number of infinitely-small scatterers placed along the particle's trajectory according to a Poisson point process with variable density. The particle moves along a parabola until it hits a 'heavy' particle (scatterer) and then it reflects uniformly with the same absolute velocity but with a reflection angle that is an independent uniform random variable. The authors of [BR14] study scaling limits of the trajectory of the particle. They show that the scaling limits are diffusions whose generators can be explicitly written down. We extend the results of [BR14] from constant forces to any radially symmetric conservative force satisfying some mild assumptions.

A model that is related to the one from[BR14] but where the scatterer density is constant can be found in [RT99]. The model from [BR14] is heuristically similar to the model known to physicists as the Galton board: a particle moving under a constant external force and bouncing off a periodic array of convex domains (scatterers). In [CD09] it is shown that the scaling limit for the trajectory of a ball in a Galton board is a diffusion that is recurrent.

These models can be seen as specific types of a random evolution, in which a system changes its law of motion because of random changes in the environment. Some physical examples of random evolutions are: a radio signal propagating through a turbulent medium in which the index of refraction is changing at random and a population of bacteria evolving in an environment that is subject to random fluctuations. See [Her74] for an overview of random evolutions.

Transport processes are a type of random evolution that model the motion of a particle whose velocity undergoes jumps of random size at random times. Diffusion approximations of transport processes in smooth domains and with compact velocity state space are studied in [Pap75] and [BPL76]. Even though the models considered in [Pap75] are very general, the coefficients of the generator of the limiting diffusion are only expressed in terms of infinite series - there are no closed form formulas. In our model we allow for unbounded velocities and we are also able to find the generator of the diffusion approximation explicitly.

General transport processes are analyzed in [Cos91]. The author studies a particle that is moving in a piecewise smooth domain of $\mathbb{R}^d$. In the interior of the domain the particle moves under the influence of a potential U; at random exponentially distributed times it changes velocity, according to a probability distribution which can depend on both the current position and velocity; when the particle hits the boundary it reflects physically, that is, the angle of reflection equals the angle of incidence. It is shown that under some assumptions and after an appropriate rescaling, the position of the particle converges to a reflecting diffusion process whose coefficients can be identified as functions of the potential U. A model where the reflections (both in the interior and on the boundary) are more general but there is no potential ($U = 0$) is studied in [CK06]. Our model cannot be analyzed in the framework of [Cos91] because some of the assumptions from [Cos91] are not verified. See Remark 2.1.2 for more details.

2.1.1 The Model

We start by defining our model more rigorously in order to be able to present our assumptions and results.

Throughout let $\mathbb{R}_+ := [0, \infty)$, $\mathbb{N}_0 := \{0, 1, 2, \dots\} = \mathbb{N} \cup \{0\}$, and $\mathcal{C}^k(S)$ be the set of real valued functions on S having k continuous derivatives.

As before, we suppose we have a particle with mass $m > 0$ that travels in $\mathbb{R}^2$ under the influence of a spherically symmetric force field with potential energy $\mathcal{U}(r)$. The particle is assumed to have fixed total energy E and evolves along a path determined by $\mathcal{U}$ until at random exponentially distributed 'reflection' times the particle undergoes jumps in velocity. The reflections leave the speed of the particle unchanged but assign the particle a new outgoing direction according to a uniform distribution on the unit circle S^1.

Let $((R(t), A(t)), t \geq 0)$ denote the polar-coordinate trajectory of this particle. We first introduce some definitions which we will use in the construction of $((R(t), A(t)), t \geq 0)$.

- $v(r)$ is the speed of a particle with mass m and energy E in a conservative field $\mathcal{U}$ when the particle is at (r, α).
- $((r(r_0, \theta, t), \alpha(r_0, \alpha_0, \theta, t)), t \geq 0)$ denotes the trajectory of a particle with mass m, total energy E, initial position (r_0, α_0) and whose velocity vector at time 0 makes angle θ with the radial vector.
- T_k denotes the random time of the k'th reflection of the process $((R(t), A(t)), t \geq 0)$.

Total energy is conserved so we can write

$$\frac{mv^2(r)}{2} + \mathcal{U}(r) = E \tag{2.1.1}$$

where $v(r)$ is the speed of the particle as a function of the radial coordinate r of the particle. As a result,

$$v(r) = \sqrt{\frac{2}{m}(E - \mathcal{U}(r))}. \tag{2.1.2}$$

and by the spherical symmetry of $\mathcal{U}$, v is not a function of the angular trajectory α. We assume that the distribution of the time between consecutive reflections is given by

$$\mathbb{P}\left\{T_{k+1} - T_k > t + s \mid T_k = s\right\} = \exp\left(-\int_s^{t+s} g(r(u))v(r(u))du\right) \tag{2.1.3}$$

for any $k \in \mathbb{N}_0$ and for some $g : \mathbb{R}_+ \mapsto \mathbb{R}_+$ that is a function of the radial position of the particle only.

The motivation for the distribution of the reflection time is the following: after a reflection at (r_0, α_0) in direction θ the particle moves according to the potential $\mathcal{U}(r)$. On the trajectory of the particle there is a Poisson point process with variable intensity g per unit length. The points of this Poisson process represent other (heavy) particles in the gas, and g their density. After hitting the first of these points, our particle reflects and the process starts anew. We assume that no energy loss occurs in between reflections, so the total energy of the tracer particle remains E. We define T_1 as the hitting time of the first point of the Poisson process and then define $T_2, \ldots, T_k, \ldots$ recursively.

The trajectory process of the particle $((R(t), A(t)), t \geq 0)$ can be constructed piecewise as follows. Set $T_0 = 0$ and let the particle start at $(R(0), A(0))$ and move in the direction Θ_0. Then for any $k \in \mathbb{N}_0$ and $t \in [T_k, T_{k+1})$ we can write

$$\begin{aligned} R(t) &= r(R(T_k), A(T_k), \Theta_k, t - T_k) \\ A(t) &= \alpha(R(T_k), A(T_k), \Theta_k, t - T_k) \end{aligned} \tag{2.1.4}$$

where $(\Theta_k)_{k\in\mathbb{N}}$ are i.i.d. uniformly distributed on S^1.

2.1.2 Assumptions

Our main interest lies in deriving a diffusion approximation for the process

$$((R(t), A(t)), t \geq 0) .$$

We will both identify the limiting diffusion and prove convergence to this limit under the following assumptions.

(A1) To obtain a nontrivial diffusion (scaling) limit we rescale the density of the scatterers

$$g_n := \sqrt{n} g,$$

the potential energy of the field

$$\mathcal{U}_n := \frac{1}{\sqrt{n}} \mathcal{U}$$

and the total energy of the particle

$$E_n := \frac{1}{\sqrt{n}} E.$$

By (2.1.2) we can write the speed as a function of E and $\mathcal{U}$; consequently the speed v is also rescaled as

$$v_n(r) := \frac{1}{n^{1/4}} v(r).$$

The trajectory of the particle with these rescaled parameters we denote by

$$((r_n(r_0, \theta, t), \alpha_n(r_0, \alpha_0, \theta, t)), t \geq 0) .$$

When no confusion will arise, we will write $(r(t), \alpha(t))$ for $(r(r_0, \theta, t), \alpha(r_0, \alpha_0, \theta, t))$ and leave the dependence on r_0, α_0, and θ implicit. Likewise, the corresponding random trajectory process with these rescaled parameters we denote by $((R^n(t), A^n(t)), t \geq 0)$. That is, $((R^n(t), A^n(t)), t \geq 0)$ is constructed in the same way as the process $((R(t), A(t)), t \geq 0)$ but with the parameters g_n, $\mathcal{U}_n$ and E_n replacing g, $\mathcal{U}$ and E respectively. Similarly, the time of the k'th reflection for the rescaled process is denoted T_k^n.

(A2) We assume that for all $n \in \mathbb{N}$ the process $((R^n(t), A^n(t)), t \geq 0)$ evolves in a domain which is represented in *polar coordinates* by $\mathcal{D} \times \mathbb{R} \subset \mathbb{R}_+ \times \mathbb{R}$ or $\mathcal{D} = [h_-, h_+]$, where $0 \leq h_- < h_+ \leq \infty$. The domain $\mathcal{D}$ is chosen so that $E - U(r) > 0$ for all $r \in \mathcal{D}^\circ$. This is equivalent to $v(r) > 0$ for all $r \in \mathcal{D}^\circ$.

We require the particle to have positive speed in the interior $\mathcal{D}^\circ$ so the time between reflections approaches 0 as n goes to infinity. This way we obtain a nontrivial diffusion limit.

(A3) On the boundary $\partial\mathcal{D}$ we have the following assumptions.

- If $h_- > 0$ then

$$\mathcal{U}(h_-) = E$$
$$-\partial_r \mathcal{U}(h_-) := -\frac{\partial \mathcal{U}}{\partial r}(h_-) > 0$$

- If $h_+ < \infty$ then

$$\mathcal{U}(h_+) = E$$
$$-\partial_r \mathcal{U}(h_+) < 0.$$

The conditions on $\mathcal{U}$ force the speed of the particle to be zero at the respective endpoints h_- and h_+, while the conditions on $\frac{\partial \mathcal{U}}{\partial r}$ ensure the force field at the endpoints is pointing towards the interior $\mathcal{D}^\circ$. Having zero speed at the boundaries and having the force field point 'inwards' prevents the particle from leaving the domain $\mathcal{D}$.

- If $h_- = 0$ then

$$E - \mathcal{U}(0) > 0.$$

  This condition ensures the particle is not trapped at the origin.

- If $h_+ = \infty$ we require that for any $\varepsilon > 0$, $\inf_{[h_- + \epsilon, \infty)}(E - \mathcal{U}(r)) > 0$.

This condition implies that for any $\varepsilon > 0$, $\inf_{[h_- + \epsilon, \infty)} v(r) > 0$ which shows that the reflection rate does not go to 0 as the process goes to infinity.

(A4) $\mathcal{U} \in \mathcal{C}^1(\mathcal{D})$

This smoothness assumption ensures that the velocity and the acceleration of the particle depend continuously on the position $r(t)$. If $\mathcal{D} = [0, h]$ we can relax this condition to $\mathcal{U} \in \mathcal{C}^1((0, h])$ so that we can allow potentials of the form $\mathcal{U}(r) = -\frac{1}{r}$ which are not defined at 0.

(A5) The density g is spherically symmetric and satisfies

$$g \in \mathcal{C}(\mathcal{D}) \cap \mathcal{C}^1(\mathcal{D}^\circ)$$

together with

$$\inf_{r \in \mathcal{D}} g(r) > 0.$$

We require some smoothness from g because the diffusion limit we get depends on the derivative g'. The second assumption is needed because we do not want to have regions where the the reflection rate goes to 0. If one allows g to approach 0, then a different scaling may be required to obtain a diffusive limit when the process is started in these regions. Examples of such situations are dealt with in Section 8 of [BR14].

2.1.3 Results

The following Theorem is our main result.

Theorem 2.1.1. *Let $(\mathscr{R}, \mathscr{A})$ be a diffusion on $\mathcal{D} \times \mathbb{R}$ whose generator $\mathscr{G}$ acts on functions $f \in C^2(\mathcal{D} \times \mathbb{R})$ with compact support in $C^2(\mathcal{D}^\circ \times \mathbb{R})$ by*

$$\mathscr{G} f(\rho, a) = \frac{v(\rho)}{2g(\rho)} f_{\rho\rho}(\rho, a) + \frac{v(\rho)}{2g(\rho)\rho^2} f_{aa}(\rho, a) + \frac{v(\rho)}{g(\rho)} \left(-\frac{g'((\rho)}{2g(\rho)} + \frac{1}{2\rho} - \frac{\partial_r \mathcal{U}(\rho)}{2mv^2(\rho)} \right) f_\rho(\rho, a).$$

Suppose the process $((R^n(t), A^n(t)), t \geq 0)$ satisfies Assumptions (A1)-(A5) above. Fix $l, u \in \mathcal{D}^\circ$ with $l < u$ and define

$$\iota^n_{l,u} = \inf\{t \geq 0 : R^n(t) \geq u \text{ or } R^n(t) \leq l\}$$

and

$$\tau_{l,u} = \inf\{t \geq 0 : \mathscr{R}(t) \geq u \text{ or } \mathscr{R}(t) \leq l\}.$$

Then as $n \to \infty$ we have the following convergence in distribution

$$\left(\left(R^n(n^{3/4} t \wedge \iota^n_{l,u}), A^n(n^{3/4} t \wedge \iota^n_{l,u}) \right), t \geq 0 \right) \to \left(\left(\mathscr{R}(t \wedge \tau_{l,u}), \mathscr{A}(t \wedge \tau_{l,u}) \right), t \geq 0 \right).$$

Futhermore, if the left boundary of $\mathcal{D}$ is inaccessible then we can remove the stopping at l while if the right boundary of $\mathcal{D}$ is inaccessible we can remove the stopping at u.

Remark 2.1.2. The approach from [Cos91] is different from the approach we take. The author looks at the Markov process given by $(X(t), V(t))$ where $X(t) \in \mathbb{R}^d$ is the position of the particle and $V(t) \in \mathbb{R}^d$ is its velocity, while we just look at the position $X(t)$ (which is not a Markov process). As a result, our methods are different from the ones in [Cos91]. Our results can not be recovered from [Cos91] because our model does not satisfy all the underlying assumptions: The function Q^{-1} from [Cos91] has to be differentiable up to the boundary of the region where the motion takes place (see assumption (H2) from [Cos91]). In our case we have singularities at the origin and when $\mathcal{D} = [0, h]$ we also have singularities at the boundary of the domain. Furthermore, the speed of the particle cannot be unbounded in [Cos91], while there is no such restriction in our model.

Remark 2.1.3. From Theorem 2.1.1 it follows that up until the first hitting time of l or u, the rescaled radius of the particle's position $R^n(n^{3/4} t)$ converges as $n \to \infty$ to a diffusion $\mathcal{R}$ on $\mathcal{D}^\circ$ with generator $\mathscr{G}_r$ that acts on functions $f \in C^2(\mathcal{D})$ with compact support in $\mathcal{D}^\circ$ by

$$\mathscr{G}_r f(\rho) = \frac{v(\rho)}{2g(\rho)} f''(\rho) + \frac{v(\rho)}{g(\rho)} \left(-\frac{g'(\rho)}{2g(\rho)} + \frac{1}{2\rho} - \frac{\partial_r \mathcal{U}(\rho)}{4(E - \mathcal{U}(\rho))} \right) f'(\rho).$$

Also, the rescaled angle of the particle's position, $A^n(n^{3/4} t)$, converges as $n \to \infty$ to

$$B \left(\int_0^t \frac{v(\mathscr{R}(s))}{\mathscr{R}^2(s) g(\mathscr{R}(s))} \, ds \right)$$

where B is a 1-dimensional Brownian motion independent of $\mathscr{R}$. So the limiting process is a skew-product in the sense that the radial part evolves as a diffusion independent of the angular process, and the angular process evolves independently from the radial product except it evolves on a clock that is dependent on the radial process.

Remark 2.1.4. Assume $\mathcal{U} \equiv 0$ and the scatterer density is constant. If we normalize the parameters of that $g \equiv 1$ and $m/E = 2$, so that $v(r) \equiv 1$, then by Remark 2.1.3 the limiting radial process $\mathscr{R}$ is a 2-dimensional Bessel process, while the limiting angular process $\mathscr{A}$ is an independent Brownian motion with a time change $t \mapsto \int_0^t \frac{1}{\mathscr{R}^2(s)}\, ds$. Hence, as one might suspect, we recover the skew-product decomposition of 2 dimensional Brownian in polar coordinates.

Theorem 2.1.1 is proved in 2.4.1. The proof is broken into several steps. First, we study the skeleton process $((R_k^n, A_k^n), k \in \mathbb{N}_0)$ which is the Markov process that observes the process only at reflection times. That is,

$$((R_k^n, A_k^n), k \in \mathbb{N}_0) := ((R^n(T_k^n), A^n(T_k^n)), k \in \mathbb{N}_0) \tag{2.1.5}$$

where T_k^n denotes the time of the the k'th reflection for the process $((R^n(t), A^n(t))\,, t \geq 0)$. In Theorem 2.3.1 we prove the continuous time step process $\left(\left(R^n_{\lfloor \mathrm{nt} \rfloor}, A^n_{\lfloor \mathrm{nt} \rfloor}\right), t \geq 0\right)$ converges in distribution to a limiting diffusion $((\mathcal{R}_t, \mathcal{A}_t), t \geq 0)$. Next, we use the convergence of the step process to show convergence of the full trajectory $\left((R^n(n^{3/4}t), A^n(n^{3/4}t)), t \geq 0\right)$. This is done by first time changing the skeleton process so that we observe reflections at their real times (with a scaling factor $\frac{1}{n^{3/4}}$) rather than at the index of how many reflections have occurred. Then we push this time change through the limit to show that a time-changed version of the step process $\left(\left(R^n_{\lfloor \mathrm{nt} \rfloor}, A^n_{\lfloor \mathrm{nt} \rfloor}\right), t \geq 0\right)$ converges to $((\mathcal{R}_t, \mathcal{A}_t), t \geq 0)$ with a time change. Finally, by showing that the time-changed step process and the full trajectory $\left((R^n(n^{3/4}t), A^n(n^{3/4}t)), t \geq 0\right)$ agree at reflection times and stay close in between reflections, we can prove the convergence for the full trajectory from the convergence of the step process.

The rest of the paper is organized as follows. In Section 2.2 we prove preliminary results about the reflection times and the skeleton process. The results are technical lemmas that are necessary to prove the convergence of the skeleton process, which is the main goal of Section 2.3. In Section 2.4 we prove convergence of the skeleton process on its natural time scale and use this result to prove convergence of the full trajectory. Lastly, in Section 2.5 we discuss how to classify the boundaries of $\mathcal{D}$ so that we can remove the stopping in Theorem 2.1.1 at a boundary that is inaccessible.

2.2 Preliminaries

In this section we prove a series of technical lemmas that are used in the sequel to prove the convergence in distribution of the Markov process $((R_k^n, A_k^n), k \in \mathbb{N}_0)$ defined in (2.1.5).

Our proofs are complicated by the fact that for a general potential $\mathcal{U}$, it is not possible to explicitly find the trajectory $((r(t), \alpha(t)), t \geq 0)$. However, the following lemma provides local estimates for the radial trajectory $r(t)$.

Lemma 2.2.1. *For any* $r_0 \in \mathcal{D}^\circ$ *and* $\theta \in [-\pi, \pi]$

$$r(t) := r(r_0, \theta, t) = r_0 + v_0 \cos(\theta) \cdot t + \frac{1}{2}\left(\frac{-\partial_r \mathcal{U}(r(\tau))}{m} + \frac{v_0^2 r_0^2 \sin^2(\theta)}{r(\tau)^3}\right) t^2$$

for some $0 \leq \tau \leq t$ *depending on* r_0, θ *and* t.

Proof. By definition $\mathbf{r}(t) := (r(r_0, \theta, t), \alpha(r_0, \alpha_0, \theta, t))$ is the solution to the equations of motion in polar coordinates for a particle of mass m in the potential $\mathcal{U}$:

$$\frac{1}{r}\frac{d}{dt}(r^2 \dot{\alpha}) e_\alpha + (\ddot{r} - r\dot{\alpha}^2) e_r = -\frac{\partial_r \mathcal{U}(r)}{m} e_r \tag{2.2.1}$$

with initial conditions

$$\begin{aligned} \mathbf{r}(0) &= (r_0, \alpha_0) \\ \frac{d}{dt}\mathbf{r}(0) &= v(r_0)\cos\theta e_r + v(r_0)\sin\theta e_\alpha \end{aligned} \tag{2.2.2}$$

where e_r is the radial unit vector, e_α is the angular unit vector and θ is the angle the initial velocity $\frac{d}{dt}\mathbf{r}(0)$ makes with $\mathbf{r}(0)$. Let $\bar{v}(t) := v(r(t))$ denote the speed of the particle as a function of time. We can also write $\bar{v}(t)$ as a function of angular and radial velocity:

$$\bar{v}(t) = \sqrt{\dot{r}^2(t) + r^2(t)\dot{\alpha}^2(t)} \tag{2.2.3}$$

where

$$\dot{r}(t) = r'(t) := \frac{dr}{dt}(t)$$

is the radial velocity and

$$\dot{\alpha}(t) = \alpha'(t) := \frac{d\alpha}{dt}(t)$$

is the angular velocity. The initial conditions (2.2.2) become

$$\begin{aligned} \dot{r}(0) &= v(r_0)\cos\theta \\ \dot{\alpha}(0) &= \frac{v(r_0)}{r_0}\sin\theta. \end{aligned} \tag{2.2.4}$$

Equation (2.2.1) implies

$$\frac{1}{r}\frac{d}{dt}(r^2\dot{\alpha}) = 0 \tag{2.2.5}$$

and

$$\ddot{r} - r\dot{\alpha}^2 = -\frac{\partial_r \mathcal{U}(r)}{m}. \tag{2.2.6}$$

As a result of (2.2.5) and (2.2.4)

$$\dot{\alpha}(t) = \frac{v(r_0) r_0 \sin\theta}{r(t)^2}. \tag{2.2.7}$$

By (2.2.6) and (2.2.7)

$$\ddot{r}(t) = \left(\frac{-\partial_r \mathcal{U}(r(t))}{m} + \frac{v(r_0)^2 r_0^2 \sin^2(\theta)}{r(t)^3} \right). \tag{2.2.8}$$

Taylor expanding $r(t)$ and using (2.2.6), (2.2.4) together with (2.2.7) yields

$$r(t) = r_0 + v(r_0)\cos(\theta) \cdot t + \frac{1}{2} \left(\frac{-\partial_r \mathcal{U}(r(\tau))}{m} + \frac{v(r_0)^2 r_0^2 \sin^2(\theta)}{r(\tau)^3} \right) t^2. \tag{2.2.9}$$

□

Remark 2.2.2. Throughout the remainder of this section we let S, S' be closed intervals satisfying

$$S \subset (S')^\circ \subset S' \subset \mathcal{D}^\circ.$$

Also, for any $\delta > 0$, we define

$$\Lambda_\delta^n(S) = \inf_{(r_0,\theta) \in S \times [-\pi,\pi]} \inf_{t \geq 0} \{t : |r_n(r_0, \theta, t) - r_0| \geq \delta\} \tag{2.2.10}$$

to be the shortest time it takes the radial displacement of the particle to change by δ when started inside S.

If

$$0 < \delta < d(S, S')^c := \inf \{|x - y| : x \in S, y \in \mathcal{D} \setminus S'\},$$

then for all $(r_0, \theta, t) \in S \times [-\pi, \pi] \times [0, \Lambda_\delta^n(S)]$ one has

$$r_n(r_0, \theta, t) \in S'.$$

The radial speed $\dot{r}_n(t)$ can be bounded above by

$$\dot{r}_n(t) \leq \sqrt{\dot{r}_n^2(t) + r_n^2(t)\dot{\alpha}_n^2(t)} := \bar{v}_n(t) := v_n(r_n(t)).$$

Since $r_n(t) \in S'$ for all $t \in [0, \Lambda_\delta^n(S)]$, we can bound $\Lambda_\delta^n(S)$ below by δ divided by the maximum of the particle speed v_n inside the interval S'. Namely,

$$\Lambda_\delta^n(S) \geq n^{1/4} \cdot \frac{\delta}{\sup_{\rho \in S'} v(\rho)} > 0$$

where $\sup_{\rho \in S'} v(\rho) < \infty$ since S' is bounded away from ∂D.

The next lemma shows that, over a fixed time interval $[0,T]$, $r_n(r_0,\theta,t)$ converges uniformly to r_0 as a function of r_0,θ and t.

Lemma 2.2.3. *Fix $T>0$. Then*

$$\lim_{n\to\infty} \sup_{(r_0,\theta,t)\in S\times[-\pi,\pi]\times[0,T]} \{|r_n(r_0,\theta,t)-r_0|\} = 0$$

and

$$\lim_{n\to\infty} \sup_{(r_0,\theta,t)\in S\times[-\pi,\pi]\times[0,T]} \left\{\left|\sqrt{n}\ddot{r}_n(t,r_0,\theta) - \left(-\frac{\partial_r\mathcal{U}(r_0)}{m} + \frac{v^2(r_0)\sin^2(\theta)}{r_0}\right)\right|\right\} = 0. \quad (2.2.11)$$

Proof. Let $\delta>0$ such that $\delta \le d(S,(S')^c)$. By Remark 2.2.2 there exists $M\in\mathbb{N}$ large enough such that $\Lambda_\delta^n(S)\ge T$ whenever $n\ge M$. Equivalently, for $n\ge M$ we have

$$|r_n(r_0,\theta,t)-r_0|\le\delta$$

for all $(r_0,\theta,t)\in S\times[-\pi,\pi]\times[0,T]$. This proves the uniform convergence of $r_n(r_0,\theta,t)$ to r_0. Define

$$\psi(\rho,r_0,\theta) := \left(-\frac{\partial_r\mathcal{U}(\rho)}{m} + \frac{v^2(r_0)r_0^2\sin^2(\theta)}{\rho^3}\right) \quad (2.2.12)$$

and note by (2.2.8) that

$$\ddot{r}_n(t,r_0,\theta) = \frac{1}{\sqrt{n}}\psi(r_n(t),r_0,\theta).$$

S' is bounded away from 0 and $\mathcal{U}\in\mathcal{C}^1(\mathcal{D}^\circ)$ imply that ψ is uniformly continuous on $S'\times S\times[-\pi,\pi]$. By construction $r_n(t)\in S'$ for all $n\ge M$ and $t\in[0,T]$. Because $r_n(r_0,\theta,t)$ converges uniformly to r_0 on compact sets, we have

$$\lim_{n\to\infty} \sup_{(r_0,\theta,t)\in S\times[-\pi,\pi]\times[0,T]} \left\{\left|\sqrt{n}\ddot{r}_n(t,r_0,\theta)-\psi(r_0,r_0,\theta)\right|\right\} = 0 \quad (2.2.13)$$

where

$$\psi(r_0,r_0,\theta) = \left(-\frac{\partial_r\mathcal{U}(r_0)}{m} + \frac{v^2(r_0)\sin^2(\theta)}{r_0}\right)$$

by (2.2.12). This completes the proof. □

Since the skeleton process tracks the process at reflection times, we want to apply the estimates for (2.2.9) between reflections. By (2.1.3) and the rescaling, we know that for every k, $T_{k+1}^n - T_k^n$ is distributed like the random variable $N^{(n)}(r_0,\theta)$ which we define by

$$\mathbb{P}(N^{(n)}(r_0,\theta)>t) = \exp\left(-\int_0^t n^{1/4}\lambda(r_n(r_0,\theta,s))\,ds\right) \quad (2.2.14)$$

where

$$\lambda(\rho) := g(\rho)v(\rho). \tag{2.2.15}$$

Throughout, we will often suppress the r_0 and θ dependencies of $N^{(n)}(r_0, \theta)$ and write $N^{(n)}$ or $N^{(n)}(\theta)$ when no confusion will arise.

For many of our proofs we require estimates that show the time between reflections approaches 0 with high probability as the scaling factor n goes to infinity. This will allow us to apply the local estimates from the expansion of $r(t)$ in Lemma 2.2.1 and to show the skeleton process does not undergo large jumps. We first prove some bounds on the moments of $N^{(n)}(r_0, \theta)$.

Lemma 2.2.4. *The family*

$$\{n^{1/4}N^{(n)}(\rho, \theta) : (\rho, \theta, n) \in S \times \mathbb{R} \times \mathbb{N}\}$$

is bounded in L^p for $1 \leq p < \infty$.

Proof. We will assume that $\mathcal{D} = [0, h]$ or $\mathcal{D} = [0, \infty)$. The cases $\mathcal{D} = [h_-, h_+]$ and $[h, \infty)$ can be treated similarly.

Case I: $\mathcal{D} = [0, h]$

By Assumptions (A3) and (A4) there exist $\delta > 0$ and $m > 0$ such that

$$\min_{[h-\delta,h]} |\partial_r U(\rho)| \geq \delta. \tag{2.2.16}$$

Let $S \subset D^\circ$ be a compact set, let $\eta > 0$ and assume our particle enters the annulus A_η with inner radius $h - \eta$ and outer radius h at time t_0. By (2.1.2)

$$v(h - \eta) = \sqrt{\frac{2}{m}(E - \mathcal{U}(h - \eta))}. \tag{2.2.17}$$

Using (2.2.6) and (2.2.7)

$$\ddot{r}(t) = -\frac{\partial_r \mathcal{U}(r(t))}{m} + r(t)\dot{\alpha}^2(t) \leq -\frac{\partial_r \mathcal{U}(r(t))}{m} + \frac{[v(h-\eta)]^2}{r}. \tag{2.2.18}$$

Let

$$t_e := \inf\{s > t_0 \ : \ r(s) = h - \eta\}$$

be the time when the particle exists the annulus. By Assumption (A3)

$$\mathcal{U}(h) = E.$$

Since $\mathcal{U}$ is continuous this means that we can make $v(r)$ as small as we like if we are close enough to $r = h$. This together with (2.2.17), (2.2.18) and (2.2.16) implies that there exist $\gamma > 0, m_\gamma > 0$ such that

$$\ddot{r}(t) \leq -m_\gamma$$

whenever $r(t) \in [h-\gamma, h]$. Set $\eta = \gamma$.

We can now find an upper bound on the time $t_e - t_0$. Note that

$$|\dot{r}(t_0)| \leq v(h-\gamma).$$

Therefore $t_e - t_0$ is bounded above by the time it would take a particle started at $r = h-\gamma$ with speed $v(h-\gamma)$ pointed along the radius and with acceleration $\ddot{r} = -m_\gamma < 0$ to return to $r = h - \gamma$. Thus,

$$t_e - t_0 \leq 2\frac{v(h-\gamma)}{m_\gamma} < \infty. \tag{2.2.19}$$

Next, define

$$t_r := \inf\{s > t_e \ : \ r(s) = h-\gamma\}.$$

This is the first return time to the annulus A_γ. We want to bound $t_r - t_e$ below. If $\dot{r}(t_0) \leq 0$ then the particle would not spend any time in the annulus A_γ, that is $t_e - t_0 = 0$. Therefore, we can assume that $\dot{r}(t_0) > 0$. It is clear by conservation of angular momentum and conservation of energy that $\dot{r}(t_e) = -\dot{r}(t_0) < 0$. Since $\ddot{r}$ is finite we immediately get that

$$t_r - t_e \geq \frac{\dot{r}(t_0)}{\left|-\frac{\partial_r \mathcal{U}(r)}{m} + r(\dot{\alpha})^2\right|} \geq \frac{\dot{r}(t_0)}{\frac{\sup_S |\partial_r \mathcal{U}(r)|}{m} + \sup_S |r(\dot{\alpha})^2|} > 0.$$

Clearly, since $\mathcal{U} \in C^1(\mathcal{D}^\circ)$ by Assumption (A4), $t_r - t_e$ is a continuous function of the initial conditions (r_0, θ). Since $S \times [-\pi, \pi]$ is compact there exists $\varrho > 0$ such that

$$\inf_{S\times[-\pi,\pi]} \{t_r(r_0,\theta) - t_e(r_0,\theta)\} = \varrho > 0. \tag{2.2.20}$$

Combining (2.2.19) and (2.2.20),

$$\sup_{S\times[-\pi,\pi]} \frac{t_e - t_0}{t_r - t_0} \leq \frac{2v(h-\gamma)}{\varrho m_\gamma} < \infty \tag{2.2.21}$$

where we assume that if $t_0 = \infty$ then $t_e - t_0 = 0$ and

$$\frac{t_e - t_0}{t_r - t_e} = 0.$$

By Assumption (A2) we know that

$$\inf_S v(\rho) > 0.$$

Suppose that $t > 2\frac{v(h-\gamma)}{m_\gamma} + \rho \geq \sup(t_e - t_0) + \inf(t_r - t_e)$.
Using (2.2.21) together with the fact that the worst case scenario is when the particle spends

the longest possible time in the 'bad region' A_γ and the least amount of time in the 'good region' $\mathcal{D} \setminus A_\gamma$, we have

$$\begin{aligned}\mathbb{P}\{N^{(n)} > t\} &= \exp\left(-\int_0^t n^{1/4} g(r_n) v(r_n)\, ds\right) \\ &\leq \exp\left(-\int_0^{\frac{1}{1+\left(\frac{2v(h-\gamma)}{\varrho m_\gamma}\right)}t} n^{1/4} \inf_S g(r_n) \inf_S v(r_n)\, ds\right) \\ &= \exp\left(-\frac{1}{1+\left(\frac{2v(h-\gamma)}{\varrho m_\gamma}\right)} t n^{1/4} \inf_S g(r_n) \inf_S v(r_n)\right)\end{aligned}$$

which decays exponentially in n as $n \to \infty$ as long as t is large. Therefore,

$$\mathbb{E}\left[\left(n^{1/4} N^{(n)}\right)^p\right] = \int_0^\infty p t^{p-1} \mathbb{P}\left(n^{1/4} N^{(n)} > t\right)\, dt < \infty.$$

Case II: $\mathcal{D} = [0, \infty)$

Set $g_{\min} := \inf_{r \in \mathcal{D}} g(r)$. We know by Assumption (A2) that there exists $\bar{\delta} > 0$ such that $\inf_{\mathcal{D}} v(r) \geq \bar{\delta}$ so by (2.2.14)

$$\mathbb{P}\{N^{(n)} > t\} \leq \exp\left(-n^{1/4} \bar{\delta} t g_{\min}\right)$$

which, like before, forces

$$\mathbb{E}\left[\left(n^{1/4} N^{(n)}\right)^p\right] = \int_0^\infty p t^{p-1} \mathbb{P}\left(n^{1/4} N^{(n)} > t\right)\, dt < \infty.$$

□

Lemma 2.2.4 also provides us with the following corollary which will prove useful in showing that the probability that $N^{(n)}(\rho, \theta)$ is larger than any fixed value decays rapidly as the scaling parameter n goes to infinity.

Corollary 2.2.5. *For all $k \in \mathbb{R}$ and for all $\varepsilon > 0$,*

$$\lim_{n \to \infty} \sup_{(\rho,\theta) \in S \times [-\pi, \pi]} n^k \mathbb{P}\left\{N^{(n)}(\rho, \theta) \geq \varepsilon\right\} = 0$$

Proof. For any $k \in \mathbb{R}$,

$$n^k \mathbb{P}\{N^{(n)}(\rho, \theta) \geq \varepsilon\} \leq \frac{n^k}{\varepsilon^{4(k+1)}} \mathbb{E}[N^{(n)}(\rho, \theta)^{4(k+1)}] = \frac{1}{n} \cdot \frac{n^{k+1}}{\varepsilon^{4(k+1)}} \mathbb{E}[N^{(n)}(\rho, \theta)^{4(k+1)}] \to 0 \tag{2.2.22}$$

as $n \to \infty$ uniformly for $(\rho, \theta) \in S \times [-\pi, \pi]$ by the uniform bounds on $\{n^{k+1} \mathbb{E}[N^{(n)}(\rho, \theta)^{4(k+1)}]\}_{n=1}^\infty$ from Lemma 2.2.4.

□

In addition, we have also the following corollary that shows that tails of the moments of $N^{(n)}(\rho,\theta)$ decay rapidly as well.

Corollary 2.2.6. *For all $k \in \mathbb{R}, l \geq 1$ and $\varepsilon > 0$*

$$\lim_{n\to\infty} \sup_{(\rho,\theta)\in S\times[-\pi,\pi]} n^k \mathbb{E}\left[(N^n(\rho,\theta))^l \, \mathbb{1}_{\{N^{(n)}>\varepsilon\}} \right] = 0.$$

Proof. By Cauchy-Schwarz,

$$\left(n^k \mathbb{E}\left[(N^n(\rho,\theta))^l \, \mathbb{1}_{\{N^{(n)}>\varepsilon\}} \right]\right)^2 \leq n^{l/2}\mathbb{E}\left[\left(N^{(n)}(\rho,\theta)\right)^{2l}\right] \cdot n^{2k-l/2}\mathbb{P}\left\{N^{(n)}(\rho,\theta) > \varepsilon\right\} \to 0 \tag{2.2.23}$$

as $n \to \infty$ by Lemma 2.2.4 and Corollary 2.2.5.

□

From Lemma 2.2.4, we have the following estimates on the moments of $N^{(n)}$.

Lemma 2.2.7. *Let $1 \leq p < \infty$. Then*

$$\lim_{n\to\infty} \sup_{(\rho,\theta)\in S\times[-\pi,\pi]} \left| \mathbb{E}\left(n^{p/4}[N^{(n)}(\rho,\theta)]^p\right) - \int_0^\infty p t^{p-1} \exp\left(-g(\rho)v(\rho)t\right)\, dt \right| = 0. \tag{2.2.24}$$

Proof. First let $M \in \mathbb{R}_+$ and note that the truncated moments $\mathbb{E}\left[n^{p/4}N^{(n)}(\rho,\theta)^p\right] \wedge M^p]$ can be written as

$$\begin{aligned} \mathbb{E}\left[n^{p/4}N^{(n)}(\rho,\theta)^p \wedge M^p\right] &= p\int_0^\infty t^{p-1}\mathbb{P}\left[n^{1/4}N^{(n)}(\rho,\theta)\wedge M > t\right] dt \\ &= p\int_0^M t^{p-1}\mathbb{P}\left[n^{1/4}N^{(n)}(\rho,\theta) > t\right] dt. \end{aligned} \tag{2.2.25}$$

Making the change of variables $u = n^{1/4}s$, we have

$$\mathbb{P}\left(n^{1/4}N^{(n)}(\rho,\theta) > t\right) = \exp\left(-\int_0^t g(r(\rho,\theta,u/n^{1/4}))v(\rho,\theta,u/n^{1/4})\right) du. \tag{2.2.26}$$

Both $r = r(\rho,\theta,t)$ $v = v(\rho,\theta,t)$ are continuous functions on $\mathcal{O} := S\times[-\pi,\pi]\times[0,M]$. Since $\mathcal{O}$ is compact, r and v are in fact uniformly continuous on S. By Assumption (A5) g is continuous on S, and therefore is uniformly continuous. This implies that $g\circ r$ is uniformly continuous on $\mathcal{O}$. By Lemma 2.2.3 it follows that

$$\lim_{n\to\infty} g(r_n(\rho,\theta,u/n^{1/4})) = g(r(\rho,\theta,0)) = g(\rho) \tag{2.2.27}$$

and

$$\lim_{n\to\infty} v(\rho,\theta,u/n^{1/4}) = v(\rho,\theta,0) := v(\rho) \tag{2.2.28}$$

both uniformly on $\mathcal{O}$. As a result, $\mathbb{P}\left(n^{1/4}N^{(n)}(\rho,\theta) > t\right)$ converges uniformly to $\exp\left(-g(\rho)v(\rho)t\right)$ on $\mathcal{O}$, which implies

$$\lim_{n\to\infty} \sup_{(\rho,\theta)\in S\times[-\pi,\pi]} \left| \mathbb{E}\left[n^{p/4}N^{(n)}(\rho,\theta)^p \wedge M^p\right] - \int_0^M pt^{p-1}\exp\left(-g(\rho)v(\rho)t\right) \right| = 0. \quad (2.2.29)$$

To extend the result to the expectation without truncation, define q as the solution to

$$1/q + p/(p+1) = 1.$$

By Hölder's inequality,

$$\begin{aligned} &\mathbb{E}\left[n^{p/4}N^{(n)}(\rho,\theta)^p - \left(n^{p/4}N^{(n)}(\rho,\theta)^p \wedge M^p\right)\right] = \mathbb{E}\left[n^{p/4}N^{(n)}(\rho,\theta)^p \mathbb{1}_{\{n^{1/4}N(\rho,\theta)>M\}}\right] \\ &\leq \mathbb{E}\left[n^{(p+1)/4}N^{(n)}(\rho,\theta)^{p+1}\right]^{p/(p+1)} \mathbb{P}\left\{n^{1/4}N^{(n)}(\rho,\theta) > M\right\}^{1/q} \\ &\leq \mathbb{E}\left[n^{(p+1)/4}N^{(n)}(\rho,\theta)^{p+1}\right]^{p/(p+1)} \mathbb{E}\left[n^{1/4}N^{(n)}(\rho,\theta)\right]^{1/q} \frac{1}{M^{1/q}}. \end{aligned} \quad (2.2.30)$$

By Lemma 2.2.4, both of these moments are uniformly bounded in ρ and θ, so the bound goes to 0 uniformly in ρ and θ as $M \to \infty$ and $n \to \infty$.

Furthermore, since v is bounded away from 0 on S by Assumption (A2) and since g is bounded away from 0 on $\mathcal{D}$ by Assumption (A5), it follows that for every $\rho \in S$

$$\lim_{M\to\infty} \int_M^\infty pt^{p-1}\exp(-g(\rho)v(\rho)t) = 0$$

uniformly on S. The proof then follows from the truncated case.

□

For the cases $p = 1, 2$ we have by Lemma 2.2.7

$$\lim_{n\to\infty} \sup_{(\rho,\theta)\in S\times[-\pi,\pi]} \left| \mathbb{E}\left[n^{1/4}N^{(n)}(\rho,\theta)\right] - \frac{1}{g(\rho)v(\rho)} \right| = 0 \quad (2.2.31)$$

and

$$\lim_{n\to\infty} \sup_{(\rho,\theta)\in S\times[-\pi,\pi]} \left| \mathbb{E}\left[n^{1/2}\left(N^{(n)}(\rho,\theta)\right)^2\right] - \frac{2}{g^2(\rho)v^2(\rho)} \right| = 0. \quad (2.2.32)$$

The next result shows that on the event $\{N^{(n)} > \varepsilon\}$ the kth moment of the difference of the radius of the particle at time 0 and at the first reflection $N^{(n)}$ decays faster than $\frac{1}{n^m}$ as $n \to \infty$.

Lemma 2.2.8. *Fix $\varepsilon > 0$, $m, k \in \mathbb{N}$. Then*

$$\lim_{n\to\infty} \sup_{(r_0,\theta)\in S\times[-\pi,\pi]} n^m \mathbb{E}\left[\left|r_n\left(N^{(n)}(\theta)\right) - r_0\right|^k \mathbb{1}_{\{N^{(n)}>\varepsilon\}}\right] = 0.$$

Proof. For any $t \geq 0$ we have

$$|r_n(t) - r_0| = \left| \int_0^t \dot{r}_n(s)ds \right| \leq \int_0^t |\dot{r}_n(s)|ds \leq \int_0^t v_n(r_n(s))ds$$

where the last inequality holds because the radial velocity is always less than or equal to the total velocity. Hence

$$\mathbb{E}\left[\left| r_n(N^{(n)}(\Theta)) - r_0 \right|^k \mathbb{1}_{\{N^{(n)} > \varepsilon\}} \right] \leq \mathbb{E}\left[\left(\int_0^{N^{(n)}} v_n(r_n(s))ds \right)^k \mathbb{1}_{\{N^{(n)} > \varepsilon\}} \right]$$

By (2.2.14) we know that $N^{(n)}(r_0, \theta)$ has density function

$$p_N^n(t) := n^{1/4}\lambda(r_n(t)) \exp\left(- \int_0^t n^{1/4}\lambda(r_n(s))ds \right), \ t \geq 0.$$

so

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_0^{N^{(n)}} v_n(r_n(s))ds \right)^k \mathbb{1}_{\{N^{(n)} > \varepsilon\}} \right] = \\
&n^{1/4} \int_\varepsilon^\infty \left(\int_0^t v_n(r_n(s))ds \right)^k \exp\left(-n^{1/4} \int_0^t \lambda(r_n(s))ds \right) \lambda(r_n(t))dt \\
&\leq \frac{1}{g_{\min}^k} \cdot \frac{1}{n^{(k-1)/4}} \int_\varepsilon^\infty \left(\int_0^t \lambda(r_n(s))ds \right)^k \exp\left(-n^{1/4} \int_0^t \lambda(r_n(s))ds \right) \lambda(r_n(t))dt \\
&= \frac{1}{g_{\min}^k} \cdot \frac{1}{n^{(k-1)/4}} \int_{L_\varepsilon^n}^{L_\infty^n} u^k e^{-n^{1/4}u} du \leq \frac{1}{g_{\min}^k} \cdot \frac{1}{n^{(k-1)/4}} \int_{L_\varepsilon^n}^\infty u^k e^{-n^{1/4}u} du.
\end{aligned}$$

where $L_\varepsilon^n := \int_0^\varepsilon \lambda(r_n(s))ds$ and $L_\infty^n := \int_0^\infty \lambda(r_n(s))ds$. The dependence of L_ε^n and L_∞^n on r_0 and θ is implicit. Now to establish uniform bounds in r_0 and θ, we bound L_ε^n uniformly away from 0 as follows. By Lemma 2.2.3, $r_n(s)$ converges uniformly on $[0, \varepsilon]$. So if we fix S' as in the proof of Lemma 2.2.3, that is we choose S' compact so that $S \subseteq (S')^\circ$ and $S' \subseteq D^\circ$. Then for n large enough, $r_n(r_0, \theta, s) \in S'$ for all $(r_0, \theta, s) \in S \times [-\pi, \pi] \times [0, \varepsilon]$. By uniform continuity of v on S', for large n we can bound L_ε^n uniformly below as follows

$$L_\varepsilon^n = \int_0^\varepsilon g(r_n(s))v(r_n(s))ds \geq \varepsilon g_{\min} \inf_{\rho \in S'} v(\rho) := L_\varepsilon > 0$$

since v is bounded away from 0 on S' by Assumption (A2).

For $u > 0$, $e^{n^{1/4}u/2}$ dominates u^k as $n \to \infty$. As a result,

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{g_{\min}^k} \cdot \frac{1}{n^{(k-1)/4}} \int_{L_\varepsilon^n}^{\infty} u^k e^{-n^{1/4}u} du &\leq \lim_{n\to\infty} \frac{1}{g_{\min}^k} \cdot \frac{1}{n^{(k-1)/4}} \int_{L_\varepsilon}^{\infty} e^{-n^{1/4}u/2} du \\ &= \lim_{n\to\infty} \frac{1}{g_{\min}^k} \cdot \frac{2}{n^{k/4}} \left(e^{-\frac{n^{1/4}}{2} L_\varepsilon} \right) \\ &= 0.\end{aligned}$$

□

A consequence of Lemma 2.2.8 is the analogous result for the angular jumps.

Corollary 2.2.9. *Fix* $\varepsilon > 0$, $m, k \in \mathbb{N}$. *Then*

$$\lim_{n\to\infty} \sup_{(r_0,\theta,\alpha_0)\in S\times[-\pi,\pi]\times\mathbb{R}} n^m \mathbb{E}\left[\left| \alpha_n(r_0, \alpha_0, N^{(n)}(\theta)) - \alpha_0 \right|^k \mathbb{1}_{\{N^{(n)}>\varepsilon\}} \right] = 0.$$

Proof. By (2.2.3) we have the bounds

$$|\dot\alpha(t))r(t)| \leq \sqrt{\dot r^2(t) + r^2(t)\dot\alpha^2(t)} := v(r(t)).$$

Let $0 < \delta < \Lambda_\delta^n(S)$. For n large enough, we have $\Lambda_\delta^n(S) \leq T$, which implies $r_n(t) \in S'$ for all $t \in [0, T]$

$$|\dot\alpha(t))| \leq \frac{v(r(t))}{\inf S'}.$$

From this we have the bounds

$$|\alpha_n(t) - \alpha_0| = \left| \int_0^t \dot\alpha_n(s) ds \right| \leq \int_0^t |\dot\alpha_n(s)| ds \leq \frac{1}{\inf S'} \int_0^t v_n(r_n(s)) ds.$$

We can now apply the estimates from the proof of Lemma 2.2.8 unchanged save for multiplying by a factor of $\frac{1}{\inf S'}$.

□

The next lemma will be used to evaluate the radial drift of the limiting diffusion for the the skeleton process $((R_k^n, A_k^n), k \in \mathbb{N}_0)$.

Lemma 2.2.10.

$$\lim_{n\to\infty} \sup_{r_0\in S} \left| n^{3/4} \cdot \mathbb{E}\left[N^{(n)}(r_0, \Theta) \cos(\Theta) \right] - \frac{1}{2g^2(r_0)v(r_0)} \cdot \left(\frac{\partial_r \mathcal{U}(r_0)}{2(E - \mathcal{U}(r_0))} - \frac{g'(r_0)}{g(r_0)} \right) \right| = 0. \tag{2.2.33}$$

Proof. First, for notational convenience, define the auxiliary function

$$F_n(r_0, \theta, t) := \int_0^t g_n(r_n(r_0, \theta, s)) v_n(r_n(r_0, \theta, s)) ds. \tag{2.2.34}$$

In the usual way, we will often suppress the r_0 and θ dependencies of F and write $F(t)$ when no confusion will arise. If we let $M_n(r_0, \alpha_0, \theta, t)$ be the number of Poisson process points on the path $((r_n(r_0, \theta, s), \alpha_n(r_0, \alpha_0, \theta, s)), 0 \leq s \leq t)$, then standard facts about Poisson processes show that

$$(M_n(r_0, \alpha_0, \theta, t) - F_n(r_0, \theta, t), t \geq 0)$$

and

$$\left((M_n(r_0, \alpha_0, \theta, t) - F_n(r_0, \theta, t))^2 - F_n(r_0, \theta, t), t \geq 0\right)$$

are martingales. An optional stopping argument shows that

$$1 = \mathbb{E}[F_n(r_0, \theta, N^{(n)}(r_0, \theta))] \tag{2.2.35}$$

and

$$2 = \mathbb{E}[F_n(r_0, \theta, N^{(n)}(r_0, \theta))^2]. \tag{2.2.36}$$

See for example (3.1-2) of [BR14].

In order to prove uniform convergence, we need to work on a compact set, so we fix a time $T > 0$ and split the expectations as

$$1 = \mathbb{E}[F_n(r_0, \theta, N^{(n)}(\theta)) \mathbb{1}_{\{N^{(n)}(\theta) < T\}}] + \mathbb{E}[F_n(r_0, \theta, N^{(n)}(\theta)) \mathbb{1}_{\{N^{(n)}(\theta) \geq T\}}] \tag{2.2.37}$$

Taylor expanding $F_n(r_0, \theta, t)$ about $t = 0$ yields

$$F_n(r_0, \theta, t) = g_n(r_0) v_n(\rho) t + \frac{1}{2} \ddot{F}_n(r_0, \theta, \tau(t)) t^2 \tag{2.2.38}$$

for some $\tau(t) \in [0, t]$. Here $\ddot{F}_n$ denotes the second derivative with respect to time t. Then after setting $t = N^{(n)}(\theta))$, multiplying both sides by $\mathbb{1}_{\{N^{(n)}(\theta) < T\}}$ and taking expectations we have

$$\begin{aligned} &\mathbb{E}\left[F_n(r_0, \theta, N^{(n)}(\theta)) \mathbb{1}_{\{N^{(n)}(\theta) < T\}}\right] \\ &= g_n(r_0) v_n(r_0) \mathbb{E}\left[N^{(n)}(\theta) \mathbb{1}_{\{N^{(n)}(\theta) < T\}}\right] + \frac{1}{2} \mathbb{E}\left[\ddot{F}_n(\tau(\theta)) (N^{(n)})^2(\theta) \mathbb{1}_{\{N^{(n)}(\theta) < T\}}\right]. \end{aligned} \tag{2.2.39}$$

If we substitute this into (2.2.37) we can write

$$\begin{aligned} \mathbb{E}\left[N^{(n)}(\theta) \mathbb{1}_{\{N^{(n)}(\theta) < T\}}\right] = \frac{1}{g_n(r_0) v_n(r_0)} \Bigg[1 - \mathbb{E}\left[F_n(r_0, \theta, N^{(n)}(r_0, \theta)) \mathbb{1}_{\{N^{(n)}(\theta) \geq T\}}\right] \\ - \frac{1}{2} \mathbb{E}\left[\ddot{F}_n(\tau(\theta)) (N^{(n)}(\theta))^2 \mathbb{1}_{\{N^{(n)}(\theta) < T\}}\right] \Bigg], \end{aligned} \tag{2.2.40}$$

and so

$$n^{3/4}\mathbb{E}\left[N^{(n)}(\theta)\right] = \frac{n^{1/2}}{g(r_0)v(r_0)}\Bigg[1 - \mathbb{E}[F_n(r_0,\theta,N^{(n)}(r_0,\theta))\mathbb{1}_{\{N^{(n)}(\theta)\geq T\}}] - \frac{1}{2}\mathbb{E}\left[\ddot{F}_n(\tau(\theta))(N^{(n)}(\theta))^2\mathbb{1}_{\{N^{(n)}(\theta)<T\}}\right]\Bigg] + n^{3/4}\mathbb{E}\left[N^{(n)}(\theta)\mathbb{1}_{\{N^{(n)}(\theta)\geq T\}}\right]. \tag{2.2.41}$$

Next we compute the limit as $n \to \infty$ of each term in the expansion (2.2.41). By applying Corollary 2.2.6 with $k = 3/4$ and $l = 1$, it follows that

$$\lim_{n\to\infty} \sup_{(r_0,\theta)\in S\times[-\pi,\pi]} n^{3/4}\mathbb{E}\left[N^{(n)}(r_0,\theta)\mathbb{1}_{\{N^{(n)}(\theta)\geq T\}}\right] = 0. \tag{2.2.42}$$

Similarly, by Cauchy-Schwarz, (2.2.36) and by an application of Corollary 2.2.5 with $k = 1$ we have

$$\begin{aligned} n^{1/2}\mathbb{E}\left[F_n(r_0,\theta,N^{(n)}(\theta))\mathbb{1}_{\{N^{(n)}(\theta)\geq T\}}\right] &\leq n^{1/2}\sqrt{\mathbb{E}[F_n^2(r_0,\theta,N^{(n)}(\theta)]\mathbb{P}\{N^{(n)}(\theta)\geq T\}} \\ &= \sqrt{2}\sqrt{n\mathbb{P}\{N^{(n)}(\theta)\geq T\}} \\ &\to 0 \end{aligned} \tag{2.2.43}$$

uniformly for $(r_0,\theta) \in S \times [-\pi,\pi]$ as $n \to \infty$. Differentiating equation (2.2.34) twice yields

$$\begin{aligned} \ddot{F}_n(t) &= v_n(r_n(t))\dot{r}_n(t)g_n'(r_n(t)) + g_n(r_n(t))v_n'(r_n(t))\dot{r}_n(t) \\ &= n^{1/4}\dot{r}_n(t)\left(v(r_n(t))g'(r_n(t)) + g(r_n(t))v'(r_n(t))\right). \end{aligned}$$

To evaluate $\lim_{n\to\infty} n^{1/4}\dot{r}_n(t)$, note that by (2.2.4) and Lemma 2.2.3

$$\begin{aligned} |n^{1/4}\dot{r}_n(t) - v(r_0)\cos\theta| &= n^{1/4}|\dot{r}_n(t) - \dot{r}_n(0)| \\ &= n^{1/4}\left|\int_0^t \ddot{r}_n(s)ds\right| \\ &= \frac{1}{n^{1/4}}\left|\int_0^t \sqrt{n}\ddot{r}_n(s)ds\right| \\ &\to 0 \end{aligned} \tag{2.2.44}$$

uniformly for $(r_0,\theta,t) \in S \times [-\pi,\pi] \times [0,T]$ as $n \to \infty$. Differentiating (2.1.2) shows

$$v'(r_0) = -\frac{\partial_r\mathcal{U}(r_0)}{mv(r_0)},$$

which together with (2.2.44), Lemma 2.2.3, and the continuity of v' and g' on S forces

$$\lim_{n\to\infty} \ddot{F}_n(t) = \ddot{F}_1(0) = \dot{r}(0)\left(v(r_0)g'(r_0) + g(r_0)v'(r_0)\right) = \left(v^2(r_0)g'(r_0) - g(r_0)\frac{\partial_r \mathcal{U}(r_0)}{m}\right)\cos\theta$$

uniformly for $(r_0, \theta, t) \in S \times [-\pi, \pi] \times [0, T]$. Thus,

$$\lim_{n\to\infty} \sup_{(r_0,\theta,t)\in S\times[-\pi,\pi]\times[0,T]} \left| \ddot{F}_n(t) - \left(v^2(r_0)g'(r_0) - g(r_0)\frac{\partial_r \mathcal{U}(r_0)}{m}\right)\cos\theta \right| = 0. \tag{2.2.45}$$

In conjunction with Lemma 2.2.7 this yields

$$\begin{aligned}
&\sup_{(r_0,\theta)\in S\times[-\pi,\pi]} \sqrt{n}\cdot\mathbb{E}\left[\left|\ddot{F}_n(r_0,\theta,N^{(n)}(\theta)) - \ddot{F}_1(r_0,\theta,0)\right| \left(N^{(n)}(\theta)\right)^2 \mathbb{1}_{\{N^{(n)}(\theta)<T\}}\right] \le \\
&\sup_{(r_0,\theta,t)\in S\times[-\pi,\pi]\times[0,T]} \left\{\left|\ddot{F}_n(r_0,\theta,t) - \ddot{F}_1(r_0,0)\right|\right\} \sqrt{n} \sup_{(r_0,\theta)\in S\times[-\pi,\pi]} \mathbb{E}\left[\left(N^{(n)}(\theta)\right)^2 \mathbb{1}_{\{N^{(n)}(\theta)<T\}}\right] \\
&\to 0
\end{aligned}$$

as $n \to \infty$. This combined with (2.2.32) shows

$$\lim_{n\to\infty} \sup_{(r_0,\theta)\in S\times[-\pi,\pi]} \mathbb{E}\left[\sqrt{n}\ddot{F}_n(r_0,\theta,N^{(n)}(\theta))\left(N^{(n)}(\theta)\right)^2 \mathbb{1}_{\{N^{(n)}(\theta)<T\}} - \frac{2}{g^2(r_0)v^2(r_0)}\ddot{F}_1(r_0,\theta,0)\right] = 0. \tag{2.2.46}$$

By the expansion (2.2.41) of $n^{3/4}N(\theta)$, we have

$$\begin{aligned}
&n^{3/4}\mathbb{E}\left[N^{(n)}(\Theta)\cos(\Theta)\right] \\
&= \frac{n^{3/4}}{2\pi}\int_{-\pi}^{\pi} \mathbb{E}\left[N^{(n)}(\theta)\right]\cos(\theta)d\theta \\
&= \frac{n^{1/2}}{2\pi}\frac{1}{g(r_0)v(r_0)}\int_{-\pi}^{\pi}\cos(\theta)d\theta \\
&\quad - \frac{1}{2\pi}\frac{1}{g(r_0)v(r_0)}\int_{-\pi}^{\pi} n^{1/2}\mathbb{E}[F_n(r_0,\theta,N^{(n)}(r_0,\theta))\mathbb{1}_{\{N^{(n)}(\theta)\ge T\}}]\cos(\theta)d\theta \\
&\quad - \frac{1}{4\pi}\frac{1}{g(r_0)v(r_0)}\int_{-\pi}^{\pi} n^{1/2}\mathbb{E}\left[\ddot{F}_n(\tau(\theta))(N^{(n)}(\theta))^2\mathbb{1}_{\{N^{(n)}(\theta)<T\}}\right]\cos(\theta)d\theta \\
&\quad + \frac{1}{2\pi}\int_{-\pi}^{\pi} n^{3/4}\mathbb{E}\left[N^{(n)}(\theta)\mathbb{1}_{\{N^{(n)}(\theta)\ge T\}}\right]\cos(\theta)d\theta.
\end{aligned} \tag{2.2.47}$$

Finally, using (2.2.42), (2.2.43) and (2.2.45)

$$\begin{aligned}
\lim_{n\to\infty} n^{3/4}\mathbb{E}\left[N^{(n)}(\Theta)\cos(\Theta)\right] &= -\frac{1}{4\pi}\frac{2}{g^3(r_0)v^3(r_0)}\left(v^2(r_0)g'(r_0) - g(r_0)\frac{\partial_r\mathcal{U}(r_0)}{m}\right)\int_{-\pi}^{\pi}\cos^2\theta d\theta \\
&= \frac{1}{2g^2(r_0)v(r_0)}\left(-\frac{g'(r_0)}{g(r_0)} + \frac{\partial_r\mathcal{U}(r_0)}{mv^2(r_0)}\right).
\end{aligned} \tag{2.2.48}$$

This completes the proof. □

With the results above, we are now in a position to prove some results about the limiting behavior of the skeleton process $((R^n_k, A^n_k), k \in \mathbb{N}_0)$ defined in (2.1.5). In the next five lemmas we identify the limiting drift, variance and covariance terms of the limiting diffusion for $((R^n_k, A^n_k), k \in \mathbb{N}_0)$. These results will be instrumental in the next section when we prove convergence of the skeleton process.

Lemma 2.2.11. *Let*

$$\mu_{r,n}(r_0) := n\mathbb{E}\left[R^n_1 - R^n_0 | R^n_0 = r_0\right]$$

be the scaled drift of the Markov process $(R^n_k, k \in \mathbb{N}_0)$*. Then*

$$\lim_{n\to\infty} \sup_{r_0 \in S} \left| \mu_{r,n}(r_0) - \frac{1}{g^2(r_0)} \left(-\frac{g'(r_0)}{2g(r_0)} + \frac{1}{2r_0} - \frac{\partial_r \mathcal{U}(r_0)}{4(E - \mathcal{U}(r_0))} \right) \right| = 0.$$

Proof. By definition, of R^n_k,

$$\mu_{r,n}(r_0) = n\mathbb{E}\left[r_n\left(N^{(n)}(\Theta)\right) - r_0\right].$$

Let $\Lambda_\delta(S) := \Lambda^1_\delta(S)$ as defined in (2.2.10). Since $\Lambda^n_\delta(S)$ is increasing in n, we have by construction that $r_n(t) \in S'$ for all $n \in \mathbb{N}$ and $t \in [0, \Lambda_\delta(S)]$. So in particular, $r_n(t)$ is bounded away from $\partial\mathcal{D}$. To compute $\mu_{r,n}$ we first split the expectation on the events $\{N^{(n)}(\Theta) \le \Lambda_\delta(S)\}$ and $\{N^{(n)}(\Theta) > \Lambda_\delta(S)\}$. This allows us to write

$$\begin{aligned} \mu_{r,n}(r_0) = n\mathbb{E}&\left[\left(r_n\left(N^{(n)}(\Theta)\right) - r_0\right) \mathbb{1}_{\{N^{(n)} \le \Lambda_\delta(S)\}}\right] \\ &+ n\mathbb{E}\left[\left(r_n\left(N^{(n)}(\Theta)\right) - r_0\right) \mathbb{1}_{\{N^{(n)} > \Lambda_\delta(S)\}}\right]. \end{aligned} \tag{2.2.49}$$

To compute the first term of (2.2.49), we utilize a second order Taylor expansion of $r_n(t)$ evaluated at $t = N^{(n)}(\Theta)$ which yields

$$r_n\left(r_0, N^{(n)}(\Theta)\right) - r_0 = v_n(r_0)\cos(\Theta) \cdot \left[N^{(n)}(\Theta)\right] + \frac{1}{2}\ddot{r}_n(\tau)\left[N^{(n)}(\Theta)\right]^2$$

for some $0 \le \tau := \tau(r_0, N^n(\Theta)) \le N^n(\Theta)$. Hence

$$\begin{aligned} n\mathbb{E}\left[\left(r_n\left(N^{(n)}(\Theta)\right) - r_0\right) \mathbb{1}_{\{N^{(n)} \le \Lambda_\delta(S)\}}\right] &= \mathbb{E}\left[n v_n(r_0)\cos(\Theta) \cdot N^{(n)}(\Theta)\mathbb{1}_{\{N^{(n)} \le \Lambda_\delta(S)\}}\right] \\ &+ \frac{n}{2}\mathbb{E}\left[\ddot{r}_n(\tau) \cdot N^{(n)}(\Theta)^2 \mathbb{1}_{\{N^{(n)} \le \Lambda_\delta(S)\}}\right]. \end{aligned} \tag{2.2.50}$$

Note that

$$\left|\mathbb{E}\left[n v_n(r_0)\cos(\Theta) \cdot N^{(n)}(\Theta)\mathbb{1}_{\{N^{(n)} > \Lambda_\delta(S)\}}\right]\right| \le \sup_S |v(r_0)| n^{3/4}\mathbb{E}\left[N^{(n)}(\Theta)\mathbb{1}_{\{N^{(n)} > \Lambda_\delta(S)\}}\right] \to 0$$

uniformly for $r_0 \in S$ as $n \to \infty$ by Corollary 2.2.6. The limit of the first term of (2.2.50) can be computed by a direct application of Lemma 2.2.10.

$$\begin{aligned} &\lim_{n\to\infty} \mathbb{E}\left[nv_n(r_0)\cos(\Theta)\cdot N^{(n)}(\Theta)\mathbb{1}_{\{N^{(n)}\le\Lambda_\delta(S)\}}\right] \\ &= \lim_{n\to\infty} v(r_0)\mathbb{E}\left[n^{3/4}\cos(\Theta)\cdot N^{(n)}(\Theta)\right] = \frac{1}{g^2(r_0)}\cdot\left(\frac{\partial_r\mathcal{U}(r_0)}{4(E-\mathcal{U}(r_0))} - \frac{g'(r_0)}{2g(r_0)}\right) \end{aligned} \tag{2.2.51}$$

uniformly for $r_0 \in S$. We now compute the limit of the second term on the right hand side of (2.2.50).

Since $\tau(N^{(n)}(\theta)) \le N^{(n)}(\theta)$, by Lemma 2.2.3 and Lemma 2.2.4

$$\sup_{(r_0,\theta)\in S\times[-\pi,\pi]} \left|\mathbb{E}\left[\left(\sqrt{n}\ddot{r}_n(\tau(N^{(n)}(\theta)) - \psi(r_0,r_0,\theta)\right)\sqrt{n}(N^{(n)}(\theta))^2\mathbb{1}_{\{N^{(n)}\le\Lambda_\delta(S)\}}\right]\right| \le$$

$$\sup_{(r_0,\theta)\in S\times[-\pi,\pi]}\left\{\sup_{t\in t\times[0,\Lambda_\delta(S)]}\left|\sqrt{n}\ddot{r}_n(r_0,\theta,t) - \psi(r_0,r_0,\theta)\right|\cdot\mathbb{E}\left[\sqrt{n}(N^{(n)})^2(\theta)\mathbb{1}_{\{N^{(n)}\le\Lambda_\delta(S)\}}\right]\right\} \to 0$$

as $n \to \infty$. This together with Corollary 2.2.6 and Lemma 2.2.7 yields

$$\lim_{n\to\infty}\sup_{(r_0,\theta)\in S\times[-\pi,\pi]}\left|\mathbb{E}\left[n\ddot{r}_n(\tau(N^{(n)}(\theta))(N^{(n)})^2(\theta)\mathbb{1}_{\{N^{(n)}\le\Lambda_\delta(S)\}}\right] - \frac{2\psi(r_0,r_0,\theta)}{g^2(r_0)v^2(r_0)}\right| = 0.$$

Since this convergence is uniform in θ, we can evaluate the limit of the second order term of equation (2.2.50) by

$$\begin{aligned} &\lim_{n\to\infty}\mathbb{E}\left[\frac{n}{2}\ddot{r}_n(\tau(N^{(n)}(\Theta))\cdot(N^{(n)})^2(\Theta)\mathbb{1}_{\{N^{(n)}(\Theta)\le\Lambda_\delta(S)\}}\right] \\ &= \frac{1}{4\pi}\int_{-\pi}^{\pi}\lim_{n\to\infty}\mathbb{E}\left[\sqrt{n}\ddot{r}_n(\tau(N^{(n)}(\theta))\cdot\sqrt{n}(N^{(n)})^2(\theta)\mathbb{1}_{\{N^{(n)}(\theta)\le\Lambda_\delta(S)\}}\right]d\theta \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\psi(r_0,r_0,\theta)}{g^2(r_0)v^2(r_0)}d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{g^2(r_0)v^2(r_0)}\left(-\frac{\partial_r\mathcal{U}(r_0)}{m} + \frac{v^2(r_0)\sin^2(\theta)}{r_0}\right)d\theta. \\ &= \frac{1}{g^2(r_0)}\left(\frac{1}{2r_0} - \frac{\partial_r U(r_0)}{mv(r_0)^2}\right) = \frac{1}{g^2(r_0)}\left(\frac{1}{2r_0} - \frac{\partial_r U(r_0)}{2(E-U(r_0))}\right). \end{aligned} \tag{2.2.52}$$

uniformly for $r_0 \in S$.

By Lemma 2.2.8, the second term of (2.2.49) converges to 0 uniformly for $r_0 \in S$. So by adding the right hand sides of 2.2.51 and 2.2.51 we have

$$\lim_{n\to\infty}\mu_{r,n}(r_0) = \frac{1}{g^2(r_0)}\left(-\frac{g'(r_0)}{2g(r_0)} + \frac{1}{2r_0} - \frac{\partial_r\mathcal{U}(r_0)}{4(E-\mathcal{U}(r_0))}\right)$$

uniformly for $r_0 \in S$.

□

Lemma 2.2.12. *Let*

$$\sigma^2_{r,n}(r_0) := n\mathbb{E}\left[(R^n_1 - R^n_0)^2 \,|R^n_0 = r_0\right].$$

be the scaled variance of the Markov process $(R^n_k, k \in \mathbb{N}_0)$*. Then*

$$\lim_{n\to\infty} \sup_{r_0 \in S} \left|\sigma^2_{r,n}(r_0) - \frac{1}{g^2(r_0)}\right| = 0.$$

Proof. We proceed as in the proof of Lemma 2.2.11 by splitting $\sigma^2_{r,n}$ on the events $\{N^{(n)}(\Theta) \leq \Lambda_\delta(S)\}$ and $\{N^{(n)}(\Theta) > \Lambda_\delta(S)\}$.

$$\begin{aligned} \sigma^2_{r,n}(r_0) &= n\mathbb{E}\left[\left(r_n(N^{(n)}(\Theta)) - r_0\right)^2 \mathbb{1}_{\{N^{(n)} \leq \Lambda_\delta(S)\}}\right] \\ &\quad + n\mathbb{E}\left[\left(r_n(N^{(n)}(\Theta)) - r_0\right)^2 \mathbb{1}_{\{N^{(n)} > \Lambda_\delta(S)\}}\right] \end{aligned} \tag{2.2.53}$$

To evaluate the limit of the first term on the right hand side of (2.2.53), we utilize a first order Taylor expansion of $r_n(t)$ evaluated at $t = N^{(n)}(\Theta)$ which yields

$$\left(r_n(r_0, N^{(n)}(\Theta)) - r_0\right)^2 = \left(\dot{r}_n(\tau) N^{(n)}(\Theta)\right)^2$$

for some $0 \leq \tau := \tau(r_0, N^n(\Theta)) \leq N^n(\Theta)$. Hence

$$n\mathbb{E}\left[\left(r_n(N^{(n)}(\Theta)) - r_0\right)^2 \mathbb{1}_{\{N^{(n)} \leq \Lambda_\delta(S)\}}\right] = n\mathbb{E}\left[\dot{r}^2_n(\tau)\left(N^{(n)}(\Theta)\right)^2 \mathbb{1}_{\{N^{(n)} \leq \Lambda_\delta(S)\}}\right]. \tag{2.2.54}$$

Using (2.2.3) we can solve for $\dot{r}^2(t)$:

$$\dot{r}^2(t) = \bar{v}^2(t) - r^2(t)\dot{\alpha}^2(t) = \bar{v}^2(t) - \frac{\bar{v}^2(0) r_0^2 \sin^2\theta}{r(t)^2}. \tag{2.2.55}$$

Define

$$\gamma(\rho, r_0, \theta) = v^2(\rho) - \frac{v^2(r_0) r_0^2 \sin^2\theta}{\rho^2} \tag{2.2.56}$$

and note that by (2.2.55) we have

$$\dot{r}^2_n(t, r_0, \theta) = \frac{1}{\sqrt{n}} \gamma(r_n(t), r_0, \theta).$$

Since S' is bounded away from 0, it follows that γ is uniformly continuous on $S' \times S \times [-\pi, \pi]$. So by Lemma 2.2.3

$$\lim_{n\to\infty} \sup_{(r_0,\theta)\in S\times[-\pi,\pi]} \left\{ \sup_{t\in[0,\Lambda_\delta(S)]} \left|\sqrt{n}\dot{r}^2_n(t, r_0, \theta) - \gamma(r_0, r_0, \theta)\right| \right\} = 0 \tag{2.2.57}$$

where

$$\gamma(r_0, r_0, \theta) = v^2(r_0) - v^2(r_0)\sin^2(\theta) = v^2(r_0)\cos^2(\theta)$$

by (2.2.56).

Since $\tau(N^{(n)}(\theta)) \le N^{(n)}(\theta)$, we can apply (2.2.57) and Lemma 2.2.4 to show

$$\sup_{(r_0,\theta)\in S\times[-\pi,\pi]} \left|\mathbb{E}\left[\left(\sqrt{n}\dot{r}_n^2(\tau) - \gamma(r_0,r_0,\theta)\right)\sqrt{n}\left(N^{(n)}(\theta)\right)^2 \mathbb{1}_{\{N^{(n)}\le\Lambda_\delta(S)\}}\right]\right| \le$$

$$\sup_{(r_0,\theta)\in S\times[-\pi,\pi]} \left\{\sup_{t\in[0,\Lambda_\delta(S)]} \left|\sqrt{n}\dot{r}_n^2(t,r_0,\theta) - \gamma(r_0,r_0,\theta)\right| \cdot \mathbb{E}\left[\left(\sqrt{n}N^{(n)}(\theta)\right)^2 \mathbb{1}_{\{N^{(n)}\le\Lambda_\delta(S)\}}\right]\right\} \to 0$$

as $n \to \infty$. Along with Corollary 2.2.6 and Lemma 2.2.7 this implies

$$\lim_{n\to\infty} n\mathbb{E}\left[\left(r_n(N^{(n)}(\theta)) - r_0\right)^2 \mathbb{1}_{\{N^{(n)}\le\Lambda_\delta(S)\}}\right] = \frac{2\gamma(r_0,r_0,\theta)}{g^2(r_0)v^2(r_0)} = \frac{2\cos^2\theta}{g^2(r_0)}$$

uniformly for $(r_0,\theta) \in S \times [-\pi,\pi]$. So by the uniform convergence in θ, we can evaluate the limit of the first term of (2.2.53)

$$\begin{aligned}
&\lim_{n\to\infty} n\mathbb{E}\left[\left(r_n(N^{(n)}(\Theta)) - r_0\right)^2 \mathbb{1}_{\{N^{(n)}\le\Lambda_\delta(S)\}}\right] \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \lim_{n\to\infty} n\mathbb{E}\left[\left(r_n(N^{(n)}(\theta)) - r_0\right)^2 \mathbb{1}_{\{N^{(n)}\le\Lambda_\delta(S)\}}\right] d\theta \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{2\cos^2\theta}{g^2(r_0)} d\theta = \frac{1}{g^2(r_0)}
\end{aligned}$$

uniformly for $r_0 \in S$. By Lemma 2.2.8, the second term of (2.2.53) converges to 0 uniformly for $r_0 \in S$. So

$$\lim_{n\to\infty} \sigma_{r,n}(r_0) = \frac{1}{g^2(r_0)}$$

uniformly for $r_0 \in S$. ☐

Lemma 2.2.13. *Let*

$$\mu_{\alpha,n}(r_0,\alpha_0) = n\mathbb{E}\left[A_1^n - A_0^n | A_0^n = \alpha_0\right]$$

be the scaled drift of A_k^n in the Markov process $((R_k^n, A_k^n), k \in \mathbb{N}_0)$. Then

$$\lim_{n\to\infty} \sup_{(r_0,\alpha_0)\in S\times\mathbb{R}} |\mu_{\alpha,n}(r_0,\alpha_0)| = 0.$$

Proof. By (2.2.7), the second order Taylor expansion for $\alpha(t)$ is given by

$$\alpha(t) - \alpha_0 = \dot{\alpha}(0)t + \frac{\ddot{\alpha}(\tau)}{2}t^2 = \frac{v(r_0)\sin(\theta)}{r_0}t - \frac{v(r_0)r_0\sin(\theta)\dot{r}(\tau)}{r^3(\tau)}t^2 \qquad (2.2.58)$$

for some $0 \leq \tau := \tau(t) \leq t$. Then by splitting the expectation on the events $\{N^{(n)}(\Theta) \leq \Lambda_\delta(S)\}$ and $\{N^{(n)}(\Theta) > \Lambda_\delta(S)\}$ and substituting (2.2.58) in the $\{N^{(n)}(\Theta) \leq \Lambda_\delta(S)\}$ term we have

$$\begin{aligned}\mu_{\alpha,n} =& \mathbb{E}\left[\frac{v(r_0)}{r_0} n^{3/4} N^{(n)}(\Theta)\sin(\Theta)\mathbb{1}_{\{N^{(n)}\leq\Lambda_\delta(S)\}}\right] \\ &- n^{3/4} v(r_0) r_0 \mathbb{E}\left[\frac{\dot{r}_n\left(\tau\left(N^{(n)}(\Theta)\right)\right)}{r_n^3\left(\tau\left(N^{(n)}(\Theta)\right)\right)}\left(N^{(n)}(\Theta)\right)^2\sin(\Theta)\mathbb{1}_{\{N^{(n)}\leq\Lambda_\delta(S)\}}\right] \\ &+ n\mathbb{E}\left[\left(\alpha_n\left(N^{(n)}(\Theta)\right)-\alpha_0\right)\mathbb{1}_{\{N^{(n)}>\Lambda_\delta(S)\}}\right].\end{aligned} \tag{2.2.59}$$

An application of Corollary 2.2.9 shows the last term on the right hand side of (2.2.59) converges to 0 uniformly for $(r_0, \alpha_0) \in S \times \mathbb{R}$. For the first term of (2.2.59), we apply (2.2.41) which yields

$$\begin{aligned}&\mathbb{E}\left[n^{3/4}N^{(n)}(\Theta)\sin(\Theta)\mathbb{1}_{\{N^{(n)}\leq\Lambda_\delta(S)\}}\right] \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\mathbb{E}\left[n^{3/4}N^{(n)}(\theta)\mathbb{1}_{\{N^{(n)}\leq\Lambda_\delta(S)\}}\right]\sin(\theta)d\theta \\ &= \frac{\sqrt{n}}{2\pi g(r_0)v(r_0)}\int_{-\pi}^{\pi}\sin(\theta)d\theta - \frac{\sqrt{n}}{2\pi g(r_0)v(r_0)}\int_{-\pi}^{\pi}\mathbb{E}\left[F_n(\theta, N^{(n)}(r_0,\theta))\mathbb{1}_{\{N^{(n)}>\Lambda_\delta(S)\}\}}\right]\sin(\theta)d\theta \\ &- \frac{\sqrt{n}}{4\pi g(r_0)v(r_0)}\int_{-\pi}^{\pi}\mathbb{E}\left[\ddot{F}_n(\tau_1)(N^{(n)}(\theta))^2\mathbb{1}_{\{N^{(n)}\leq\Lambda_\delta(S)\}\}}\right]\sin(\theta)d\theta\end{aligned} \tag{2.2.60}$$

for some $0 \leq \tau_1 := \tau_1\left(N^{(n)}(\theta)\right) \leq N^{(n)}(\theta)$. Now

$$\int_{-\pi}^{\pi}\sqrt{n}\sin(\theta)d\theta = 0$$

for all n, and

$$\lim_{n\to\infty}\sup_{(r_0,\theta)\in S\times[-\pi,\pi]}\left|\sqrt{n}\mathbb{E}[F_n(\theta, N^{(n)}(\rho,\theta))\mathbb{1}_{\{N^{(n)}>\Lambda_\delta(S)\}}]\right| = 0$$

by equation (2.2.43). To compute the limit of the last term, we first observe that

$$\lim_{n\to\infty}\sqrt{n}\mathbb{E}\left[\ddot{F}_n(\tau_1)(N^{(n)}(\theta))^2\mathbb{1}_{\{N^{(n)}\leq\Lambda_\delta(S)\}}\right] = \frac{2}{g^2(r_0)v^2(r_0)}\left(v(r_0)^2 g'(r_0) - g(r_0)\frac{\partial_r\mathcal{U}(r_0)}{m}\right)\cos(\theta)$$

uniformly in $(r_0, \theta) \in S \times [-\pi, \pi]$ by equation (2.2.45). By the uniform convergence in θ we have

$$\lim_{n\to\infty} \mathbb{E}\left[n^{3/4}N^{(n)}(\Theta)\sin(\Theta)\mathbb{1}_{\{N^{(n)}(\Theta)\le\Lambda_\delta(S)\}}\right] = \int_{-\pi}^{\pi}\lim_{n\to\infty}\mathbb{E}\left[\ddot{F}_n(\tau_1)(N^{(n)}(\theta))^2\mathbb{1}_{\{N^{(n)}\le\Lambda_\delta(S)\}}\right]\sin(\theta)d\theta$$
$$= -\frac{1}{2\pi g^3(r_0)v^3(r_0)}\left(v(r_0)^2g'(r_0) - g(r_0)\frac{\partial_r\mathcal{U}(r_0)}{m}\right)\int_{-\pi}^{\pi}\cos(\theta)\sin(\theta)d\theta = 0 \tag{2.2.61}$$

uniformly for $r_0 \in S$. To evaluate the limit of second term of (2.2.59), we first compute

$$\lim_{n\to\infty}\frac{n}{2}\mathbb{E}\left[\ddot{\alpha}_n(0)\left(N^{(n)}(\Theta)\right)^2\mathbb{1}_{\{N^{(n)}\le\Lambda_\delta(S)\}}\right] = -\lim_{n\to\infty}\frac{v_0^2}{r_0^2}\sqrt{n}\mathbb{E}\left[\left(N^{(n)}(\Theta)\right)^2\sin(\Theta)\,\mathbb{1}_{\{N^{(n)}\le\Lambda_\delta(S)\}}\right]$$
$$= -\frac{v_0^2}{r_0^2}\int_{-\pi}^{\pi}\lim_{n\to\infty}\sqrt{n}\mathbb{E}\left[\left(N^{(n)}(\theta)\right)^2\mathbb{1}_{\{N^{(n)}\le\Lambda_\delta(S)\}}\right]\sin(\theta)d\theta$$
$$= -\frac{2}{g^2(r_0)r_0^2}\int_{-\pi}^{\pi}\sin(\theta)d\theta = 0 \tag{2.2.62}$$

by (2.2.32). Using Lemma 2.2.3 and (2.2.44)

$$\begin{aligned}\lim_{n\to\infty} n^{1/4}\left(\frac{\dot{r}_n(t)}{r_n^3(t)} - \frac{\dot{r}_n(0)}{r_n^3(0)}\right) &= \lim_{n\to\infty}\frac{n^{1/4}\dot{r}_n(t)\left(r_n^3(0) - r_n^3(t)\right) + n^{1/4}\left(\dot{r}_n(t) - \dot{r}_n(0)\right)r_n^3(t)}{r_n^3(t)r_0^3}\\ &= \frac{v(r_0)\cos(\theta)\cdot 0 + 0\cdot r_0^3}{r_0^6}\\ &= 0\end{aligned} \tag{2.2.63}$$

uniformly for $(r_0,\theta) \in S\times[-\pi,\pi]$. By Lemma (2.2.7), Corollary (2.2.6) and (2.2.63) we have

$$\begin{aligned}&\lim_{n\to\infty}\sup_{(r_0,\alpha_0)\in S\times\mathbb{R}}\frac{n}{2}\mathbb{E}\left[\left\{\ddot{\alpha}_n(0) - \ddot{\alpha}_n\left(\tau\left(N^{(n)}(\theta)\right)\right)\right\}\left(N^{(n)}(\Theta)\right)^2\mathbb{1}_{\{N^{(n)}\le\Lambda_\delta(S)\}}\right]\\ &\le \lim_{n\to\infty}\sup_{(r_0,\alpha_0,t)\in S\times\mathbb{R}\times[0,\Lambda_\delta(S)]}\frac{\sqrt{n}}{2}\left|\ddot{\alpha}_n(0) - \ddot{\alpha}_n(t)\right|\sup_{(r_0,\alpha_0)\in S\times\mathbb{R}}\mathbb{E}\left[\sqrt{n}\left(N^{(n)}(\Theta)\right)^2\mathbb{1}_{\{N^{(n)}\le\Lambda_\delta(S)\}}\right]\\ &= -v(r_0)r_0\lim_{n\to\infty}\sup_{(r_0,\alpha_0,t)\in S\times\mathbb{R}\times[0,\Lambda_\delta(S)]}\left|n^{1/4}\frac{\dot{r}_n(t)}{r_n^3(t)} - \frac{\dot{r}_n(0)}{r_n^3(0)}\right|\\ &\quad\times\lim_{n\to\infty}\sup_{(r_0,\alpha_0)\in S\times\mathbb{R}}\mathbb{E}\left[\sqrt{n}\left(N^{(n)}(\Theta)\right)^2\mathbb{1}_{\{N^{(n)}\le\Lambda_\delta(S)\}}\right]\\ &= 0\times\frac{2}{g^2(r_0)v^2(r_0)}\\ &= 0.\end{aligned} \tag{2.2.64}$$

Combining (2.2.62) and (2.2.64) yields

$$\lim_{n\to\infty} \sup_{r_0 \in S} n^{3/4} v(r_0) r_0 \mathbb{E}\left[\frac{\dot{r}_n\left(\tau\left(N^{(n)}(\theta)\right)\right)}{r_n^3\left(\tau\left(N^{(n)}(\theta)\right)\right)}\left(N^{(n)}(\Theta)\right)^2 \sin(\Theta) \mathbb{1}_{\{N^{(n)} \leq \Lambda_\delta(S)\}}\right] = 0.$$

Since we have shown that all the three terms from (2.2.59) converge to zero uniformly for $(r_0, \alpha_0) \in S \times \mathbb{R}$ we are done. □

Lemma 2.2.14. *Let*

$$\sigma_{\alpha,n}^2(r_0, \alpha_0) = n\mathbb{E}\left[(A_1^n - A_0^n)^2 \,|A_0^n = \alpha_0\right]$$

be the scaled variance of A_k^n *in the Markov process* $((R_k^n, A_k^n), k \in \mathbb{N}_0)$*. Then*

$$\lim_{n\to\infty} \sup_{(r_0,\alpha_0) \in S\times\mathbb{R}} \left|\sigma_{\alpha,n}^2(r_0, \alpha_0) - \frac{1}{g^2(r_0) r_0^2}\right| = 0.$$

Proof. By (2.2.58)

$$\begin{aligned}\sigma_{\alpha,n}^2 = \mathbb{E}&\left[\frac{v^2(r_0) r_0^2 \sin^2(\Theta)}{r_n^4(\tau)} n^{1/2}\left(N^{(n)}(\Theta)\right)^2 \mathbb{1}_{\{N^{(n)} \leq \Lambda_\delta(S)\}}\right] \\ &+ n\mathbb{E}\left[\left(\alpha_n\left(N^{(n)}(\Theta)\right) - \alpha_0\right)^2 \mathbb{1}_{\{N^{(n)} > \Lambda_\delta(S)\}}\right]\end{aligned} \tag{2.2.65}$$

for some $\tau := \tau\left(N^{(n)}(\Theta)\right) \leq N^{(n)}(\Theta)$.

The second term of (2.2.65) converges uniformly to 0 by Corollary 2.2.9. To compute the limit of the first term, we have by Lemma 2.2.4 and Lemma 2.2.3 that

$$\begin{aligned}&\sup_{r_0\in S}\left|\mathbb{E}\left[\left(\frac{v^2(r_0) r_0^2 \sin^2(\Theta)}{r_n^4(\tau)} - \frac{v^2(r_0)\sin^2(\Theta)}{r_0^2}\right) n^{1/2}\left(N^{(n)}(\Theta)\right)^2 \mathbb{1}_{\{N^{(n)} \leq \Lambda_\delta(S)\}}\right]\right| \\ &\leq \sup_{r_0\in S}\{v^2(r_0)\} \sup_{(r_0,\theta,t)\in S\times[-\pi,\pi]\times[0,\Lambda_\delta(S)]}\left\{\left|\frac{r_0^2}{r_n^4(t)} - \frac{1}{r_0^2}\right|\right\} \cdot \sup_{r_0\in S}\mathbb{E}\left|n^{1/2}\left(N^{(n)}(\Theta)\right)^2 \mathbb{1}_{\{N^{(n)} \leq \Lambda_\delta(S)\}}\right| \\ &= 0.\end{aligned} \tag{2.2.66}$$

Corollary 2.2.6 and (2.2.32) yield

$$\lim_{n\to} \sup_{r_0\in S}\left|\frac{v^2(r_0)}{r_0^2}\mathbb{E}\left[n^{1/2}\left(N^{(n)}(\Theta)\right)^2 \sin^2(\Theta) \mathbb{1}_{\{N^{(n)} \leq \Lambda_\delta(S)\}}\right] - \frac{1}{g^2(r_0) r_0^2}\right| = 0.$$

This together with (2.2.66) implies

$$\lim_{n\to\infty} \sup_{(r_0,\alpha_0)\in S\times\mathbb{R}}\left|\sigma_{\alpha,n}^2(r_0, \alpha_0) - \frac{1}{g^2(r_0) r_0^2}\right| = 0. \tag{2.2.67}$$

□

Next we will show the scaled covariance of R^n and A^n goes to 0 uniformly as a function of the initial position (r_0, α_0) as $n \to \infty$.

Lemma 2.2.15.

$$\lim_{n\to\infty} \sup_{(r_0,\alpha_0)\in S\times\mathbb{R}} |n\mathbb{E}\left[(R_1^n - R_0^n)(A_1^n - A_0^n)|(R_0^n, A_0^n) = (r_0, \alpha_0)\right]| = 0$$

Proof. By (2.2.9) and (2.2.7) we have

$$(\alpha_n(t) - \alpha_0)(r_n(t) - r_0) = \dot{\alpha}_n(\tau_\alpha)\dot{r}_n(0)t^2 + \frac{\dot{\alpha}(\tau_\alpha)\ddot{r}_n(\tau_r)}{2}t^3 \tag{2.2.68}$$

for some $0 \le \tau_\alpha(t), \tau_r(t) \le t$. So by splitting the expectation on the events $\{N^{(n)}(\Theta) \le \Lambda_\delta(S)\}$ and $\{N^{(n)}(\Theta) > \Lambda_\delta(S)\}$ it follows that

$$\begin{aligned}
&n\mathbb{E}\left[\left(\alpha_n\left(N^{(n)}(\Theta)\right) - \alpha_0\right)\left(r_n\left(N^{(n)}(\Theta)\right) - r_0\right)\right] = \\
&n\mathbb{E}\left[\dot{\alpha}_n(\tau_\alpha)\dot{r}_n(0)\left(N^{(n)}(\Theta)\right)^2 \mathbb{1}_{\{N^{(n)}\le\Lambda_\delta(S)\}}\right] + n\mathbb{E}\left[\frac{\dot{\alpha}_n(\tau_\alpha)\ddot{r}_n(\tau_r)}{2}\left(N^{(n)}(\Theta)\right)^3 \mathbb{1}_{\{N^{(n)}\le\Lambda_\delta(S)\}}\right] \\
&\quad + n\mathbb{E}\left[\left(\alpha_n\left(N^{(n)}(\Theta)\right) - \alpha_0\right)\left(r_n\left(N^{(n)}(\Theta)\right) - r_0\right)\mathbb{1}_{\{N^{(n)}>\Lambda_\delta(S)\}}\right].
\end{aligned} \tag{2.2.69}$$

We will show each of these terms converges to 0 uniformly for $(r_0, \alpha_0) \in S \times \mathbb{R}$.

For the first term, we first note that by (2.2.4) and (2.2.7)

$$\dot{\alpha}_n(\tau_\alpha)\dot{r}_n(0) = \frac{1}{\sqrt{n}}\frac{v^2(r_0)r_0}{r_n(\tau_\alpha)}\sin(\theta)\cos(\theta)$$

Furthermore, by Lemma 2.2.3

$$\lim_{n\to\infty} \sup_{(r_0,\theta)\in S\times[-\pi,\pi]} \sup_{t\in[0,\Lambda_\delta(S)]} \left|\sqrt{n}\dot{\alpha}_n(t)\dot{r}_n(0) - v^2(r_0)\sin(\theta)\cos(\theta)\right| = 0.$$

Corollary 2.2.6 and (2.2.32) show

$$\begin{aligned}
&= \lim_{n\to\infty}\sup_{r_0\in S}\sqrt{n}\mathbb{E}\left[v^2(r_0)\left(N^{(n)}(\Theta)\right)^2\sin\Theta\cos\Theta\mathbb{1}_{\{N^{(n)}\le\Lambda_\delta(S)\}}\right] \\
&= \frac{v^2(r_0)}{4\pi}\int_{-\pi}^{\pi}\lim_{n\to\infty}\sup_{r_0\in S}\sqrt{n}\mathbb{E}\left[\left(N^{(n)}(\theta)\right)^2\mathbb{1}_{\{N^{(n)}\le\Lambda_\delta(S)\}}\right]\sin(2\theta)d\theta \\
&= \frac{1}{2\pi g^2(r_0)}\int_{-\pi}^{\pi}\sin(2\theta)d\theta = 0.
\end{aligned} \tag{2.2.70}$$

Since none of the bounds involve α_0 it follows that

$$\lim_{n\to\infty} \sup_{(r_0,\alpha_0)\in S\times\mathbb{R}} \left|n\mathbb{E}\left[\dot{\alpha}_n(\tau_\alpha)\dot{r}_n(0)\left(N^{(n)}(\Theta)\right)^2\mathbb{1}_{\{N^{(n)}\le\Lambda_\delta(S)\}}\right]\right| = 0.$$

For the second term of (2.2.69) note that by Lemma 2.2.3 and (2.2.13)

$$\lim_{n\to\infty} \sup_{(r_0,\theta,t,s)\in S\times[-\pi,\pi]\times[0,\Lambda_\delta(S)]\times[0,\Lambda_\delta(S)]} \left| n^{3/4}\dot{\alpha}_n(s)\ddot{r}_n(t) - \frac{v(r_0)}{r_0}\sin(\theta)\psi(r_0,r_0,\theta)\right| = 0.$$

By Lemma 2.2.4

$$\sup_{(r_0,\theta)\in S\times[-\pi,\pi]} \mathbb{E}\left[\left(N^{(n)}(r_0,\theta)\right)^3\right] = \mathcal{O}\left(\frac{1}{n^{3/4}}\right),$$

so

$$\sup_{(r_0,\theta)\in S\times[-\pi,\pi]} n\mathbb{E}\left[\frac{\dot{\alpha}_n(\tau_\alpha)\ddot{r}_n(\tau_r)}{2}\left(N^{(n)}(\Theta)\right)^3 \mathbb{1}_{\{N^{(n)}\le\Lambda_\delta(S)\}}\right] = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)$$

which shows the second term of (2.2.69) converges uniformly to 0. Finally, for the last term we have by Cauchy-Schwarz, Lemma 2.2.8 and Corollary 2.2.9

$$\begin{aligned} & n\left(\mathbb{E}\left[\left(\alpha_n\left(N^{(n)}(\Theta)\right)-\alpha_0\right)\left(r_n\left(N^{(n)}(\Theta)\right)-r_0\right)\mathbb{1}_{\{N^{(n)}>\Lambda_\delta(S)\}}\right]\right)^2 \\ & \le n\cdot\mathbb{E}\left[\left(\alpha_n\left(N^{(n)}(\Theta)\right)-\alpha_0\right)^2\mathbb{1}_{\{N^{(n)}>\Lambda_\delta(S)\}}\right]\cdot\mathbb{E}\left[\left(r_n\left(N^{(n)}(\Theta)\right)-r_0\right)^2\mathbb{1}_{\{N^{(n)}>\Lambda_\delta(S)\}}\right] \to 0 \end{aligned}$$

uniformly for $r_0 \in S$. Finally, note that none of the upper bounds depend on α_0, which shows that the convergence of the covariance to 0 is uniform in α_0 as well. This completes the proof.

□

2.3 Convergence of the Skeleton Process

In this section we will prove the convergence of the skeleton process $((R^n_k, A^n_k), k \in \mathbb{N}_0)$ defined in (2.1.5). By construction, the transition operator P^n for $((R^n_k, A^n_k), k \in \mathbb{N}_0)$ is given by

$$P^n f(r_0,\alpha_0) = \mathbb{E}\left[f\left\{r_n(r_0,\Theta,N^{(n)}(r_0,\Theta)), \alpha_n(r_0,\alpha_0,\Theta,N^{(n)}(r_0,\Theta))\right\}\right]. \tag{2.3.1}$$

The following is the main theorem of this section.

Theorem 2.3.1. *Let $(\mathcal{R},\mathcal{A})$ be a diffusion on $\mathcal{D}\times\mathbb{R}$ whose generator G acts on functions $f \in C^2(\mathcal{D}\times\mathbb{R})$ with compact support in $C^2(\mathcal{D}^\circ\times\mathbb{R})$ by*

$$Gf(\rho,a) = \frac{1}{2g^2(\rho)}f_{\rho\rho}(\rho,a) + \frac{1}{2g^2(\rho)\rho^2}f_{aa}(\rho,a) + \frac{1}{g^2(\rho)}\left(-\frac{g'((\rho)}{2g(\rho)} + \frac{1}{2(\rho)} - \frac{\partial_r\mathcal{U}(\rho)}{2mv^2(\rho)}\right)f_\rho(\rho,a) \tag{2.3.2}$$

Suppose $((R^n_k, A^n_k), k \in \mathbb{N}_0)$ is a Markov process whose transition operator P^n is given by (2.3.1). Consider any $l, u \in \mathcal{D}^\circ$ with $l < u$ and start the process $((R^n_k, A^n_k), k \in \mathbb{N}_0)$ at (r_0,α_0), where $l < R^n_0 = r_0 < u$ and $A^n_0 := \alpha_0 \in \mathbb{R}$. Define the stopping times

$$\tau^n_{l,u} := \inf\{k \in \mathbb{N}_0 : R^n_k \ge u \text{ or } R^n_k \le l\}$$

and

$$\tau_{l,u} := \inf\{t \geq 0 : \mathcal{R}_t^n \geq u \text{ or } \mathcal{R}_t^n \leq l\}.$$

Then, as $n \to \infty$*, the family of continuous time processes* $((R^n_{\lfloor \mathrm{nt} \rfloor \wedge \tau^n_{l,u}}, A^n_{\lfloor \mathrm{nt} \rfloor \wedge \tau^n_{l,u}}), t \geq 0)$ *converges in distribution on the Skorokhod space to the diffusion* $\big((\mathcal{R}_{t\wedge\tau_{l,u}}, \mathcal{A}_{t\wedge\tau_{l,u}}), t \geq 0\big)$.

Remark 2.3.2. If the left boundary point of $\mathcal{D}$ is inaccessible for the diffusion $\mathcal{R}$ then we do not need the stopping at l, and we have that, as $n \to \infty$, $((R^n_{\lfloor \mathrm{nt} \rfloor \wedge \tau^n_u}, A^n_{\lfloor \mathrm{nt} \rfloor \wedge \tau^n_u}), t \geq 0)$ converges in distribution on the Skorokhod space to the diffusion $((\mathcal{R}_{t\wedge\tau_u}, \mathcal{A}_{t\wedge\tau_u}), t \geq 0)$ where τ_u and τ_u^n denote the hitting times of u by $\mathcal{R}$ and R^n respectively.

Similarly, if the right boundary point of $\mathcal{D}$ is inaccessible one can remove the stopping at u. See Section 2.5 for how one can determine when a point is inaccessible.

We prove Theorem 2.3.1 at the end of this section. The proof utilizes Theorem IX.4.21 from [JS03] which gives sufficient conditions to prove a continuous time step process converges to a diffusion. We reproduce the result here for completeness.

Theorem 2.3.3. *Suppose that for each* $n \in \mathbb{N}$*,* X^n *is a pure step Markov process. That is, its generator has the form*

$$A^n f(x) = \int [f(x+y) - f(x)] K^n(x, dy)$$

where K^n *is a finite transition kernel on* $\mathbb{R}^d$*. Then define* b^n *and* c^n *by*

$$b^n(x) = \int y K^n(x, dy), \ \ c^{n,ij}(x) = \int y^i y^j K^n(x, dy). \tag{2.3.3}$$

Let b*,* c *be continuous functions on* $\mathbb{R}^d$ *and suppose* X *is a diffusion whose generator is given by*

$$Gf(x) = \sum_{i\leq d} b_i(x) D_i f(x) + \frac{1}{2} \sum_{1\leq i,j\leq d} c^{i,j}(x) D_{i,j} f(x) \tag{2.3.4}$$

and defines a martingale problem with a unique solution (see Assumption IX.4.3 from [JS03]*). Assume that*

(i) $b^n \to b$*,* $c^n \to c$ *locally uniformly;*

(ii) $\sup_{x:|x|\leq a} \int K^n(x, dy)|y|^2 \mathbb{1}_{\{|y|>\varepsilon\}} \to 0$ *as* $n \uparrow \infty$ *for all* $\varepsilon > 0$*;*

(iii) $\nu_n \to \nu$ *weakly, where* ν_n *and* ν *are the initial distributions of* X_0^n *and* X_0 *respectively.*

Then the laws $\mathcal{L}(X^n)$ *converge weakly to* $P = \int P_x \nu(dx)$*, the law of the diffusion process* X *started with the initial distribution* ν*.*

We define

$$\tilde{r}_n(r_0,\theta) := r_n(r_0, N^{(n)}(r_0,\theta)) - r_0$$

and

$$\tilde{\alpha}_n(r_0,\alpha_0,\theta) := \alpha_n(r_0,\alpha_0, N^{(n)}(r_0,\theta)) - \alpha_0$$

to represent the radial and angular displacements respectively.

We suppose $\left(\left(R_k^{l,u,n}, A_k^{l,u,n}\right), k \in \mathbb{N}_0\right)$ is a Markov process whose transition operator $P_{l,u}^n$ is given by

$$P_{l,u}^n f(r_0,\alpha_0) = \mathbb{E}\left[f\left(r_0 + \tilde{r}_n((r_0 \vee l)\wedge u, \Theta)\right), \alpha_0 + \tilde{\alpha}_n\left((r_0 \vee l)\wedge u, \alpha_0, \Theta\right)\right]. \tag{2.3.5}$$

We refer to $\left(\left(R_k^{l,u,n}, A_k^{l,u,n}\right), k \in \mathbb{N}_0\right)$ as the *cut-off process.* $\left(\left(R_k^{l,u,n}, A_k^{l,u,n}\right), k \in \mathbb{N}_0\right)$ evolves as the original Markov process when $R_k^{l,u,n} \in (l,u)$, but is cut-off near $\partial\mathcal{D}$.

More precisely, the transition operator $P_{l,u}^n$ has the following properties:

- When $r_0 \in (l,u)$ the operator $P_{l,u}^n$ coincides with P^n.
- When $r_0 < l$ the operator $P_{l,u}^n$ describes a process that starts at r_0 but evolves like the process defined by P^n started at l.
- When $r_0 > u$ the operator $P_{l,u}^n$ describes a process that starts at r_0 but evolves like the process defined by P^n started at u.

The *cut-offs* act to prevent the coefficients of the generator from blowing up when $R^{l,u,n}$ approaches $\partial\mathcal{D}$. Note also that $P^{l,u,n}f(r_0,\alpha_0)$ is defined for all $r_0 \in \mathbb{R}$. We will proceed with the proof of Theorem 2.3.1 by first proving the following lemma which the shows convergence of the cut-off process. Then since the transition operators P^n and $P_{l,u}^n$ agree on the interval $[l,u]$, by stopping at the first exit time of this interval, we can pass to a convergence result for the original skeleton process.

Lemma 2.3.4. *Fix $l,u \in \mathcal{D}$ with $l < u$ and (r_0,α_0) in $(l,u)\times\mathbb{R}$. Suppose $\left(\left(R_k^{l,u,n}, A_k^{l,u,n}\right), k \in \mathbb{N}_0\right)$ is a Markov process with transition operator $P_{l,u}^n$ started from (r_0,α_0), and let $(\Gamma_t, t \geq 0)$ be a Poisson process with rate 1 that is independent of $\left(\left(R_k^{l,u,n}, A_k^{l,u,n}\right), k \in \mathbb{N}_0\right)$. Then the continuous time step processes $\left(\left(R_{\Gamma_{nt}}^{l,u,n}, A_{\Gamma_{nt}}^{l,u,n}\right), t \geq 0\right)$ converge as $n \to \infty$ in distribution on the Skorokhod space $D(\mathbb{R}^+,\mathbb{R}) \times D(\mathbb{R}^+,\mathbb{R})$ to the diffusion $\left(\left(\mathcal{R}_t^{l,u}, \mathcal{A}_t^{l,u}\right), t \geq 0\right)$ with generator $\mathcal{G}_{l,u}$ and started at (r_0,α_0). The generator $\mathcal{G}_{l,u}$ acts on functions $f \in C^2(\mathbb{R}^2)$ as*

$$\mathcal{G}_{l,u}f(r_0,\alpha_0) = \frac{1}{2g^2((r_0\vee l)\wedge u)} f_{\rho\rho}(r_0,\alpha_0) + \frac{1}{2g^2((r_0\vee l)\wedge u)((r_0\vee l)\wedge u)^2} f_{aa}(r_0,\alpha_0)$$
$$+ \frac{1}{g^2((r_0\vee l)\wedge u)}\left(-\frac{g'((r_0\vee l)\wedge u)}{2g((r_0\vee l)\wedge u)} + \frac{1}{2((r_0\vee l)\wedge u)} - \frac{\partial_r\mathcal{U}((r_0\vee l)\wedge u))}{4((E-\mathcal{U}(r_0\vee l)\wedge u))}\right) f_\rho(r_0,\alpha_0). \tag{2.3.6}$$

Proof. Step I:

We work with $\left(R^{l,u,n}_{\Gamma_{nt}}, A^{l,u,n}_{\Gamma_{nt}}\right)$ because one needs a continuous time pure jump process to apply Theorem 2.3.3. To determine the generator of $\left(\left(R^{l,u,n}_{\Gamma_{nt}}, A^{l,u,n}_{\Gamma_{nt}}\right), t \geq 0\right)$, let $f \in C^2_c(\mathbb{R}^2)$ and compute

$$\begin{aligned}\frac{1}{t}\mathbb{E}[f(R^{l,u,n}_{\Gamma_{nt}}, A^{l,u,n}_{\Gamma_{nt}}) - f(r_0,\alpha_0)] &= \sum_{k=0}^{\infty}\mathbb{P}(\Gamma_{nt}=k)\frac{1}{t}\mathbb{E}[f(R^{l,u,n}_k, A^{l,u,n}_k) - f(r_0,\alpha_0)] \\ &= 0 + \frac{1}{t}P(\Gamma_{nt}=1)\mathbb{E}[f(R^{l,u,n}_1, A^{l,u,n}_1) - f(r_0,\alpha_0)] \\ &\quad + \sum_{k=2}^{\infty}\frac{1}{t}\mathbb{P}(\Gamma_{nt}=k)\mathbb{E}[f(R^{l,u,n}_k, A^{l,u,n}_k) - f(r_0,\alpha_0)]\end{aligned} \tag{2.3.7}$$

where the first equality follows by the independence of Γ_{nt} and $\left(R^{l,u,n}, A^{l,u,n}\right)$. Furthermore,

$$\lim_{t\to 0}\frac{1}{t}\mathbb{P}(\Gamma_{nt}=1) = \lim_{t\to 0}\frac{e^{-nt}nt}{t} = n$$

and

$$\lim_{t\to 0}\frac{1}{t}\mathbb{P}(\Gamma_{nt}>1) = \lim_{t\to 0}\frac{e^{-nt}}{t}(e^{nt}-1-nt) = \lim_{t\to 0}e^{-nt}\,\mathcal{O}(t) = 0.$$

As a result

$$\lim_{t\to 0}\left|\sum_{k=2}^{\infty}\mathbb{P}(\Gamma_{nt}=k)\frac{1}{t}\mathbb{E}[f(R^{l,u,n}_k, A^{l,u,n}_{\Gamma_{nt}}) - f(r_0,\alpha_0)]\right| \leq 2f_\infty \lim_{t\to 0}\frac{1}{t}\mathbb{P}(\Gamma_{nt}>1) = 0$$

which forces

$$\lim_{t\to 0}\frac{1}{t}\mathbb{E}_x[f(R^{l,u,n}_{\Gamma_{nt}}, A^{l,u,n}_{\Gamma_{nt}}) - f(r_0,\alpha_0)] = n\mathbb{E}_x[f(R^{l,u,n}_1) - f(r_0,\alpha_0)] = n(P^n_{l,u}-1)f(r_0,\alpha_0).$$

This shows that $\left(R^{l,u,n}_{\Gamma_{nt}}, A^{l,u,n}_{\Gamma_{nt}}\right)$ is a continuous time process whose generator is given by $\mathcal{G}^n_{l,u}f(r_0,\alpha_0) := n(P^n_{l,u}-I)f(r_0,\alpha_0)$. Define the kernel

$$\begin{aligned}&K^n_{l,u}((r_0,\alpha_0),\cdot) \\ &:= n\mathbb{P}\left(\left((r_n((r_0\vee l)\wedge u, N^{(n)}(\Theta)) - (r_0\vee l)\wedge u, \alpha_n((r_0\vee l)\wedge u, \alpha_0, N^{(n)}(\Theta)) - \alpha_0\right)\in\cdot\right) \\ &= n\mathbb{P}\left((\tilde{r}_n((r_0\vee l)\wedge u,\Theta), \tilde{\alpha}_n((r_0\vee l)\wedge u,\alpha_0,\Theta)\in\cdot\right).\end{aligned} \tag{2.3.8}$$

Using $\mathcal{G}^n_{l,u}f(r_0,\alpha_0) := n(P^n_{l,u}-1)f(r_0,\alpha_0)$, it follows that

$$\mathcal{G}^n_{l,u}f(r_0,\alpha_0) = \int_{\mathbb{R}^2}\left(f(r_0+\rho,\alpha_0+a) - f(r_0,\alpha_0)\right)K^n_{l,u}((r_0,\alpha_0),d\rho da).$$

This shows that $\left(R^{l,u,n}_{\Gamma_{nt}}, A^{l,u,n}_{\Gamma_{nt}}\right)$ is a pure jump process.

We proceed by checking all the assumptions of Theorem 2.3.3 hold.

Step II:

First, we will show $\mathcal{G}_{l,u}$ defines a martingale problem with a unique solution. By Section 8.3 from [Øks03], this holds for G as defined in Theorem 2.3.3 if the matrix $c(x) := (c^{i,j}(x))$ is everywhere positive definite, $c^{i,j}(x)$ is continuous, $b(x) := (b_i(x))$ is measurable and there exists a D such that

$$|b(x)| + |c(x)|^{1/2} \leq D(1 + |x|)$$

From (2.3.6) we observe the diffusion matrix of of $\mathcal{G}_{l,u}$ is diagonal with diagonal entries

$$\frac{1}{2g^2((\cdot \vee l) \wedge u)((\cdot \vee l) \wedge u)^2}, \quad \frac{1}{2g^2((\cdot \vee l) \wedge u)}.$$

Since the density g is continuous and bounded away from 0 by Assumptions (A3) and (A5) both terms above are clearly positive, continuous and bounded.

The drift vector for $\mathcal{G}_{l,u}$ has sole nonzero entry

$$\frac{1}{g^2((r_0 \vee l) \wedge u)} \left(-\frac{g'((r_0 \vee l) \wedge u)}{2g((r_0 \vee l) \wedge u)} + \frac{1}{2((r_0 \vee l) \wedge u)} - \frac{\partial_r \mathcal{U}((r_0 \vee l) \wedge u))}{4((E - \mathcal{U}(r_0 \vee l) \wedge u))} \right)$$

which is bounded since:

- $\partial_r \mathcal{U}$ is bounded on $[l, u]$ by Assumption (A4).
- g and $E - \mathcal{U}$ are both bounded away from 0 on $[l, u] \subset \mathcal{D}^\circ$ by Assumptions (A4) and (A5).

Step III:

Next we check that Assumption (i) of Theorem 2.3.3 holds. If we let $S = [l, u]$, then Lemmas 2.2.11, 2.2.12, 2.2.13, 2.2.14 and 2.2.15 show that as $n \to \infty$

$$\begin{aligned}
&\mu_{r,n}(r_0) := n\mathbb{E}\left[\tilde{r}_n(r_0, \Theta)\right] \to \frac{1}{g^2(r_0)} \left(-\frac{g'(r_0)}{2g(r_0)} + \frac{1}{2r_0} - \frac{\partial_r \mathcal{U}(r_0)}{4(E - \mathcal{U}(r_0))} \right) \\
&\sigma^2_{r,n}(r_0) := n\mathbb{E}\left[\tilde{r}^2_n(r_0, \Theta)\right] \to \frac{1}{g^2(r_0)} \\
&\mu_{\alpha,n}(r_0, \alpha_0) := n\mathbb{E}\left[\tilde{\alpha}_n(r_0, \alpha_0, \Theta)\right] \to 0 \\
&\sigma^2_{\alpha,n}(r_0, \alpha_0) := n\mathbb{E}\left[\tilde{\alpha}^2_n(r_0, \alpha_0, \Theta)\right] \to \frac{1}{g^2(r_0) r_0^2} \\
&n\mathbb{E}\left[\tilde{r}_n(r_0, \Theta) \cdot \tilde{\alpha}_n(r_0, \alpha_0, \Theta)\right] \to 0
\end{aligned} \tag{2.3.9}$$

uniformly for $r_0 \in S$ and $\alpha_0 \in \mathbb{R}$. This implies

$$\begin{aligned}
&n\mathbb{E}\left[\tilde{r}_n((r_0 \vee l) \wedge u, \Theta)\right] \to \frac{1}{g^2((r_0 \vee l) \wedge u)}\left(-\frac{g'((r_0 \vee l) \wedge u)}{2g((r_0 \vee l) \wedge u)} + \frac{1}{2(r_0 \vee l) \wedge u} - \frac{\partial_r \mathcal{U}(r_0)}{4(E - \mathcal{U}((r_0 \vee l) \wedge u))}\right) \\
&n\mathbb{E}\left[\tilde{r}_n^2((r_0 \vee l) \wedge u, \Theta)\right] \to \frac{1}{g^2((r_0 \vee l) \wedge u)} \\
&n\mathbb{E}\left[\tilde{\alpha}_n((r_0 \vee l) \wedge u, \alpha_0, \Theta)\right] \to 0 \\
&n\mathbb{E}\left[\tilde{\alpha}_n^2((r_0 \vee l) \wedge u, \alpha_0, \Theta)\right] \to \frac{1}{g^2((r_0 \vee l) \wedge u)((r_0 \vee l) \wedge u)^2} \\
&n\mathbb{E}\left[\tilde{r}_n((r_0 \vee l) \wedge u, \Theta) \cdot \tilde{\alpha}_n((r_0 \vee l) \wedge u, \alpha_0, \Theta)\right] \to 0
\end{aligned} \tag{2.3.10}$$

uniformly for all $(r_0, \alpha_0) \in \mathbb{R}^2$ as $n \to \infty$. This verifies Assumption (i).

Step IV:

We now show Assumption (ii) holds. Fix $\varepsilon > 0$. By Lemma 2.2.3 and (2.2.58) we have

$$\lim_{n\to\infty} \sup_{(r_0,\theta)\in[l,u]\times[-\pi,\pi]} \left\{ \sup_{t\in[0,T]} \|r_n(r_0,\theta,t) - \alpha_0\| \right\} = 0.$$

and

$$\lim_{n\to\infty} \sup_{(r_0,\alpha_0,\theta)\in[l,u]\times\mathbb{R}\times[-\pi,\pi]} \left\{ \sup_{t\in[0,T]} \|\alpha_n(r_0,\theta,t) - r_0\| \right\} = 0.$$

As a result there exists a $\delta > 0$ such that for n large enough, and for any $(r_0, \alpha_0) \in [l, u] \times \mathbb{R}$, $|r_n(r_0, t) - r_0| > \frac{\varepsilon}{\sqrt{2}}$ implies $t > \delta$, and $|\alpha_n(r_0, \alpha_0, t) - \alpha_0| > \frac{\varepsilon}{\sqrt{2}}$ implies $t > \delta$. Thus,

$$\begin{aligned}
&\{\|(\tilde{r}_n((\rho \vee l) \wedge u, \Theta), \tilde{\alpha}_n((\rho \vee l) \wedge u, \Theta)\| > \varepsilon\} \\
&\subseteq \{|\tilde{r}_n((\rho \vee l) \wedge u, \Theta)| > \frac{\varepsilon}{\sqrt{2}}\} \cup \{|\tilde{\alpha}_n((\rho \vee l) \wedge u, \alpha_0, \Theta)| > \frac{\varepsilon}{\sqrt{2}}\} \\
&\subseteq \{N^{(n)}((\rho \vee l) \wedge u, \Theta) > \delta\}.
\end{aligned}$$

So for large n,

$$\begin{aligned}
&\sup_{(r_0,\alpha_0)\in\mathbb{R}^2} \int (\rho^2 + a^2)\mathbb{1}_{\{\|\hat{\rho}\|>\varepsilon\}} K^n_{l,u}(\hat{r}_0, d\hat{\rho}) \\
&= n \sup_{(r_0,\alpha_0)\in\mathbb{R}^2} \mathbb{E}\left[\left(\tilde{r}_n^2((\rho \vee l) \wedge u, \Theta) + \tilde{\alpha}_n^2((\rho \vee l) \wedge u, \alpha_0, \Theta)\right)\mathbb{1}_{\{\|(\tilde{r}_n,\tilde{\alpha}_n)\|>\varepsilon\}}\right] \\
&\leq n \sup_{(r_0,\alpha_0)\in\mathbb{R}^2} \mathbb{E}\left[\tilde{r}_n^2((\rho \vee l) \wedge u, \Theta)\mathbb{1}_{\{N^{(n)}>\delta\}}\right] + n \sup_{(r_0,\alpha_0)\in\mathbb{R}^2} \mathbb{E}\left[\tilde{\alpha}_n^2((\rho \vee l) \wedge u, \alpha_0, \Theta)\mathbb{1}_{\{N^{(n)}>\delta\}}\right] \\
&\to 0
\end{aligned} \tag{2.3.11}$$

where the last limit follows from Lemma 2.2.8 and Corollary 2.2.9. This shows assumption (*ii*) holds. Assumption (iii) is trivially satisfied since the starting points are fixed to be (r_0, α_0) for all n. Therefore, all the assumptions of Theorem 2.3.3 hold and so $(R^{l,u,n}_{\Gamma_{nt}}, A^{l,u,n}_{\Gamma_{nt}})$ converges in distribution to the diffusion with generator $\mathcal{G}_{l,u}$. □

To apply the convergence of the cut-off process from Lemma 2.3.4 to the proof of Theorem 2.3.1, we first require a technical lemma about the continuity of first passage times. The following lemma shows that if a sequence of functions converges in the Skorokhod topology, then their first passage times converge as well as long as the level is not a local extrema of the limiting function.

Lemma 2.3.5. *For $\gamma \in D[0,\infty)$ define $\tau_x(\gamma) = \inf\{t : \gamma(t) \geq x\}$ to be the first passage time of γ across x. Consider an $f \in C[0,\infty)$ and an $a > 0$ such that*

$$0 < \tau_a(f) < \infty \quad \textit{and} \quad \inf\{t > \tau_a(f) : f(t) > a\} = \tau_a(f).$$

If $(f_n)_{n\geq 1}$ is a sequence in $D[0,\infty)$ such that $f_n \to f$ in $D[0,\infty)$, then $\tau_a(f_n) \to \tau_a(f)$.

Proof. Fix an arbitrary $\epsilon > 0$. Since f is continuous, f_n converges to f uniformly on compact sets. In particular, we may fix $T > \tau_a(f)$ and consider f_n converging to f uniformly on $[0,T]$. Since $\inf\{t > \tau_a(f) : f(t) > a\} = \tau_a(f)$, we can find s_0 such that $0 < s_0 - \tau_a(f) < \epsilon$ and $f(s_0) > a$. Moreover, because f is a continuous function, $\sup\{f(t) : t \leq \tau_a(f) - \epsilon/2\} < a$. Since $f_n \to f$ uniformly on $[0,T]$, there exists N such that $n \geq N$ implies that

$$\sup_{0\leq t\leq T} |f_n(t) - f(t)| < \min\left(\frac{f(s_0) - a}{2}, \frac{a - \sup\{f(t) : t \leq \tau_a(f) - \epsilon/2\}}{2}\right).$$

For $n \geq N$, we have

$$f_n(s_0) \geq f(s_0) - \frac{f(s_0) - a}{2} = \frac{f(s_0) + a}{2} > a$$

and therefore $\tau_a(f_n) \leq s_0 < \tau_a(f) + \epsilon$. Moreover, for $n \geq N$,

$$\begin{aligned}\sup\{f_n(t) : t \leq \tau_a(f) - \epsilon/2\} &\leq \sup\{f(t) : t \leq \tau_a(f) - \epsilon/2\} + \frac{a - \sup\{f(t) : t \leq \tau_a(f) - \epsilon/2\}}{2} \\ &= \frac{a + \sup\{f(t) : t \leq \tau_a(f) - \epsilon/2\}}{2} \\ &< a,\end{aligned}$$

and thus $\tau_a(f_n) \geq \tau_a(f) - \epsilon/2$. Hence for $n \geq N$ we have $|\tau_a(f_n) - \tau_a(f)| < \epsilon$. Since ϵ is arbitrary, the lemma is proved. □

We now prove Theorem 2.3.1.

Proof. (Theorem 2.3.1) By the Skorokhod representation theorem (see for example Theorem 6.7 from [Bil99]) and by Lemma 2.3.4 we can construct the relevant process on a single probability space so that for any $T > 0$

$$\Big((R^{l,u,n}_{\Gamma_{nt}}, A^{l,u,n}_{\Gamma_{nt}}), 0 \le t \le T\Big) \to \Big((\mathcal{R}^{l,u}_t, \mathcal{A}^{l,u}_t), 0 \le t \le T\Big)$$

a.s. in the Skorokhod topology on $D[0,T]$. Since $\lim_{n\to\infty} \frac{\Gamma_{nt}}{n} = t$ uniformly on every compact interval we have

$$\Big((R^{l,u,n}_{\lfloor \mathrm{nt} \rfloor}, A^{l,u,n}_{\lfloor \mathrm{nt} \rfloor}), 0 \le t \le T\Big) \to \Big((\mathcal{R}^{l,u}_t, \mathcal{A}^{l,u}_t), 0 \le t \le T\Big) \tag{2.3.12}$$

a.s. in the Skorokhod topology on $D[0,T]$. Since the radial diffusion term of $\mathcal{G}_{l,u}$, $\frac{1}{2g^2((r_0\vee l)\wedge u)}$ is bounded away from 0, $\mathcal{R}^{l,u}_t$ almost surely fluctuates across fixed levels after first hitting them. So, in the notation of Lemma 2.3.5, we have $\tau_u(\mathcal{R}^{l,u}_\cdot) = \inf\{t : R^{l,u}(t) \ge u\}$. By applying Lemma 2.3.5, it follows that $\tau_u(R^{l,u,n}_{\lfloor \mathrm{n}\cdot \rfloor}) \to \tau_u(\mathcal{R}^{l,u}_\cdot)$ as $n \to \infty$. Likewise, if for any $\gamma \in D[0,\infty)$ we define $\tau_{\cdot^-}(\gamma) := \inf\{t : \gamma(t) < l\}$, then $\tau_{l^-}(R^{l,u,n}_{\lfloor \mathrm{n}\cdot \rfloor}) \to \tau_{l^-}(\mathcal{R}^{l,u}_\cdot)$ as well. To simplify notation, let

$$\Delta_n := \tau_u(R^{l,u,n}_{\lfloor \mathrm{n}\cdot \rfloor}) \wedge \tau_{l^-}(R^{l,u,n}_{\lfloor \mathrm{n}\cdot \rfloor})$$

and

$$\Delta := \tau_u(\mathcal{R}^{l,u}_\cdot) \wedge \tau_{l^-}(\mathcal{R}^{l,u}_\cdot)$$

Then it follows that

$$\Big((R^{l,u,n}_{\lfloor \mathrm{nt} \rfloor \wedge \Delta_n}, A^{l,u,n}_{\lfloor \mathrm{nt} \rfloor \wedge \Delta_n}), 0 \le t \le T\Big) \to \Big((\mathcal{R}^{l,u}_{t\wedge\Delta}, \mathcal{A}^{l,u}_{t\wedge\Delta}), 0 \le t \le T\Big) \tag{2.3.13}$$

a.s. in the Skorokhod topology. By (2.3.5), the transition operators $P^n(r_0, \alpha_0)$ and $P^n_{l,u}(r_0, \alpha_0)$ agree when $r_0 \in (l,u)$. Furthermore, $\mathcal{G}_{l,u} f = \mathcal{G} f$ for any $f \in C^2_c([l,u], \mathbb{R})$. Since $\mathcal{G}_{l,u}$ defines a martingale problem with a unique solution (see the proof of Lemma 2.3.4), we have by (2.3.13) that

$$\Big((R^n_{\lfloor \mathrm{nt} \rfloor \wedge \tau^n_{l,u}}, A^n_{\lfloor \mathrm{nt} \rfloor \wedge \tau^n_{l,u}}), 0 \le t \le T\Big) \to \big((\mathcal{R}_{t\wedge\tau_{l,u}}, \mathcal{A}_{t\wedge\tau_{l,u}}), 0 \le t \le T\big)$$

a.s. in the Skorokhod topology, hence also in distribution. □

2.4 The Process on its Natural Time Scale

In this section, we study the convergence of the full trajectory of the particle. To this end, we need to keep track of the reflection times and the angle of reflection. That is, we will look at the Markov process $((R^n_k, A^n_k, \Delta^n_k, \Theta_k), k \in \mathbb{N}_0)$ with transition operator

$$\hat{P}^n f(r_0, \alpha_0, z_0, \theta_0) = \mathbb{E}\left[f\left(r_n(N^{(n)}(\Theta)), \alpha_n(N^{(n)}(\Theta)), n^{1/4} N^{(n)}(\Theta), \Theta\right)\right].$$

The real time of the k-th reflection is given by $T_k^n = n^{-1/4}\sum_{j=0}^k \Delta_j^n$, where we set $\Delta_0^n = 0$. Furthermore, we extend T^n to all $[0,\infty)$ by linear interpolation:

$$T_s^n := T_{\lfloor s\rfloor}^n + (s - \lfloor s\rfloor)(T_{\lfloor s\rfloor+1}^n - T_{\lfloor s\rfloor}^n),\ \forall s \in [0,\infty) \tag{2.4.1}$$

By (2.1.4) the full trajectory $(R^n(t), A^n(t)), t \geq 0)$ of the particle after rescaling is given by

$$\begin{aligned} R^n(t) &:= r_n\left(R_k^n, \Theta_k, t - T_k^n\right) \\ A^n(t) &:= \alpha_n\left(R_k^n, A_k^n, \Theta_k, t - T_k^n\right) \end{aligned} \tag{2.4.2}$$

for $t \in [T_k^n, T_{k+1}^n)$. Recall that $r_n(r_0, \theta_0, t)$ and $\alpha_n(r_0, \alpha_0, \theta_0, t)$ represent the radius and angle respectively of a particle at time t in the potential $\mathcal{U}_n$ when started at position (r_0, α_0) and having total energy E_n. θ_0 is the angle the initial velocity vector makes with the the initial position vector.

The following theorem is the main result of this section.

Theorem 2.4.1. *Let $(R^n(t), A^n(t))$ denote the full trajectory of the particle as defined in* (2.4.2) *and let*

$$\iota_{l,u}^n := \inf\{t : R^n(t) \leq l \text{ or } R^n(t) \geq u\}$$

denote the first time $R^n(t)$ leaves the interval (l,u). Then for any fixed $l, u \in \mathcal{D}^\circ$ with $l < u$, as $n \to \infty$ we have the following convergence in distribution on $D(\mathbb{R}_+, \mathbb{R})$:

$$\left(\left(R^n(n^{3/4}t \wedge \iota_{l,u}^n), A^n(n^{3/4}t \wedge \iota_{l,u}^n)\right), t \geq 0\right) \to \left((\mathcal{R}(\Omega(t) \wedge \tau_{l,u})\mathcal{A}(\Omega(t) \wedge \tau_{l,u})\right), t \geq 0)$$

where Ω is the time change given by

$$\Omega(t) := \mathcal{I}\left(\int_0^\cdot \frac{ds}{\lambda(R(s \wedge \tau_{l,u}))}\right)(t)$$

and $\mathcal{I}$ is the inverse operator defined by $\mathcal{I}(f)(t) = \inf\{s : f(s) > t\}$. Furthermore, the generator for the time changed process $((\mathcal{R}(\Omega(t)), \mathcal{A}(\Omega(t)), t \geq 0))$ is $\mathscr{G}$, the generator of the diffusion $((\mathscr{R}(t), \mathscr{A}(t)), t \geq 0)$ from Theorem 2.1.1.

Remark 2.4.2. The stopping at u and/or at l can be removed when the left and/or right boundary points of $\mathcal{D}$ are inaccessible for the diffusion $\mathcal{R}$. See Remark 2.3.2 and Section 2.5.

The idea of the proof is to time change the step-process $R_{\lfloor n\cdot\rfloor}^n$ by a function $\Omega_n(t)$ such that

- $\Omega_n \to \Omega$ as $n \to \infty$,
- $R^n(n^{3/4}\cdot)$ and $R_{n\Omega_n(\cdot)}^n$ agree at reflection times, as do $A^n(n^{3/4}\cdot)$ and $A_{n\Omega_n(\cdot)}^n$.

More specifically, Ω_n is chosen to satisfy $n\Omega_n\left(\frac{1}{n^{3/4}}T_k^n\right) = k$ so that

$$\left(R^n(n^{3/4}t), A^n(n^{3/4}t)\right) = \left(R^n_{n\Omega_n(t)}, A^n_{n\Omega_n(t)}\right) = (R^n_k, A^n_k)$$

whenever $t = \frac{1}{n^{3/4}}T_k^n$. We can then utilize our previous convergence results for the step-process $R^n_{\lfloor \mathrm{n}\cdot\rfloor}$ to show that $\left(R_n(n^{3/4}\cdot), A_n(n^{3/4}\cdot)\right)$ and $\left(R^n_{n\Omega_n(\cdot)}, A^n_{n\Omega_n(\cdot)}\right)$ remain close on each interval $\left[\frac{1}{n^{3/4}}T^n_{k-1}, \frac{1}{n^{3/4}}T^n_k\right)$.

We will also need to make use of the cut-off process $\left((R^{l,u}_k, A^{l,u}_k, \Delta^{l,u,n}_k), k \in \mathbb{N}_0\right)$ with transition operator

$$\begin{aligned}&\tilde{P}^{l,u,n} f(r_0, \alpha_0) = \\ &\mathbb{E}\left[f\left(r_n(N^{(n)}(((r_0 \vee l) \wedge u, \Theta)), \alpha_n(N^{(n)}((r_0 \vee l) \wedge u, \Theta)), n^{1/4}N^{(n)}((r_0 \vee l) \wedge u, \Theta)\right)\right]\end{aligned}$$

and the corresponding cut-off reflection-time process:

$$T^{l,u,n}_s := \frac{1}{n^{1/4}} \sum_{j=1}^{\lfloor \mathrm{s} \rfloor} \Delta^{l,u,n}_j + \frac{1}{n^{1/4}}(s - \lfloor \mathrm{s} \rfloor)(\Delta^{l,u,n}_{\lfloor \mathrm{s} \rfloor + 1} - \Delta^{l,u,n}_{\lfloor \mathrm{s} \rfloor}).$$

Note that $T^{l,u,n}$ is just the cut-off version of T^n from (2.4.1).

To motivate the definition of the time change Ω, we first need to evaluate the limiting behavior of $T^{l,u,n}$ as $n \to \infty$. The next lemma shows that the following convergence in distribution holds in $D(\mathbb{R}_+, \mathbb{R})$ for the cut-off time process $T^{l,u,n}$.

Lemma 2.4.3. *Let $\left(\left(\mathcal{R}^{l,u}_t, \mathcal{A}^{l,u}_t\right), t \geq 0\right)$ be a diffusion whose generator extends the operator $\mathcal{G}_{l,u}$ defined in* (2.3.6). *Then we have the following convergence in distribution on the product space $D(\mathbb{R}_+, \mathbb{R}) \times D(\mathbb{R}_+, \mathbb{R}) \times D(\mathbb{R}_+, \mathbb{R}_+)$.*

$$\left(\left(R^{l,u,n}_{\lfloor \mathrm{nt} \rfloor}, A^{l,u,n}_{\lfloor \mathrm{nt} \rfloor}, \frac{1}{n^{3/4}}T^{l,u,n}_{ns}\right), t, s \geq 0\right) \to \left(\left(\mathcal{R}^{l,u}_t, \mathcal{A}^{l,u}_t, \int_0^s \frac{1}{\lambda((\mathcal{R}^{l,u}_{\tilde{s}} \vee l) \wedge u)} d\tilde{s}\right), t, s \geq 0\right) \tag{2.4.3}$$

In particular, we have

$$\left(\frac{1}{n^{3/4}}T^{l,u,n}_{ns}, s \geq 0\right) \to \left(\int_0^s \frac{1}{\lambda((\mathcal{R}^{l,u}_{\tilde{s}} \vee l) \wedge u)} d\tilde{s}, s \geq 0\right) \tag{2.4.4}$$

as $n \to \infty$.

Proof. Let $\mathcal{F}_k = \sigma\left((R^{l,u,n}_j, A^{l,u,n}_j)), 0 \leq j \leq k\right)$, and consider the martingale with respect to the filtration $(\mathcal{F}_k)_{k \geq 0}$ given by

$$W^n_k := \sum_{j=1}^{k} \left(\Delta^{l,u,n}_j - \mathbb{E}[\Delta^{l,u,n}_j | \mathcal{F}_{j-1}]\right), \quad k \in \mathbb{N}_0. \tag{2.4.5}$$

Define

$$\phi_n(\rho) = n^{1/4}\mathbb{E}\left(N^{(n)}((\rho \vee l) \wedge u), \Theta)\right). \tag{2.4.6}$$

It follows from the Markov property that

$$\mathbb{E}[\Delta_j^{l,u,n}|\mathcal{F}_{j-1}] = \phi_n(R_{j-1}^{l,u,n}). \tag{2.4.7}$$

By Lemma 2.2.4 we have

$$\sup_{n,\rho} \phi_n(\rho) < \infty$$

which in turn implies

$$\xi := \sup_{(j,n)\in(\mathbb{N}_0,\mathbb{N})} \mathbb{E}\left[\left(\Delta_j^{l,u,n} - \mathbb{E}[\Delta_j^{l,u,n}|\mathcal{F}_{j-1}]\right)^2\right] < \infty. \tag{2.4.8}$$

Hence by Doob's maximal inequality, for every $\varepsilon > 0$, and for every $m \geq 1$

$$\mathbb{P}\left\{\sup_{1\leq k\leq mn} |W_k^n| \geq n\varepsilon\right\} \leq \frac{1}{n^2\varepsilon^2}\mathbb{E}\left[|W_{mn}^n|^2\right] \leq \frac{m\xi}{n\varepsilon^2} \tag{2.4.9}$$

from which it follows that

$$\sup_{1\leq k\leq mn}\left|\frac{1}{n}W_k^n\right| \to 0 \tag{2.4.10}$$

in probability as $n \to \infty$. By Lemma 2.2.7

$$\lim_{n\to\infty}\sup_{\rho\in\mathcal{D}}\left|\phi_n(\rho) - \frac{1}{\lambda((\rho \vee l) \wedge u)}\right| = 0.$$

It follows that

$$\lim_{n\to\infty}\sup_{1\leq k\leq mn}\frac{1}{n}\sum_{j=1}^{k}\left|\phi_n\left(R_{j-1}^{l,u,n}\right) - \frac{1}{\lambda((R_{j-1}^{l,u,n} \vee u) \wedge v)}\right| \to 0 \tag{2.4.11}$$

almost surely.

Hence, by (2.4.5), (2.4.6), (2.4.7), (2.4.10), (2.4.11) and the triangle inequality

$$\lim_{n\to\infty}\sup_{1\leq k\leq mn}\left|\frac{1}{n}\sum_{j=1}^{k}\Delta_j^{l,u,n} - \frac{1}{n}\sum_{j=1}^{k}\frac{1}{\lambda((R_{j-1}^{l,u,n} \vee u) \wedge v)}\right| = 0, \tag{2.4.12}$$

in probability. Let

$$\psi^{l,u}(f)(t) := \int_0^t \frac{1}{\lambda((f(s) \vee l) \wedge u)}\,ds. \tag{2.4.13}$$

Then since $R^{l,u,n}_{\lfloor n\cdot \rfloor}$ is a step function with step size $\frac{1}{n}$,

$$\psi^{l,u}\left(R^{l,u,n}_{\lfloor n\cdot \rfloor}\right)\left(\frac{k}{n}\right) = \frac{1}{n}\sum_{j=1}^{k}\frac{1}{\lambda(R^{l,u,n}_{j-1}\vee l)\wedge u)}$$

which implies together with (2.4.12)

$$\lim_{n\to\infty}\sup_{1\le k\le mn}\left|\frac{1}{n}\left(\sum_{j=1}^{k}\Delta^{l,u,n}_{j}\right) - \psi^{l,u}\left(R^{l,u,n}_{\lfloor n\cdot \rfloor}\right)\left(\frac{k}{n}\right)\right| = 0$$

in probability. Or equivalently,

$$\lim_{n\to\infty}\sup_{k\le m}\left|\frac{1}{n^{3/4}}T^{l,u,n}_{nk} - \psi^{l,u}\left(R^{l,u,n}_{\lfloor n\cdot \rfloor}\right)(k)\right| = 0 \qquad (2.4.14)$$

in probability. Now $f \to \psi^{l,u}(f)$ is a continuous mapping in the Skorokhod topology. By (2.3.12) we have the convergence in distribution of the cut-off processes:

$$\left(\left(R^{l,u,n}_{\lfloor nt \rfloor}, A^{l,u,n}_{\lfloor nt \rfloor}\right), t\ge 0\right) \to \left(\left(\mathcal{R}^{l,u,}_{t}, \mathcal{A}^{l,u}_{t}\right), t\ge 0\right)$$

as $n\to\infty$, which implies

$$\left(R^{l,u,n}_{\lfloor nt \rfloor}, A^{l,u,n}_{\lfloor nt \rfloor}, \psi^{l,u}(R^{l,u,n}_{\lfloor n\cdot \rfloor})(s), t, s\ge 0\right) \to \left(\mathcal{R}^{l,u}_{\cdot}, \mathcal{A}^{l,u}_{\cdot}, \psi^{l,u}(\mathcal{R}^{l,u}_{\cdot})(s), t, s\ge 0\right)$$

in distribution as $n\to\infty$. The above together with the convergence of the time process (2.4.14) yield the following convergence in distribution

$$\left(R^{l,u,n}_{\lfloor nt \rfloor}, A^{l,u,n}_{\lfloor nt \rfloor}, n^{-3/4}T^{l,u,n}_{ns}, t, s\ge 0\right) \to \left(\mathcal{R}^{l,u}_{\cdot}, \mathcal{A}^{l,u}_{\cdot}, \psi^{l,u}(\mathcal{R}^{l,u}_{\cdot})(s), t, s\ge 0\right)$$

as $n\to\infty$.

□

As in the proof of Theorem 2.3.1, if we apply the previous lemma and the fact that the transition operators $P^n(r_0,\alpha_0)$ and $P^n_{l,u}(r_0,\alpha_0)$ agree when $r_0\in(l,u)$, we can show analogous an result for the non cut-off processes stopped before hitting u or l. The exact statement is given by the following lemma.

Lemma 2.4.4. *Suppose* $(R^n_k, A^n_k, \Delta^n_k, k\in\mathbb{N}_0)$ *is a Markov process with transition operator* $\tilde{U}_n$*, and* $(\mathcal{R}(t),\mathcal{A}(t), t\ge 0)$ *a diffusion with generator* G *defined in* (2.3.2)*. If we fix* $l < u \in \mathcal{D}(:= [0,h]$ *or* $\mathbb{R}^+)$*, with* $r_0\in(l,u)$*, and let* $\tau^n_{l,u} = \inf\{k\ge 0 : R^n_k \ge u$ *or* $R^n_k \le l\}$ *and* $\tau_{l,u} = \inf\{t\ge 0 : \mathcal{R}_t\ge u$ *or* $\mathcal{R}_t \le l\}$*, then*

$$\left(\left(R^n_{\lfloor nt \rfloor\wedge\tau^n_{l,u}}, A^n_{\lfloor nt \rfloor\wedge\tau^n_{l,u}}, n^{-3/4}T^n_{ns\wedge\tau^n_{l,u}}\right), t, s\ge 0\right)$$
$$\to \left(\left(\mathcal{R}(t\wedge\tau_{l,u}), \mathcal{A}(t\wedge\tau_{l,u}), \int_0^{s\wedge\tau_{l,u}}\frac{d\tilde{s}}{\lambda(\mathcal{R}(\tilde{s}\wedge\tau_{l,u}))}\right), t, s\ge 0\right)$$

in distribution on $D(\mathbb{R}_+,\mathbb{R}_+)\times D(\mathbb{R}_+,\mathbb{R})\times D(\mathbb{R}_+,\mathbb{R}_+)$*.*

Let us now define the time change $\Omega_n : \mathbb{R}_+ \to \mathbb{R}_+$ by

$$\Omega_n(t) = \mathcal{I}\left(\frac{1}{n^{3/4}} T^n_{(n\cdot)\wedge\tau^n_{l,u}} + (\cdot - \frac{1}{n}\tau^n_{l,u})^+ \frac{1}{\lambda(R(\tau_{l,u}))}\right)(t) \tag{2.4.15}$$

Remark 2.4.5. The reason for the $(\cdot - \frac{1}{n}\tau^n_{l,u})^+ \frac{1}{\lambda(R(\tau_{l,u}))}$ term is so that we fall in the setting of Lemma 5.2 from [BR14], which we reproduce here for convenience.

Lemma 2.4.6. *If $f \in D(\mathbb{R}_+, \mathbb{R}_+)$ is continuous and strictly increasing with $\lim_{t\to\infty} f(t) = \infty$, then $\mathcal{I}(f) \in D(\mathbb{R}_+, \mathbb{R}_+)$ and $\mathcal{I}$ is continuous at f.*

This result is useful for proving our next lemma which shows weak convergence of the time-changed step process.

Lemma 2.4.7.

$$\left(\left(R_{\lfloor \mathrm{n}\Omega_\mathrm{n}(\mathrm{t})\rfloor\wedge\tau^n_{l,u}}, A_{\lfloor \mathrm{n}\Omega_\mathrm{n}(\mathrm{t})\rfloor\wedge\tau^n_{l,u}}\right), t \geq 0\right) \to \left(\left(\mathcal{R}(\Omega(t)\wedge\tau_{l,u}), \mathcal{A}(\Omega(t)\wedge\tau_{l,u})\right), t \geq 0\right)$$

in distribution on $D(\mathbb{R}_+, \mathbb{R}_+) \times D(\mathbb{R}_+, \mathbb{R})$ as $n \to \infty$.

Proof. By Lemma 2.4.4 and the Skorokhod representation theorem (see for example Theorem 6.7 of [Bil99]) we can construct the relevant processes on a single probability space so that

$$\begin{aligned}&\left(\left(R^n_{\lfloor \mathrm{nt}\rfloor\wedge\tau^n_{l,u}}, A^n_{\lfloor \mathrm{nt}\rfloor\wedge\tau^n_{l,u}}, n^{-3/4}T^n_{ns\wedge\tau^n_{l,u}}\right), t, s \geq 0\right)\\ &\to \left(\left(\mathcal{R}(t\wedge\tau_{l,u}), \mathcal{A}(t\wedge\tau_{l,u}), \int_0^{s\wedge\tau_{l,u}} \frac{d\tilde{s}}{\lambda(\mathcal{R}(\tilde{s}\wedge\tau_{l,u}))}\right), t, s \geq 0\right)\end{aligned}$$

a.s. in the Skorokhod topology on $D(\mathbb{R}_+, \mathbb{R}_+) \times D(\mathbb{R}_+, \mathbb{R})$.

Since $\int_0^s \frac{d\tilde{s}}{\lambda(\mathcal{R}(\tilde{s}\wedge\tau_{l,u}))}$ increases linearly in s after $s \geq \tau_{l,u}$, it is clear to see that $f_n(s) := \frac{1}{n^{3/4}} T^n_{(ns)\wedge\tau^n_{l,u}} + (s - \frac{1}{n}\tau^n_{l,u})^+ \frac{1}{\lambda(R(\tau_{l,u}))} \to \int_0^s \frac{d\tilde{s}}{\lambda(\mathcal{R}(\tilde{s}\wedge\tau_{l,u}))}$ for $s \geq \tau_{l,u}$ as well as $s \leq \tau_{l,u}$. Let

$$f(s) := \int_0^s \frac{d\tilde{s}}{\lambda(\mathcal{R}(\tilde{s}\wedge\tau_{l,u}))}$$

Since λ is bounded away from 0 on $[l, u]$, f is strictly increasing. The function f is also differentiable by the Fundamental Theorem of Calculus, so f meets the assumptions of Lemma 2.4.6. Hence $\mathcal{I}$ is continuous at f and, as $n \to, \infty$ $\Omega_n(t) = \mathcal{I}(f_n)(t) \to \mathcal{I}(f)(t) = \Omega(t)$ a.s. on $D(\mathbb{R}_+, \mathbb{R}_+)$. What is more, by Lemma 2.4.4 we have that

$$\left(R^n_{\lfloor \mathrm{nt}\rfloor\wedge\tau^n_{l,u}}, A^n_{\lfloor \mathrm{nt}\rfloor\wedge\tau^n_{l,u}}, \Omega_n(s), t, s \geq 0\right) \to \left(\mathcal{R}(t\wedge\tau_{l,u}), \mathcal{A}(t\wedge\tau_{l,u}), \Omega(s), t, s \geq 0\right)$$

a.s. on $D(\mathbb{R}_+,\mathbb{R}_+)\times D(\mathbb{R}_+,\mathbb{R})\times D(\mathbb{R}_+,\mathbb{R}_+)$ as $n\to\infty$, hence also in distribution. By the Continuous Mapping Theorem, the result will follow if we can show the function

$$\Psi : D(\mathbb{R}_+,\mathbb{R})\times D(\mathbb{R}_+,\mathbb{R})\times D(\mathbb{R}_+,\mathbb{R}_+)\to D(\mathbb{R}_+,\mathbb{R})\times D(\mathbb{R}_+,\mathbb{R}),$$
$$\Psi(f_1,f_2,\phi) := (f_1\circ\phi, f_2\circ\phi)$$

is a.s. continuous at $(\mathcal{R},\mathcal{A},\Omega)$. By considering the coordinate projections, this is equivalent to showing the composition map $(f,\phi)\mapsto f\circ\phi$ on $D(\mathbb{R}_+,\mathbb{R})$ is continuous. It can be shown that continuity of f implies continuity of the composition map in the Skorokhod topology (see for example [Bil99][pg 151]). Now, $(\mathcal{R},\mathcal{A})$ has a.s. continuous paths because it is a diffusion. The function $s\mapsto\int_0^s \frac{d\tilde{s}}{\lambda(\mathcal{R}(\tilde{s}\wedge\tau_{l,u}))}$ is a.s. strictly increasing and differentiable with derivative bounded below on compact sets. Thus $\Omega(s)=\mathcal{I}\left(\int_0^s \frac{d\tilde{s}}{\lambda(\mathcal{R}(\tilde{s}\wedge\tau_{l,u}))}\right)$ is a.s. continuous. In conclusion, Ψ is a.s. continuous at $(\mathcal{R}(\cdot\wedge\tau_{l,u}),\mathcal{A}(\cdot\wedge\tau_{l,u}),\Omega(\cdot))$ and so by the continuous mapping theorem

$$\left(R^n_{\lfloor \mathrm{n}\Omega_\mathrm{n}(\mathrm{t})\rfloor\wedge\tau^n_{l,u}}, A^n_{\lfloor \mathrm{n}\Omega_\mathrm{n}(\mathrm{t})\rfloor\wedge\tau^n_{l,u}}, t\geq 0\right)\to(\mathcal{R}(\Omega(t)\wedge\tau_{l,u}),\mathcal{A}(\Omega(t)\wedge\tau_{l,u}),t\geq 0)$$

in distribution $D(\mathbb{R}_+,\mathbb{R}_+)\times D(\mathbb{R}_+,\mathbb{R})$.

□

We can now prove the main result of this section: Theorem 2.4.1.

Proof. (Theorem 2.4.1.) To simplify the arguments, we note that $\Omega_n(t)=\mathcal{I}\left(\frac{1}{n^{3/4}}T^n_{(n\cdot)\wedge\tau^n_{l,u}}\right)(t)$ when $t\leq\frac{1}{n^{3/4}}T^n_{\tau^n_{l,u}}$.

By construction, $T^n_{\lfloor \mathrm{n}\cdot\rfloor}$ maps the reflection index k/n to the kth reflection time. Hence, the inverse map $\mathcal{I}(T^n_{\lfloor \mathrm{n}\cdot\rfloor})(t)$ maps the kth reflection time T^n_k to the reflection index k/n.

Since $\mathcal{I}\left(\frac{1}{n^{3/4}}T^n_{(n\cdot)\wedge\tau^n_{l,u}}\right)(t)=\mathcal{I}\left(T^n_{(n\cdot)\wedge\tau^n_{l,u}}\right)(n^{3/4}t)$, we have that $n\Omega_n\left(\frac{1}{n^{3/4}}T^n_k\right)=k$ whenever $k\leq\tau^n_{l,u}$. Or equivalently,

$$\mathcal{R}(n^{3/4}t)=R^n_{n\Omega_n(t)}=R^n_k$$

whenever $t=\frac{T^n_k}{n^{3/4}}$ and $k\leq\tau^n_{l,u}$. As a consequence, $n^{3/4}t\in[T^n_{k-1},T^n_k]$ implies $\Omega_n(t)\in[k-1,k)$, so by the piecewise definition of the full trajectory $R^n(t)$ (equation (2.4.2)), if we fix $S>0$ we have

$$\begin{aligned}&\sup_{0\leq t\leq S}\left|R^n(n^{3/4}t\wedge T^n_{\tau^n_{l,u}})-R^n_{\lfloor \mathrm{n}\Omega_\mathrm{n}(\mathrm{t})\rfloor\wedge\tau^n_{l,u}}\right|\\ &\leq\sup_{k\leq n\Omega_n(S)\wedge T^n_{\tau^n_{l,u}}}\ \sup_{t\in[T^n_{k-1},T^n_k]}\left|r_n\left(R^n_{k-1},\Theta_{k-1},t-T^n_{k-1}\right)-R^n_{k-1}\right|.\end{aligned}\tag{2.4.16}$$

By fixing $M, C > 0$, applying a union bound and utilizing the fact that for each k, $T_k^n - T_{k-1}^n$ is distributed like $N^{(n)}$, we have the following bounds

$$\begin{aligned}
&\mathbb{P}\left\{\sup_{0\leq t\leq S} |R^n(n^{3/4}t \wedge T^n_{\tau^n_{l,u}}) - R^n_{\lfloor n\Omega_n(t)\rfloor \wedge \tau^n_{l,u}}| > \epsilon, \Omega_n(S \wedge T^n_{\tau^n_{l,u}}) \leq M\right\} \\
&\leq nM \sup_{(r_0,\theta)\in[l,u]\times[-\pi,\pi]} \mathbb{P}\left\{\sup_{t\in[0,N^{(n)}(\theta)]} |r_n(r_0,\theta,t) - r_0| > \epsilon, \Omega_n(S \wedge T^n_{\tau^n_{l,u}}) \leq M\right\} \\
&\leq nM \sup_{(r_0,\theta)\in[l,u]\times[-\pi,\pi]} \mathbb{P}\left\{\sup_{t\in[0,N^{(n)}(\theta)\wedge C]} |r_n(r_0,\theta,t) - r_0| > \epsilon, \Omega_n(S \wedge T^n_{\tau^n_{l,u}}) \leq M\right\} \\
&+ nM \sup_{(r_0,\theta)\in[l,u]\times[-\pi,\pi]} \mathbb{P}\left\{N^{(n)}(r_0,\Theta) > C\right\}.
\end{aligned}$$

By Lemma 2.2.3, for n large enough

$$\sup_{(r_0,\theta)\in[l,u]\times[-\pi,\pi]} \sup_{t\in[0,C]} |r_n(r_0,\theta,t) - r_0| < \epsilon.$$

So for such n, $\mathbb{P}\{\sup_{t\in[0,C]} |r_n(r_0,\theta,t) - r_0| > \epsilon\} = 0$ for all (r_0,θ) in $[l,u]\times[-\pi,\pi]$. Also, by Corollary 2.2.5, $nM \sup_{(r_0,\theta)\in[l,u]\times[-\pi,\pi]} \mathbb{P}\left\{N^{(n)}(r_0,\Theta) > C\right\}$ approaches 0 as $n \to \infty$. So

$$\lim_{n\to\infty} \mathbb{P}\left\{\sup_{0\leq t\leq S} |R^n(n^{3/4}t \wedge T^n_{\tau^n_{l,u}}) - R^n_{\lfloor n\Omega_n(t)\rfloor \wedge \tau^n_{l,u}}| > \epsilon, \Omega_n(S \wedge T^n_{\tau^n_{l,u}}) \leq M\right\} = 0. \tag{2.4.17}$$

By the proof of Lemma 2.4.7, we have

$$\left((\Omega_n(t), n^{-1}\tau^n_{l,u}), t \geq 0\right) \to ((\Omega(t), \tau_{l,u}), t \geq 0)$$

in distribution as $n \to \infty$. So for $M > 0$ large enough we have

$$\lim_{n\to\infty} \mathbb{P}\left\{\Omega_n(t) \wedge T^n_{\tau^n_{l,u}} \leq M\right\} = 1.$$

From this and (2.4.17) it follows that

$$\lim_{n\to\infty} \mathbb{P}\left\{\sup_{0\leq t\leq S} |R^n(n^{3/4}t \wedge T^n_{\tau^n_{l,u}}) - R^n_{\lfloor n\Omega_n(t)\rfloor \wedge \tau^n_{l,u}}| > \epsilon\right\} = 0.$$

for any any $\epsilon, S > 0$. Furthermore, by (2.2.58), we can show that

$$\sup_{(r_0,\alpha_0,\theta)\in[l,u]\times\mathbb{R}\times[-\pi,\pi]} \sup_{t\in[0,C]} |\alpha_n(r_0,\alpha_0,\theta,t) - \alpha_0| < \epsilon.$$

Calculations identical to those in (2.4.17) show that

$$\lim_{n\to\infty} \mathbb{P}\left\{\sup_{0\leq t\leq S} |A^n(n^{3/4}t \wedge T^n_{\tau^n_{l,u}}) - A^n_{\lfloor n\Omega_n(t)\rfloor \wedge \tau^n_{l,u}}| > \epsilon\right\} = 0.$$

So by Lemma 2.4.7 it follows that

$$\left(\left(R^n(n^{3/4}t \wedge T^n_{\tau^n_{l,u}}), A^n(n^{3/4}t \wedge T^n_{\tau^n_{l,u}})\right), t \geq 0\right) \to ((\mathcal{R}(\Omega(t) \wedge \tau_{l,u}), \mathcal{A}(\Omega(t) \wedge \tau_{l,u})), t \geq 0)$$

in distribution. Recall that that $\iota^n_{l,u} = \inf\{t : R^n(t) \leq l \text{ or } R^n(t) \geq u\}$ is the first time the full time trajectory $R^n(t)$ exits the interval (l, u) and $\tau^n_{l,u} = \inf\{k : R^n_k \leq l \text{ or } R^n_k \geq u\}$. Together these imply $T^n_{\tau^n_{l,u}}$ is the time of first reflection after the Markov process R^n_k exits (l, u). An immediate consequence is that $\iota^n_{l,u} \leq T^n_{\tau^n_{l,u}}$. Then by continuity of $\left(R^n(n^{3/4}t \wedge T^n_{\tau^n_{l,u}}), A^n(n^{3/4}t \wedge T^n_{\tau^n_{l,u}})\right)$ and $(\mathcal{R}(\Omega(t) \wedge \tau_{l,u}), \mathcal{A}(\Omega(t) \wedge \tau_{l,u}))$ we conclude that

$$\left(R^n(n^{3/4}t \wedge \iota^n_{l,u}), A^n(n^{3/4}t \wedge \iota^n_{l,u}), t \geq 0\right) \to ((\mathcal{R}(\Omega(t) \wedge \tau_{l,u}), \mathcal{A}(\Omega(t) \wedge \tau_{l,u})), t \geq 0)$$

in distribution.

Lastly, by Theorem 8.5.1 from [Øks03], if $\mathcal{G}$ is the generator of $(\mathcal{R}(t), \mathcal{A}(t), t \geq 0)$ then the generator of $(\mathcal{R}(\Omega(t)), \mathcal{A}(\Omega(t)), t \geq 0)$ will be $\lambda(\rho)\mathcal{G}$. Since

$$\mathscr{G} = \lambda(\rho)\mathcal{G}$$

this completes the proof. □

2.5 Classifying the boundaries of $\mathcal{D}$

In this section we want to give conditions under which the boundary points of $\mathcal{D}$ are inaccessible. This would imply that the boundary points cannot be reached in finite time. These results can then be used to remove the stopping at u and/or l in Theorem 2.1.1.

Suppose we have a regular diffusion X with state space the interval (ℓ, r). Every diffusion has two basic characteristics: the speed measure $m(dx)$ and the scale function $s(x)$.

We assume the infinitesimal generator $G : \mathcal{D}(G) \mapsto \mathcal{C}_b(I)$ of X is a second order differential operator

$$Gf(x) = \frac{1}{2}\sigma^2(x)\partial_{xx}f(x) + \mu(x)\partial_x f(x)$$

where $\sigma, \mu \in C(I)$ and $\sigma^2(x) > 0$ for all $x \in I$. Then it is well-known that

- The speed measure is absolutely continuous with respect to Lebesgue measure and has density

$$m'(x) = 2\sigma^{-2}(x)e^{B(x)}$$

- The scale function has density

$$s'(x) = e^{-B(x)}$$

with $B(x) := \int_a^x 2\sigma^{-2}(y)\mu(y)\,dy$ for some arbitrary (fixed) $a \in I$. The domain $\mathcal{D}(G)$ consists of all functions in $\mathcal{C}_b(I)$ such that $Gf \in \mathcal{C}_b(I)$ together with the appropriate boundary conditions.

The boundary point ℓ is called *accessible* when

$$\int_\ell^x \left(\int_y^x m'(\eta)\,d\eta\right) s'(y)\,dy < \infty$$

and *inaccessible* when

$$\int_\ell^x \left(\int_y^x m'(\eta)\,d\eta\right) s'(y)\,dy = \infty.$$

Similarly, one can classify the boundary r.

As we have shown above, the one-dimensional diffusion defining the limiting radial process has generator $\mathcal{G}_r$ that acts on compactly supported functions in $C^2(\mathcal{D}^\circ)$ as

$$\begin{aligned}\mathcal{G}_r f(\rho) &:= \mu_r(\rho)f'(\rho) + \frac{\sigma_r^2(\rho)}{2}f''(\rho)\\ &= \frac{1}{g^2(\rho)}\left(-\frac{g'(\rho)}{2g(\rho)} + \frac{1}{2\rho} - \frac{\partial_r\mathcal{U}(\rho)}{4(E-\mathcal{U}(\rho))}\right)f'(\rho) + \frac{1}{2g^2(\rho)}f''(\rho).\end{aligned} \tag{2.5.1}$$

The density of the scale function will be

$$\begin{aligned}s'(y) &= \exp\left(-\int_a^y \frac{2\mu_r(\rho)}{\sigma_r^2(\rho)}d\rho\right)\\ &= \exp\left(\int_a^y\left(\frac{g'(\rho)}{g(\rho)} - \frac{1}{\rho} + \frac{\partial_r\mathcal{U}(\rho)}{2(E-\mathcal{U}(\rho))}\right)d\rho\right)\\ &= \exp\left(\ln(g(y)) - \ln(y) - \frac{1}{2}\ln(E-\mathcal{U}(y)) - C\right) = \exp\left(\ln\left(\frac{g(y)}{y\sqrt{E-\mathcal{U}(y)}}\right) - C\right)\\ &= \frac{a\sqrt{E-\mathcal{U}(a)}}{g(a)}\left(\frac{g(y)}{y\sqrt{E-\mathcal{U}(y)}}\right)\end{aligned}$$

where $C := \ln\left(\frac{g(a)}{a\sqrt{E-\mathcal{U}(a)}}\right)$. As a result, if we fix an arbitrary $c \in \mathcal{D}^\circ$ the scale function will be given by

$$s(x) = \int_c^x s'(y)dy = \frac{a\sqrt{E-\mathcal{U}(a)}}{g(a)}\int_c^x\left(\frac{g(y)}{y\sqrt{E-\mathcal{U}(y)}}\right)dy. \tag{2.5.2}$$

The speed measure density for the radial diffusion will be

$$m'(x) = \frac{2}{\sigma_r^2(x)s'(x)} = \frac{2g(a)}{a\sqrt{E-\mathcal{U}(a)}}\left(g(x)x\sqrt{E-\mathcal{U}(x)}\right). \tag{2.5.3}$$

One can then find the speed measure by setting

$$m(J) := \int_J m'(x)dx = \frac{2g(a)}{a\sqrt{E-\mathcal{U}(a)}} \int_J g(x)x\sqrt{E-\mathcal{U}(x)}dx$$

for any Lebesque measurable $J \subseteq D^\circ$.

2.5.1 Examples

As an illustrative example, we consider case when the potential is a gravitational force directed towards the origin. That is, the potential function is given by

$$\mathcal{U}(\rho) = -\frac{k}{\rho}$$

for some constant $k > 0$. Suppose the the total energy of the particle is positive $E > 0$. We have

$$s(x) = \frac{\sqrt{(Ea+k)a}}{g(a)} \int_c^x \left(\frac{g(y)}{\sqrt{(Ey+k)y}} \right) dy.$$

Since g is bounded above and bounded away from 0 on $\mathbb{R}_+$, for $y \downarrow 0$,

$$\frac{g(y)}{\sqrt{(Ey+k)y}} = \mathcal{O}\left(\frac{1}{\sqrt{y}}\right)$$

which implies $\lim_{x\downarrow 0} |s(x)| < \infty$. For $y \uparrow \infty$ one has

$$\frac{g(y)}{\sqrt{(Ey+k)y}} = \mathcal{O}\left(\frac{1}{y}\right)$$

so that $\lim_{x\uparrow\infty} s(x) = \infty$. The speed measure density is given by

$$m'(x) = \frac{g(a)}{a\sqrt{E-\mathcal{U}(a)}} g(x)\sqrt{(Ex+k)x}$$

and the speed measure

$$M(J) = \frac{g(a)}{a\sqrt{E-\mathcal{U}(a)}} \int_J g(x)\sqrt{(Ex+k)x}dx$$

for $x \in \mathbb{R}_+$. If we fix an arbitrary $b > 0$, then we can observe that $\lim_{x\downarrow 0} M([x,b]) < \infty$, and and $\lim_{x\uparrow\infty} M([b,x]) = \infty$. We can then apply the boundary classification from [KT81][Table 6.2].

$$\lim_{x\downarrow 0} |s(x)| < \infty, \quad \lim_{x\downarrow 0} M([x, b]) < \infty$$

which implies that 0 is a regular boundary, and

$$\lim_{x\uparrow\infty} s(x) = \infty, \quad \lim_{x\uparrow\infty} M([b, x]) = \infty$$

which implies that ∞ is a natural boundary. As a result 0 is accessible while ∞ is inaccessible.

Chapter 3

Reconstruction of a manifold from the intrinsic metric of an associated Markov chain

3.1 Introduction

We study the Carathéodory metric ρ_N associated with the infinitesimal generator G_N of a continuous time Markov Chain on a finite weighted graph Λ_N. Our goal is to prove a rescaling of the sequence $\{\rho_N\}_N$ converges when the corresponding sequence of rescaled generators $\{G_N\}_N$ converges to a diffusive limit. Our basic set up is the following: the vertices of Λ_N are a set of random data points $\{x_i\}_{i=1}^N$ sampled independent from a distribution P on a compact Riemannian manifold $\mathcal{M}$ which is embedded in $\mathbb{R}^n$ and equipped with Riemannian volume measure. Λ_N is a complete graph on these data points, and has edge weights assigned by a kernel function K. The transition probabilities for the random walk are constructed using these edge weights which are chosen to reflect how close different points are in the ambient Euclidean space, but also attempt to just capture "local" similarities. The specifics of the construction are discussed in complete detail in the following section.

The Carathéodory metric, also known as the intrinsic metric, is a metric associated with a Dirichlet form. Since the generator of a reversible Markov chain defines a Dirichlet form in a natural way, we can associate a Carathéodory metric with this form. These metrics are studied extensively in [Stu94] and in [Dav93] where they are used to establish bounds on heat kernels on weighted graphs. More specifically, the Carathéodory metric defined on a graph with edge weights can be found in Section 4 of [Dav93], where it is assigned the notation d_3. It is observed in that paper that this metric arises naturally from considerations in noncommutative geometry.

Under a proper rescaling and with a suitable choice of weight functions, as the number of data points we sample goes to infinity, the infinitesimal generators G_N converge in the supremum norm to the Laplace-Beltrami operator on $\mathcal{M}$ (see for example Theorem 3 of [THJ11]). In other words, a suitably sped up version of the Markov chain associated with G_N converges to the natural Brownian motion on the manifold. This leads us to conjecture that after a suitable rescaling, the Carathéodory metric ρ_N associated with G_N also converges in some sense to the Carathéodory metric associated with the Laplace-Beltrami operator, which is the geodesic distance metric on $\mathcal{M}$. A more precise statement of the conjecture is given below in Theorem 3.2.5. If this theorem holds, then for large N, ρ_N rescaled will be able to approximate the distances between the points $\{x_1, \dots, x_N\}$ in our embedded manifold. Since the geodesic metric uniquely identifies the Riemannian inner product structure on a manifold, this shows that in essence we can reconstruct the manifold $\mathcal{M}$ by sampling a large number of data points from the distribution P and computing the intrinsic metric distances between these data points.

Before going further, we first give a precise definition for the Carathéodory metric. Suppose $\mathscr{E}$ is a Dirichlet form with domain $\mathcal{D}(\mathscr{E})$ on a real Hilbert Space $H = L^2(X, m)$, where X is a locally compact separable Hausdorff with positive measure m fully supported on X (see [Stu94] for the full details). Then it can be show than $\mathscr{E}$ has the form

$$\mathscr{E}(u, v) = \int_X d\Gamma(u, v) \tag{3.1.1}$$

where Γ is a symmetric bilinear form on $\mathcal{D}(\mathscr{E})$ with values in the signed Radon measures on X. Γ is known as the energy measure for $\mathscr{E}$. $\Gamma(u, u)$ is defined by

$$\int_X \phi d\Gamma(u, u) = \mathscr{E}(u, \phi u) - \frac{1}{2}\mathscr{E}(u^2, \phi)$$

for every $u \in \mathcal{D}(\mathscr{E}) \cap L^\infty(X, m)$ and every $\phi \in \mathcal{D}(\mathscr{E}) \cap \mathcal{C}_0(X)$. Γ can be extended to any $u, v \in \mathcal{D}(\mathscr{E}) \cap L^\infty(X, m)$ by polarization:

$$\Gamma(u, v) = \frac{1}{4}\left(\Gamma(u + v, u + v) - \Gamma(u - v, u - v)\right).$$

We can then define the Carathéodory metric associated with $\mathscr{E}$ in terms of the corresponding energy measure Γ by

$$\rho(x, y) = \sup\left\{u(x) - u(y) : u \in \mathcal{D}_{\mathbf{loc}}(\mathscr{E}) \cap \mathcal{C}(X), d\Gamma(u, u) \leq dm \text{ on } X\right\} \tag{3.1.2}$$

where $\mathcal{D}_{\mathbf{loc}}(\mathscr{E})$ is the set of all functions in $\mathcal{D}(\mathscr{E})$ whose restriction to any compact set $S \subset X$ is in $L^2(S)$, and $d\Gamma(u, u) \leq dm$ means the energy measure $\Gamma(u, u)$ is absolutely continuous with respect to the reference measure m and with Radon-Nikodym derivative $\frac{d}{dm}\Gamma(u, u) \leq 1$.

For example, suppose $\Delta_\mathcal{M}$ is the Laplace-Beltrami operator on the compact Riemmanian manifold $\mathcal{M}$. Then the Dirichlet form associated with $\Delta_\mathcal{M}$ is given by

$$\mathscr{E}_\mathcal{M}(u, v) := -\int_\mathcal{M} \langle u(x), \Delta_\mathcal{M} v(x)\rangle d\text{vol}_\mathcal{M} x$$

with domain $\mathcal{D}(\mathscr{E}_{\mathcal{M}}) = \mathcal{C}^\infty(\mathcal{M})$, and where $d\text{vol}_{\mathcal{M}}x$ is the volume element associated with $\mathcal{M}$. By the definition of $\Delta_{\mathcal{M}}$ we have

$$\mathscr{E}_{\mathcal{M}}(u,u) = \int_{\mathcal{M}} |\nabla u|^2 \, d\text{vol}_{\mathcal{M}}x$$

which shows the energy density with respect to the volume element $d\text{vol}_{\mathcal{M}}x$ is given by $d\Gamma_{\mathcal{M}} = |\nabla u|^2$. By (3.1.2), the Carathéodory metric associated with $\mathscr{E}_{\mathcal{M}}$ is defined by

$$\rho_{\mathcal{M}}(x,y) := \sup\left\{u(x) - u(y) : |\nabla u|^2 \leq 1 \text{ on } \mathcal{M}\right\}, \quad x, y \in \mathcal{M}$$

which is the geodesic distance on $\mathcal{M}$.

3.2 Construction of the Graph Carathéodory Metric

We now provide a derivation of the second order cone problem that defines our main object of study: the Carathéodory metric ρ_N associated with the infinitesimal generator G_N for the Markov chain which we also now define. We let $\mathcal{D}_N = \{x_1, \cdots, x_N\}$ denote the data points which constitute the vertex set of the weighted graph Λ_N. We let $W_N = (w_{ij})_{i,j=1}^N$ denote the $N \times N$ symmetric matrix of edge *weights* between the points of $\mathcal{D}_N$. Here w_{ij} represents the weight assigned to the edge connecting x_i and x_j. w_{ij} is assumed to have the form

$$w_{ij} = K\left(\frac{\|x_i - x_j\|^2}{h_N^2}\right)$$

for some *kernel* function $K : \mathbb{R} \mapsto \mathbb{R}_+$, and some parameter $h_N \in \mathbb{R}_+$ known as the *bandwidth.* $\|\cdot\|$ is the usual Euclidean distance in the ambient space $\mathbb{R}^n$. K is assumed to be exponential decreasing, and as $N \to \infty$, $h_N \downarrow 0$ so that only edges connecting points close in the ambient space are assigned numerically significant weights. For convenience, we introduce the function

$$K_h(x,y) : \mathcal{M} \times \mathcal{M} \mapsto \mathbb{R}, \quad K_h(x,y) := K\left(\frac{\|x-y\|^2}{h^2}\right).$$

The exact assumptions on K are given in the next section. However, a standard choice for K is $K(x) = e^{-x}$ so that

$$w_{ij}^N = \exp(-\|x_i - x_j\|^2/h_N^2), 1 \leq i, j \leq N,$$

Remark 3.2.1. Note that the weights w_{ij} depend on N, but to simplify notation we leave this dependence on N implicit. We follow this same convention throughout for the entries of other matrices and vectors which depend on N.

For a given matrix of weights, W_N, we define a stochastic matrix $\Pi_N = (p_{ij})_{i,j=1}^N$ via

$$p_{ij} = \frac{w_{ij}}{\sum_{k=1}^N w_{ik}}, \quad 1 \le i, j \le N. \tag{3.2.1}$$

The matrix Π_N is reversible with respect to the stationary distribution π given by

$$\pi_i = \frac{\sum_{k=1}^N w_{ik}}{\sum_{j=1}^N \sum_{k=1}^N w_{jk}};$$

that is, $\pi_i p_{ij} = \pi_j p_{ji}$ for $1 \le i, j \le N$.

The matrix $G_N := \Pi_N - I = (g_{ij})_{i,j=1}^N$ is the infinitesimal generator of a continuous time Markov chain. G_N is also reversible with respect to π, and $\sum_j g_{ij} = 0$ for $1 \le i \le N$.

As with any reversible Markov process, we can associate the Markov chain with infinitesimal generator G_N to the corresponding *Dirichlet form*. In our case, this will be the nonnegative definite, symmetric bilinear form $\mathscr{E}_N$ on $\mathbb{R}^N$ given by

$$\begin{aligned}
\mathscr{E}_N(u,v) &= -\sum_i \pi_i u_i (G_N v)_i \\
&= -\sum_{i,j} u_i v_j \pi_i g_{ij} \\
&= \frac{1}{2} \sum_{i,j} (u_i - u_j)(v_i - v_j) \pi_i g_{ij} \\
&= \frac{1}{2} \sum_{i,j} (u_i - u_j)(v_i - v_j) \pi_i p_{ij}.
\end{aligned}$$

Following [Stu94], we associate each pair $u, v \in \mathbb{R}^N$ with the signed energy measure $\Gamma_N(u,v)$ on $\mathcal{D}_N$ defined in (3.1.1). So

$$\mathscr{E}_N(u,v) = \int_{\mathcal{D}_N} d\Gamma(u,v). \tag{3.2.2}$$

Of course, in our case we can just think of $\Gamma_N(u,v)$ as an element of $\mathbb{R}^N$. Specializing the general recipe for Γ_N to our situation, we have that

$$\int_{\mathcal{D}_N} \phi \, d\Gamma_N(u,u) = \mathscr{E}_N(u, \phi u) - \frac{1}{2} \mathscr{E}_N(u^2, \phi) \tag{3.2.3}$$

for all $\phi \in \mathbb{R}^N$. Here we think of u as a function from $\mathcal{D}_N$ to $\mathbb{R}$, and so u^2 and ϕu are defined coordinate-wise. Thus,

$$\begin{aligned}\Gamma_N(u,u)_k &= \int_{\mathcal{D}_N} \delta_k \, d\Gamma_N(u,u) \\ &= \mathscr{E}_N(u,\delta_k u) - \frac{1}{2}\mathscr{E}_N(u^2,\delta_k) \\ &= -\sum_i u_i u_k \pi_i q_{ik} + \frac{1}{2}\sum_i u_i^2 \pi_i g_{ik} \\ &= \frac{1}{2}\sum_i (u_i - u_k)^2 \pi_i g_{ik} \\ &= \frac{1}{2}\sum_i (u_i - u_k)^2 \pi_i p_{ik}.\end{aligned}$$

Furthermore, $\Gamma_N(u,v)$ is given by polarization; that is,

$$\Gamma_N(u,v)_k = \frac{1}{4}\left[\Gamma_N(u+v,u+v)_k - \Gamma_N(u-v,u-v)_k\right],$$

so that

$$\begin{aligned}\Gamma_N(u,v)_k &= \frac{1}{2}\sum_i (u_i-u_k)(v_i-v_k)\pi_i g_{ik} \\ &= \frac{1}{2}\sum_i (u_i-u_k)(v_i-v_k)\pi_i p_{ik}.\end{aligned}$$

Because $\pi_i p_{ij} = \pi_j p_{ji}$ and $\pi_i g_{ij} = \pi_j g_{ji}$, we have the alternative formulae

$$\begin{aligned}\Gamma_N(u,v)_k &= \frac{1}{2}\sum_i (u_i-u_k)(v_i-v_k)\pi_k g_{ki} \\ &= \frac{1}{2}\sum_i (u_i-u_k)(v_i-v_k)\pi_k p_{ki}.\end{aligned}$$

In our setting, the procedure for constructing a metric ρ_N on $\mathcal{D}$ from Γ_N reduces to the formula

$$\rho_N(x_s,x_t) = \sup\{u_s - u_t : u \in \mathbb{R}^N,\ \Gamma_N(u,u)_k \le \pi_k,\ 1 \le k \le N\}, \quad s,t \in [N]. \tag{3.2.4}$$

That is,

$$\rho_N(x_s,x_t) = \sup\{u_s - u_t : \frac{1}{2}\sum_i (u_i-u_k)^2 p_{ki} \le 1,\ 1 \le k \le N\}, \quad s,t \in [N]. \tag{3.2.5}$$

For the sake of completeness, we give the proof that ρ_N is a genuine metric on $\mathcal{D}$ when the matrix Π_N is irreducible. It is clear that $\rho_N(x_s, x_s) = 0$ for all $s \in [N]$. Since all the vectors u in some ball around 0 satisfy the constraints $\frac{1}{2}\sum_i (u_i - u_k)^2 p_{ki} \leq 1$, $1 \leq k \leq N$, it is clear that $\rho_N(x_s, x_t) > 0$ for all $s, t \in [N]$ with $s \neq t$. Also, if the vector u satisfies these constraints, then so does the vector $-u$, and hence $\rho_N(x_s, x_t) = \rho_N(x_t, x_s)$ for all $s, t \in [N]$. By the irreducibility of Π_N, if $s, t \in [N]$, then there exist $s = i_1, i_2, \ldots, i_k = t$ in $[N]$ such that $p_{i_j, i_{j+1}} > 0$ for $1 \leq j \leq k-1$. If the vector u satisfies the constraints, then $(u_{i_{j+1}} - u_{i_j})^2 p_{i_j, i_{j+1}} \leq 2$ and so

$$\rho_N(s,t) \leq \sum_{j=1}^{k-1} |u_{i_{j+1}} - u_{i_j}| \leq \sqrt{2} \sum_{j=1}^{k-1} (p_{i_j, i_{j+1}})^{-\frac{1}{2}} < \infty.$$

Lastly, by the triangle inequality $\rho_n(x_r, x_t) \leq \rho_N(x_r, x_s) + \rho_N(x_s, x_t)$ for all $r, s, t \in [N]$ is immediate from the observation that $u_r - u_t = (u_r - u_s) + (u_s - u_t)$. For fixed $s, t \in [N]$ then, we can pose the computation of $\rho_N(x_s, x_t)$ as a second-order cone program as discussed in [LVBL98]. If we note that maximizing $u_s - u_t$ is equivalent to minimizing $u_t - u_s$, then $-\rho(s,t)$ is the value of the following second-order cone program:

$$\begin{aligned} &\text{minimize } (e_t - e_s)^T u \\ &\text{subject to } \|A_k^N u\| \leq 1, \quad 1 \leq k \leq N, \end{aligned} \tag{3.2.6}$$

where e_i the column vector that has a 1 in coordinate i and 0 elsewhere, the superscript T denotes the transpose, and A_k^N is the $N \times N$ matrix $\frac{1}{\sqrt{2}}\text{diag}(\sqrt{p_{k\cdot}})(I - \tilde{1}e_k^T)$, where $p_{k\cdot}$ is the row vector given by the k^{th} row of the matrix P and $\tilde{1}$ is the column vector whose entries are all 1.

The problem (3.2.6) is clearly *strictly feasible* in the terminology of [LVBL98]; that is, there is a vector u such that $\|A_k^N u\| < 1$ for $1 \leq k \leq N$ (all the vectors in some ball around 0 satisfy this condition). Moreover, as we have shown above, the value of (3.2.6) is finite for all $s, t \in [N]$ and so any vector $u \in \mathbb{R}^N$ that satisfies $A_k^N u = 0$ for $1 \leq k \leq N$ must also satisfy $u_t - u_s = 0$ for all $s, t \in [N]$ and hence be a multiple of $\tilde{1}$.

From (29) of [LVBL98], the dual problem to (3.2.6) is

$$\begin{aligned} &\text{maximize } -\sum_{k=1}^N w_k \\ &\text{subject to } \sum_{k=1}^N (A_k^N)^T z_k = e_t - e_s \\ &\qquad \text{and } \|z_k\| \leq w_k, \quad 1 \leq k \leq N \end{aligned} \tag{3.2.7}$$

where the unknowns z_k belong to $\mathbb{R}^N$ and the unknowns w_k belong to $\mathbb{R}$.

As observed in Section 4 of [LVBL98], the strong feasibility of the primal problem (3.2.6) implies that the dual problem (3.2.7) is *feasible*, the optima are attained in both the primal and the dual problems, and these optima are equal.

It is clear that we can rewrite the dual problem (3.2.7) as

$$\begin{aligned} \text{maximize} \quad & -\sum_{k=1}^{N} \|z_k\| \\ \text{subject to} \quad & \sum_{k=1}^{N} (A_k^N)^T z_k = e_t - e_s. \end{aligned}$$

Because of the sign change we made leading to (3.2.6), we see that $\rho_N(x_s, x_t)$ is the value of the program

$$\begin{aligned} \text{minimize} \quad & \sum_{k=1}^{N} \|z_k\| \\ \text{subject to} \quad & \sum_{k=1}^{N} (A_k^N)^T z_k = e_t - e_s. \end{aligned}$$

Sufficient conditions for optimality of the primal problem are given by (32),(33) and (34) in Section 4 of [LVBL98]. These conditions generalize the *complementary slackness* conditions between the primal and dual solutions in the linear programming setting. That is, a point u is optimal for (3.2.6) if u is feasible for (3.2.6) and there exists a dual-feasible set $\{z_k\}_{k=1}^N$ such that u and $\{z_k\}_{k=1}^N$ satisfy

$$\begin{aligned} \|A_k^N u\| < 1 &\implies \|z_k\| = 0 \\ \|A_k^N u\| = 1 &\implies z_k = -\|z_k\| A_k^N u \end{aligned} \tag{3.2.8}$$

If both the primal and dual problems are strictly feasible, then the generalized complementary slackness conditions are also necessary.

3.2.1 Assumptions

We now describe in detail our underlying assumptions on the data points, the manifold $\mathcal{M}$ and the kernel K. These assumptions will be required to show convergence of the metrics $\{\rho_N\}_N$.

Assumption 3.2.2. (Data Point Assumptions)

1. The data points $\{x_i\}_{i=1}^\infty$ are independent random variables sampled from a distribution P on $\mathcal{M}$. $\mathcal{D}_N = \{x_1, \ldots, x_N\}$ denotes the first N points of this set.

2. We suppose P has density $p \in \mathcal{C}^3(\mathcal{M})$ with respect to the natural volume element $d\text{vol}_{\mathcal{M}}$ on $\mathcal{M}$, and p is bounded below by some $p_{\min} > 0$ on $\mathcal{M}$.

Assumption 3.2.3. (Manifold Assumptions) We assume $\mathcal{M}$ is an m-dimensional compact manifold and is isometrically and smoothly embedded in $\mathbb{R}^n$. The embedding allows us to think of the tangent spaces to the manifold as hyperplanes in $\mathbb{R}^n$ and we can use the usual Euclidean inner product on $\mathbb{R}^n$ to put a Riemannian inner product on these tangent spaces. In addition, we will assume our m-dimensional manifold $\mathcal{M}$ satisfies the conditions from Assumption 1 of [HAvL05] which we restate here:

1. The boundary $\partial\mathcal{M}$ of $\mathcal{M}$ is empty.
2. $\mathcal{M}$ has a bounded second fundamental form.
3. $\mathcal{M}$ has bounded sectional curvature.
4. For any $x \in \mathcal{M}$, the regularity radius r at x is defined by

$$r(x) = \sup\left\{r > 0 |\ \|x - y\|^2 \geq \frac{1}{2} d^2_{\mathcal{M}}(x, y) \quad \forall y \in B_{\mathcal{M}}(x, y)\right\}.$$

   We assume r is continuous and for any $x \in \mathcal{M}$, $r(x) > 0$.

5. for any $x \in \mathcal{M}$, $\delta(x) := \inf_{y \in M \setminus B_{\mathcal{M}}(x, \frac{1}{3}\min \text{inj}(x), r(x))} \|x - y\| > 0$ where inj denotes the injectivity radius x.

We suppose also that we have have a kernel K satisfying Assumption 2 of [HAvL05], which we also restate here.

Assumption 3.2.4. (Kernel assumptions) $K : \mathbb{R}_+ \mapsto \mathbb{R}$ is measurable, non-negative and non-increasing. In addition,

1. $K \in \mathcal{C}^2(\mathbb{R}_+)$. In particular, this implies K and $\frac{\partial^2 K}{\partial x^2}$ are bounded.
2. K, $|\frac{\partial K}{\partial x}|$ and $|\frac{\partial^2 K}{\partial x^2}|$ have exponential decay; that is, there exist c, α, and A in $\mathbb{R}_+$ such that for any $t \geq A$, $f(t) \leq ce^{-\alpha t}$, where $f(t) = \max\{K(t), |\frac{\partial K}{\partial x}|(t), |\frac{\partial^2 K}{\partial x^2}|(t)\}$.

These assumptions allow us to apply the machinery from Theorem 4 of [THJ11] and Theorem 3 of [HAvL05]. Both of these theorems show convergence of weighted graph Laplacians to a weighted Laplace-Beltrami operator on $\mathcal{M}$.

Under these assumptions, we conjecture the sequence $\{\rho_N\}_N$ converges in the following sense.

Theorem 3.2.5. *Let $d_{\mathcal{M}}$ denote the geodesic distance metric on $\mathcal{M}$. Then*

$$\lim_{N\to\infty} \|Ch_N \cdot \rho_N - d_{\mathcal{M}}|_{\mathcal{D}_N \times \mathcal{D}_N}\|_\infty = 0 \quad a.s.$$

where $\{h_N\}$ is an appropriately defined sequence of bandwidth functions going to 0 as $N \to \infty$, and C is a constant depending on $\mathcal{M}$ and K. In particular, $C = \sqrt{M_2/2M_0}$ where $M_0 := \int_{\mathbb{R}^m} K(\|y\|^2)dy$ and $M_2 := \int_{\mathbb{R}^m} K(\|y\|^2)y_1^2 dy$. For the lower bound estimates, we require that $\lim_{N\to\infty} h_N^{m+4} \frac{N}{\log N} \to \infty$.

We let $\tilde{\rho}_N := \sqrt{\frac{M_2}{2M_0}}\rho_N$ denote the rescaled metric. In the next section we prove $d_{\mathcal{M}}$ is a limiting lower bound for the sequence $\{\tilde{\rho}_N\}_N$ in the sense that $d_{\mathcal{M}}|_{\mathcal{D}_N \times \mathcal{D}_N} \leq \tilde{\rho}_N + \mathcal{O}(h_N)$. In section 3.4 we give some heuristic arguments for why $d_{\mathcal{M}}$ is a limiting upper bound, but we have not yet found a rigorous proof.

3.3 Lower bound limits for the graph Carathéodory metric

In this section we prove the geodesic metric $d_{\mathcal{M}}$ is a limiting lower bound for the rescaled Carathéodory metric in the sense of the following proposition.

Proposition 3.3.1. *As above, let $\tilde{\rho}_N := \sqrt{\frac{M_2}{2M_0}} h \cdot \rho_N$. Then almost surely*

$$d_{\mathcal{M}}|_{\mathcal{D}_N \times \mathcal{D}_N} \leq \tilde{\rho}_N + \mathcal{O}(h_N)$$

where $\mathcal{O}(h_N)$ is a function depending on $\mathcal{M}$ and $\|p\|_{\mathcal{C}^3}$. For any $k \in \mathbb{N}$, the norm $\|\cdot\|_{\mathcal{C}^k(\mathcal{M})}$ is given by

$$\|f\|_{\mathcal{C}^k(\mathcal{M})} := \sup_{|\alpha|\leq k, x\in\mathcal{M}} |\partial_\alpha f(x)|$$

where α denotes a multi-index.

We first note that by (3.2.6)

$$\begin{aligned} Ch_N \cdot \rho_N &= \sup\{Ch_N(u_s - u_t) : \frac{1}{2}\sum_i (u_i - u_k)^2 p_{ki} \leq 1, 1 \leq k \leq N\} \\ &= \sup\{u_s - u_t : \frac{1}{2}\sum_i (u_i - u_k)^2 p_{ki} \leq \mathcal{C}^2 h_N^2, 1 \leq k \leq N\} \end{aligned} \tag{3.3.1}$$

The proof of Proposition 3.3.1 centers around estimating the limiting value of the left hand side of the constraint in (3.3.1) as $N \to \infty$.

If we let $P_N = \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$ be the empirical distribution on $\mathcal{D}_N$, then we can rewrite the left hand side of the constraints in the problem 3.3.1 in the form:

$$\begin{aligned}\frac{1}{2}\sum_i (u_i - u_k)^2 p_{ki} &= \frac{1}{2}\left(\frac{1}{N}\sum_i (u_i-u_k)^2 K_h(x_k,x_i)\right)\left(\frac{1}{\frac{1}{N}\sum_j K_h(x_k,x_j)}\right)\\ &= \frac{1}{2}P_N\left((f(\cdot)-u_k)^2 K_h(x_k,\cdot)\right)\frac{1}{P_N(K_h(x_k,\cdot))}\end{aligned}$$

where we let $Q(\phi) := \mathbb{E}_Q[\phi]$ for any measure Q and any function ϕ integrable with respect to this measure. The limiting behavior of this constraint is established in the following lemma.

Lemma 3.3.2. *Suppose $f \in \mathcal{C}^\infty(\mathcal{M})$ and let $u_N := f|_{\mathcal{D}_N} \in \mathbb{R}^N$ denote the restriction of f to the data points $\mathcal{D}_N$. Suppose $h_N \downarrow 0$ as $N \to \infty$, and $\lim_{N\to\infty} h_N^{m+4}\frac{N}{\log N} \to \infty$. Then we have almost surely for all $k = 1, \ldots, N$,*

$$\begin{aligned}\frac{1}{2}\sum_i (u_i - u_k)^2 p_{ki} &= \frac{1}{2}P_N\left((f(\cdot)-u_k)^2 K_{h_N}(x_k,\cdot)\right)\frac{1}{P_N(K_{h_N}(x_k,\cdot))}\\ &= \left(\frac{M_2}{2M_0}|\nabla f(x_k)|^2\right)h_N^2 + \mathcal{O}(h_N^3)\end{aligned} \tag{3.3.2}$$

where $\mathcal{O}(h_N^3)$ is a function of $\|f\|_{\mathcal{C}^3}$, $\|p\|_{\mathcal{C}^3}$ and $\mathcal{M}$, and a bounded function in x_k.

To prove Lemma 3.3.2, we make use of Proposition 2 of [HAvL05], which gives estimates for integrals of functions against the kernel K_h with respect to the measure P. For completeness, we reproduce the proposition here.

Proposition 3.3.3. *if $\mathcal{M}$ satisfies Assumption 3.2.3 and K satisfies Assumption 3.2.4, and if P satisfies Assumption (2) of 3.2.2, then we have for any $x \in \mathcal{M}\backslash\partial\mathcal{M}$ there exists an $h_0(x) > 0$ for any $g \in \mathcal{C}^3(\mathcal{M})$ such that for all $h < h_0(x)$,*

$$\begin{aligned}&\frac{1}{h^m}P\left(K_h(x,\cdot)g(\cdot)\right)\\ &= M_0 p(x)g(x) + \frac{h^2}{4}M_2\left(p(x)g(x)\left[-R + \frac{1}{2}\|\sum_a \Pi(\partial_a,\partial_a)\|^2_{T_x\mathbb{R}^n}\right] + 2(\Delta(pg))(x)\right) + \mathcal{O}(h^3).\end{aligned}$$

where R is the scalar curvature, $M_0 := \int_{\mathbb{R}^m} K(\|y\|^2)dy < \infty$ and $M_2 := \int_{\mathbb{R}^m} K(\|y\|^2)y_1^2 dy < \infty$, and $\mathcal{O}(h^3)$ is a function depending on x, $\|g\|_{\mathcal{C}^3}$, and $\|p\|_{\mathcal{C}^3}$.

Remark 3.3.4. When $\mathcal{M}$ is compact and without boundary, as in our situation, the estimates in Proposition 3.3.3 are uniform in x, so that we can remove the x dependence of $\mathcal{O}(h^3)$. For example see Lemma 8 of [CL06] which states similar result to Proposition 3.3.3, but with estimates uniform in x for all x bounded away from $\partial\mathcal{M}$, which is empty in our setting.

Proof. (Lemma 3.3.2) The proof is done in two steps. First, we show that the quantity

$$\frac{1}{2}P_N\left((f(\cdot)-u_k)^2K_{h_N}(x_k,\cdot)\right)\frac{1}{P_N(K_{h_N}(x_k,\cdot))}$$

can be estimated with the quantity

$$\frac{1}{2}P\left((f(\cdot)-u_k)^2K_{h_N}(x_k,\cdot)\right)\frac{1}{P(K_{h_N}(x_k,\cdot))}$$

with error $\mathcal{O}\left(h_N^3\right)$. Next, we apply Proposition 3.3.3 to show that this quantity in turn estimates $\left(\frac{M_2}{2M_0}|\nabla f(x_k)|^2\right)h_N^2$ with error that is also $\mathcal{O}\left(h_N^3\right)$. To simplify notation throughout this proof, we will leave the depedence of h_N on N implicit and instead write just h. We now investigate the convergence of the integrals with respect to the empirical distribution P_N. In particular, we need to show

$$\lim_{N\to\infty}\sup_{k\in[N]}\frac{1}{h^3}\left|\frac{P\left((f(\cdot)-f(x_k))^2K_h(x_k,\cdot)\right)}{P(K_h(x_k,\cdot))}-\frac{P_N\left((f(\cdot)-f(x_k))^2K_h(x_k,\cdot)\right)}{P_N(K_h(x_k,\cdot))}\right|=0\quad\text{a.s.} \tag{3.3.3}$$

where the convergence is uniform for $f\in\mathcal{C}^\infty(\mathcal{M})$ such that $|\nabla f|^2\le 1$. By the triangle inequality, for any $x\in\mathcal{M}$ we have the bounds

$$\begin{aligned}
&\frac{1}{h^3}\left|\frac{P\left((f(\cdot)-f(x_k))^2K_h(x_k,\cdot)\right)}{P(K_h(x_k,\cdot))}-\frac{P_N\left((f(\cdot)-f(x_k))^2K_h(x_k,\cdot)\right)}{P_N(K_h(x_k,\cdot))}\right|\\
&\le\left|\frac{1}{\frac{1}{h^m}P(K_h(x,\cdot))}\right|\cdot\frac{1}{h^{m+3}}\left|P_N\left((f(\cdot)-f(x))^2K_h(x,\cdot)\right)-P\left((f(\cdot)-f(x))^2K_h(x,\cdot)\right)\right|\\
&+\left|\frac{1}{h^m}P_N\left((f(\cdot)-f(x))^2K_h(x,\cdot)\right)\right|\cdot\frac{1}{h^3}\left|\frac{1}{\frac{1}{h^m}P_N(K_h(x,\cdot))}-\frac{1}{\frac{1}{h^m}P(K_h(x,\cdot))}\right|.
\end{aligned} \tag{3.3.4}$$

The quantity

$$\frac{1}{h^{m+3}}\left(P_N\left((f(\cdot)-f(x))^2K_h(x,\cdot)\right)-P\left((f(\cdot)-f(x))^2K_h(x,\cdot)\right)\right)$$

we can rewrite as $\frac{1}{h^{m+3}}\sum_{j=1}^N Z_j$, where

$$Z_j(x):=\frac{1}{N}\left((f(x_j)-f(x))^2K_h(x,x_j)-P\left((f(\cdot)-f(x))^2K_h(x,\cdot)\right)\right)$$

are i.i.d. centered random variables.

Then we have the following bounds for all $j=1,\dots,N$ and $x\in\mathcal{M}$

$$|Z_j(x)|\le\frac{8}{N}|K_{\max}||C_{\mathcal{M}}|^2:=\frac{C_1}{N}.$$

where $K_{\max}$ and $C_{\mathcal{M}}$ are constants such that $K(\cdot) \leq K_{\max} < \infty$ on $\mathcal{M}$, and $f(\cdot) \leq C_{\mathcal{M}} < \infty$ for all $f \in \mathcal{C}^\infty(\mathcal{M})$ with $|\nabla f|^2 \leq 1$ and $f(x_i) = 0$ for some $x_i \in \mathcal{D}$. By the compactness of $\mathcal{M}$, it is possible to choose $C_{\mathcal{M}}$ to be finite.

Furthermore, the proof of Theorem 3 from [THJ11] shows that the first and second moments of K^h are $\mathcal{O}\left(h^{m+2}\right)$, so we can bound $P[Z_j(x)^2]$ by

$$P[Z_j(x)^2] := \mathrm{Var}(\frac{1}{N} K_h(x, x_j)) \leq C_2 \frac{h^{m+2}}{N^2}$$

for some constant $C_2 < \infty$.

Then an application of Bernstein's inequality shows for all $x \in \mathcal{M}$ and $\epsilon > 0$,

$$P\left\{\frac{1}{h^{m+3}}\left|\sum_{j=1}^N Z_j(x)\right| > \epsilon\right\} \leq 2\exp\left(-\frac{3h^{m+4}\epsilon^2}{6\sum_{j=1}^N P[Z_j(x)^2] + 2\frac{C_1}{N}h^{m+3}\epsilon}\right)$$
$$\leq 2\exp\left(-\frac{3Nh^{m+4}\epsilon^2}{6C_2h^{m+2} + 2C_1h^{m+3}\epsilon}\right).$$

Since this holds for all x, by a union bound we have that for arbitrary collection of N points $y_k \in \mathcal{M}$ $(k = 1, \ldots, N)$

$$P\left\{\sup_{k \leq N}\left|\frac{1}{h^{m+3}}\sum_{j=1}^N Z_j(y_k)\right| > \epsilon\right\} \leq 2N\exp\left(-\frac{3Nh^{m+4}\epsilon^2}{6C_2h^{m+2} + 2C_1h^{m+3}\epsilon}\right). \tag{3.3.5}$$

A sufficient condition for the left hand side of (3.3.5) to be summable in N is that

$$\lim_{N\to\infty} h_N^{m+4}\frac{N}{\log N} \to \infty. \tag{3.3.6}$$

If we replace the points y_k above with the random variables x_k for $k = 1, \ldots, N$, then by the Borel-Cantelli lemma

$$\lim_{N\to\infty} \frac{1}{h^{m+3}}\left|P_N\left((f(\cdot) - f(x_k))^2 K_h(x_k, \cdot)\right) - P\left((f(\cdot) - f(x_k))^2 K_h(x_k, \cdot)\right)\right| = 0 \quad \text{a.s.} \tag{3.3.7}$$

uniformly in k and in f. To estimate the second term of (3.3.4), we again apply the bounds $\left|\frac{1}{a+t} - \frac{1}{a}\right| \leq C|t|$ with $a = \frac{1}{h^m}P(K_h(x, \cdot)) = M_0 p(x) + \mathcal{O}(h^2) \geq M_0 p_{\min} + \mathcal{O}(h^2) > 0$ for h small enough, and $t = P_N(K_h(x_k, \cdot)) - P(K_h(x_k, \cdot))$, we see that

$$\frac{1}{h^3}\left|\frac{1}{\frac{1}{h^m}P_N(K_h(x, \cdot))} - \frac{1}{\frac{1}{h^m}P(K_h(x, \cdot))}\right| = \frac{1}{h^3}\mathcal{O}\left(\left|\frac{1}{h^m}P_N(K_h(x_k, \cdot)) - \frac{1}{h^m}P(K_h(x_k, \cdot))\right|\right).$$

By an argument identical to the one above, we can show

$$\lim_{N\to\infty}\sup_{k\in[N]}\frac{1}{h^{m+3}}\left|P_N(K_h(x_k, \cdot)) - P(K_h(x_k, \cdot))\right| = 0 \quad \text{a.s.} \tag{3.3.8}$$

since (3.3.6) holds by assumption.

Then for all $x \in \mathcal{M}$ and for all small enough $h > 0$,

$$\frac{1}{\frac{1}{h^m} P(K_h(x,\cdot))} \leq \frac{1}{M_0 p_{\min}} + \mathcal{O}(h^2) < \infty \tag{3.3.9}$$

by Proposition 3.3.3. Furthermore, we have the estimates

$$\frac{1}{h^m} P_N\left((f(\cdot) - f(x))^2 K_h(x,\cdot)\right) \leq 4M_0 p_{\max} \mathcal{C}_{\mathcal{M}}^2 < \infty. \tag{3.3.10}$$

Combining (3.3.7) with (3.3.9) shows the first term on the right hand side of (3.3.4) converges a.s. to 0. Similarly, (3.3.8) and (3.3.10) show the second term also converges a.s. to 0. This proves (3.3.3).

We now show convergence of the quantity

$$\frac{1}{2} P\left((f(\cdot) - u_k)^2 K_{h_N}(x_k,\cdot)\right) \frac{1}{P(K_{h_N}(x_k,\cdot))}.$$

For the function $g(\cdot) = (f(\cdot) - u_k)^2$, we have

$$\begin{aligned}\nabla g(\cdot) &= 2\left(f(\cdot) - u_k\right) \nabla f(\cdot), \\ \Delta g(\cdot) &= 2(f(\cdot) - u_k)\Delta f(\cdot) + 2|\nabla f(\cdot)|^2.\end{aligned}$$

In particular, $g(x_k) = 0$, $\nabla g(x_k) = 0$, and $\Delta g(x_k) = 2|\nabla f(x_k)|^2$, so applying Proposition 3.3.3 to g we have that

$$\begin{aligned}\frac{1}{h^m} P\left(g(\cdot) K_h(x_k,\cdot)\right) &= M_2 \frac{h^2}{4}\left(2(\Delta(pg))(x_k)\right) + \mathcal{O}(h^3) \\ &= M_2 \frac{h^2}{2}\left(p(x_k)(\Delta g)(x_k) + 2(\nabla p)(x_k)(\nabla g)(x_k) + (\Delta p)(x_k) g(x_k)\right) + \mathcal{O}(h^3). \\ &= h^2 M_2 p(x_k)|\nabla f(x_k)|^2 + \mathcal{O}(h^3).\end{aligned} \tag{3.3.11}$$

Applying Proposition 3.3.3 with $f = 1$:

$$\frac{1}{h^m} P(K_h(x_k,\cdot)) = M_0 p(x_k) + \mathcal{O}(h^2). \tag{3.3.12}$$

Thus, $P(K_h(x_k,\cdot)) \geq M_0 p_{\min} + \mathcal{O}(h^2) > 0$ for h small enough. The first order Taylor expansion of the function $f(x) = 1/x$ centered at $a \neq 0$ shows

$$\left|\frac{1}{a+t} - \frac{1}{a}\right| \leq C|t|$$

for t in a neighborhood of 0. If we apply this bound with $a = P(K_h(x_k, \cdot))$ and $t = \frac{1}{h^m} P(K_h(x_k, \cdot)) - M_0 p(x_k)$ it follows that

$$\frac{1}{\frac{1}{h^m} P(K_h(x_k, \cdot))} = \frac{1}{M_0 p(x_k)} + \mathcal{O}\left(\frac{1}{h^m} P(K_h(x_k, \cdot)) - M_0 p(x_k)\right) = \frac{1}{M_0 p(x_k)} + \mathcal{O}(h^2). \tag{3.3.13}$$

By multiplying (3.3.11) and (3.3.13),

$$\frac{1}{2} \frac{1}{h^m} P\left((f(\cdot) - u_k)^2 K_h(x_k, \cdot)\right) \frac{1}{\frac{1}{h^m} P(K_h(x_k, \cdot))} = \left(\frac{M_2}{2M_0} |\nabla f(x_k)|^2\right) h^2 + \mathcal{O}(h^3). \tag{3.3.14}$$

Combined with (3.3.3), this proves Lemma 3.3.2. □

We are now able to prove Proposition 3.3.1.

Proof. (Proposition 3.3.1) We first introduce some new terminology. A function $f \in \mathcal{C}^\infty(\mathcal{M})$ is said to be $d_{\mathcal{M}}$-admissible if $|\nabla f|^2 \leq 1$. in other words, f satisfies the constraints of the optimization problem

$$\sup \{u(x) - u(y) : |u \in \mathcal{C}^\infty(\mathcal{M}), |\nabla u| \leq 1\} \tag{3.3.15}$$

whose optimal value is the geodesic distance $d_{\mathcal{M}}(x, y)$. Likewise, a vector $u \in \mathbb{R}^N$ is $\tilde{\rho}_N$-admissible if u satisfies all the constraints of the primal problem (3.3.1). That is,

$$\frac{1}{2} \sum_i (u_i - u_k)^2 p_{ki} \leq \mathcal{C}^2 h_N^2$$

for $1 \leq k \leq N$. Suppose f is $d_{\mathcal{M}}$-admissible, and let $u := f|_{\mathcal{D}_N}$. Then by Lemma 3.3.2,

$$\frac{1}{2} \sum_i (u_i - u_k)^2 p_{ki} \frac{2M_0}{M_2} \frac{1}{h^2} \leq |\nabla f(x_k)|^2 + \mathcal{O}(h_N) \leq 1 + \mathcal{O}(h_N). \quad (k = 1, \ldots, N)$$

This shows that if f is $d_{\mathcal{M}}$-admissible then $\frac{u}{\sqrt{1+\mathcal{O}(h_N)}}$ is $\tilde{\rho_N}$-admissible. Since the problems (3.3.1) and are invariant under translations $f \mapsto f + c$, without loss of generality we can assume $f(x_0) = 0$ for some fixed $x_0 \in \mathcal{M}$. Now the set

$$\mathcal{S} := \{f \in \mathcal{C}^\infty(\mathcal{M}), f(x_0) = 0, |\nabla f| \leq 1\}$$

is uniformly bounded in $\mathcal{C}^3(\mathcal{M})$. Since the value of $\mathcal{O}(h_N)$ depends on f through $\|f\|_{\mathcal{C}_3}$, it is uniformly bounded over all $d_{\mathcal{M}}$-admissable f.

Now

$$\frac{u}{\sqrt{1 + \mathcal{O}(h_N)}} = u + u\mathcal{O}(h_N) = u(1 + \mathcal{O}(h_N))$$

since $u = f|_{\mathcal{D}_N}$ can be bounded uniformly for all N and for all $f \in \mathcal{S}$. So we have shown that every $d_{\mathcal{M}}$-admissable function f defines a $\tilde{\rho}_N$-admissable vector $f|_{\mathcal{D}_N}(1+\mathcal{O}(h_N))$. Furthermore, the difference between the objective function of the problem for (3.3) evaluated at f and the objective function of the problem (3.3.1) evaluated at $f|_{\mathcal{D}_N}(1+\mathcal{O}(h_N))$ is $\mathcal{O}(h_N)$. It follows that $d_{\mathcal{M}}|_{\mathcal{D}_N\times\mathcal{D}_N} \leq \tilde{\rho}_N + \mathcal{O}(h_N)$.

□

3.4 Uppper bound limits for the graph Carathéodory metric

In this section we provide some heuristic arguments which suggest why the geodesic metric may be a limiting upper bound for the sequence of metrics $\{\tilde{\rho}_N\}_N$. In the previous section we showed that $d_{\mathcal{M}}$ was a limiting lower bound for $\{\tilde{\rho}_N\}_N$ by examining the limiting behavior of the constraint for the rescaled primal problem (3.3.1). We will analyze the rescaled dual problem to try to establish upper bounds for the metric $\tilde{\rho}_N$. Consider the rescaled primal problem:

$$\begin{aligned} &\text{maximize } \sqrt{\frac{M_2}{2M_0}} h_N (e_t - e_s)^T u \\ &\text{subject to } \|A_k^N u\| \leq 1, \quad 1 \leq k \leq N. \end{aligned}$$

It is equivalent to the problem:

$$\begin{aligned} &\text{maximize } (e_t - e_s)^T u \\ &\text{subject to } \|A_k^N u\| \leq \sqrt{\frac{M_2}{2M_0}} \cdot h_N, \quad 1 \leq k \leq N. \end{aligned} \tag{3.4.1}$$

The corresponding dual problem is given by

$$\begin{aligned} &\text{minimize } \left(\sqrt{\frac{M_2}{2M_0}} h_N\right) \cdot \sum_{k=1}^{N} \|z_k\| \\ &\text{subject to } \sum_{k=1}^{N} (A_k^N)^T z_k = e_t - e_s. \end{aligned} \tag{3.4.2}$$

From the complementary slackness conditions (3.2.8), we know that if u^* is the optimal solution for the primal problem (3.4.1), then the optimal dual solution $(z_k^*)_{k=1}^N$ has the form $z_k^* = c_k A_k^N u^*$, for some constants $\{c_k\}_{k=1}^N$.

The strategy we propose for attempting to show $d_{\mathcal{M}}$ is a limiting upper bound for the sequence $\{\tilde{\rho}_N\}_N$ is the following. We estimate the primal optimal solution u^* with the function

$d_{\mathcal{M}}(x_s, \cdot)|_{\mathcal{D}_N}$, which we can identify with the vector $(d_{\mathcal{M}}(x_s, x_j))_{j=1}^N$. By the complementary slackness conditions (3.2.8), we will in turn estimate the dual optimal solution $(z_k^*)_{k=1}^N$ by $(\bar{z}_k)_{k=1}^N$, where

$$\bar{z}_k := c_k A_k^N \left(d_{\mathcal{M}}(x_s, x_j)\right)_{j=1}^N \tag{3.4.3}$$

for some constants $\{c_k\}_{k=1}^N$ which are yet to be determined. Without loss of generality, we can assume that $(\bar{z}_k)_{k=1}^N$ is a feasible point for a perturbation of (3.4.2) which has the form

$$\begin{aligned} &\text{minimize} \quad \left(\sqrt{\frac{M_2}{2M_0}} h_N\right) \cdot \sum_{k=1}^N \|z_k\| \\ &\text{subject to} \quad \sum_{k=1}^N (A_k^N)^T z_k = e_t - e_s + \nu_N \end{aligned} \tag{3.4.4}$$

where

$$\nu_N := \sum_{k=1}^N (A_k^N)^T \bar{z}_k - (e_t - e_s).$$

We want to show that as $N \to \infty$, the optimal value of (3.4.4) approaches $d_{\mathcal{M}}(x_s, x_t)$ and also that the difference between the optimal values of (3.4.4) and (3.4.2) goes to 0. That is, we need to find a set of constants $\{c_k\}_{k=1}^N$ such that

$$\lim_{N\to\infty} \left| \left(\sqrt{\frac{M_2}{2M_0}} h_N\right) \cdot \sum_{k=1}^N \|\bar{z}_k\| - d_{\mathcal{M}}(x_s, x_t) \right| = 0 \tag{3.4.5}$$

and

$$\lim_{N\to\infty} \left(\sqrt{\frac{M_2}{2M_0}} h_N\right) \left| \sum_{k=1}^N \|\bar{z}_k^*\| - \sum_{k=1}^N \|z_k^*\| \right| = 0 \tag{3.4.6}$$

where $(\bar{z}_k^*)_{k=1}^N$ is optimal for (3.4.4), and the convergence holds either almost surely or with probability going to 1 as $N \to \infty$. Furthermore, we want these bounds to be uniform in the indices $s, t \in [N]$ of the objective functions of (3.4.2) and (3.4.4). Since (3.4.4) is a minimization problem, the objective value of any point feasible for (3.4.4) is an upper bound for the optimal value. In particular, $\left(\sqrt{\frac{M_2}{2M_0}} h_N\right) \cdot \sum_{k=1}^N \|\bar{z}_k\|$ is an upper bound for $\left(\sqrt{\frac{M_2}{2M_0}} h_N\right) \cdot \sum_{k=1}^N \|\bar{z}_k^*\|$. If (3.4.6) holds, then $\left(\sqrt{\frac{M_2}{2M_0}} h_N\right) \cdot \sum_{k=1}^N \|\bar{z}_k\|$ would also be an approximate upper limit for the optimal value of the unperturbed problem (3.4.2). Combined with (3.4.5), this would show that the geodesic distance $d_{\mathcal{M}}(x_s, x_t)$ is an approximate upper bound for $\tilde{\rho}_N$.

We now investigate selecting constants c_k to make (3.4.5) hold. Let γ denote a path length minimizing geodesic from x_s to x_t. We will assume the coefficients c_k have the form:

$$c_k = \frac{1}{h_N^2} \frac{2M_0}{M_2} \cdot \mu_k \tag{3.4.7}$$

for some $\mu_k \geq 0$ which are chosen so that

$$\sum_i \mu_i = |\gamma| = d_{\mathcal{M}}(x_s, x_t) \tag{3.4.8}$$

where $|\gamma|$ denotes the arc length of γ, and the last equality holds because γ is a geodesic. For this choice of c_k, the dual objective function evaluated at $(\bar{z}_k)_{k=1}^N$ is given by:

$$\begin{aligned} \left(\sqrt{\frac{M_2}{2M_0}} h_N\right) \cdot \sum_{k=1}^N \|\bar{z}_k\| &= \left(\sqrt{\frac{2M_0}{M_2}}\right) \sum_{k=1}^N \left(\mu_k \cdot \left\| \frac{A_k^N \left(d_{\mathcal{M}}(x_s, x_j)\right)_{j=1}^N}{h_N} \right\|\right) \\ = \left(\sum_{k=1}^N \mu_k\right) \left(1 + \mathcal{O}\left(h_N\right)\right) &= d_{\mathcal{M}}(x_s, x_t) \left(1 + \mathcal{O}\left(h_N\right)\right). \end{aligned} \tag{3.4.9}$$

In the second equality, we applied Proposition 3.3.2 with $f(\cdot) = d_{\mathcal{M}}(x_s, \cdot)$ to show

$$\frac{\|A_k^N \left(d_{\mathcal{M}}(x_s, x_j)\right)_{j=1}^N \|^2}{h_N^2} = \frac{M_2}{2M_0} |\nabla d_{\mathcal{M}}(x_s, \cdot)|^2 + \mathcal{O}\left(h_N\right),$$

followed by the first order Taylor expansion of the function $f(x) = \sqrt{x}$ centered at $a \neq 0$, which shows

$$|\sqrt{a+t} - \sqrt{a}| \leq C|t|$$

for t in a neighborhood of 0. We apply this with $a = \frac{M_2}{2M_0} |\nabla d_{\mathcal{M}}(x_s, \cdot)|^2$ and $t = h_N$, which gives the estimate

$$\frac{\|A_k^N \left(d_{\mathcal{M}}(x_s, x_j)\right)_{j=1}^N \|}{h_N} = \sqrt{\frac{M_2}{2M_0}} + \mathcal{O}\left(h_N\right),$$

since $|\nabla_x d_{\mathcal{M}}(x_s, x)| = 1$ for $x \neq x_s$.

The estimates in (3.4.9) show that if we choose $\{c_k\}_{k=1}^N$ according to (3.4.7) and (3.4.8), our approximation for the dual solution, $(\bar{z})_{k=1}^N$ satisfies almost surely(3.4.5) with error $\mathcal{O}\left(h_N\right)$.

To show (3.4.6) holds, we want to show first that the perturbation $\nu_N \to 0$ in some sense as $N \to \infty$; then we must establish conditions under which this implies the optimal values of (3.4.2) and (3.4.4) converge with high probability. To this end, we will show that ν_N converges to 0 in a weak sense. That is, we will show

$$\lim_{N \to \infty} |\langle \nu_N, f|_{\mathcal{D}_N} \rangle| = 0.$$

for all f belonging to a set of test functions $\mathcal{F}$ which is yet to be determined.

We begin by rewriting the constraint for (3.4.2). When $\{z_k\}_k^N$ has the form $z_k = c_k A_k^N u$ for some $u \in \mathbb{R}^N$, the dual constraint equation for (3.4.2) becomes

$$\sum_{k=1}^{N} c_k (A_k^N)^T A_k^N u = e_t - e_s.$$

Since $A_k^N = \frac{1}{\sqrt{2}}\mathrm{diag}(\sqrt{p_{k\cdot}})(I - \tilde{1}e_k^T)$, by expanding we see that

$$\begin{aligned}\sum_k c_k (A_k^N)^T A_k^N u &= \frac{1}{2}\sum_k c_K (I - e_k\tilde{1}^T)\mathrm{diag}(p_{k\cdot})(I - \tilde{1}e_k^T)u^* \\ &= \frac{1}{2}\sum_k c_k \left[(p_{kj}u_j)_{j=1}^N - (p_{kj}u_k)_{j=1}^N - \left(\sum_{l=1}^N p_{kl}u_l\right)e_k + u_k e_k\right].\end{aligned}$$

Let $f : \mathcal{M} \mapsto \mathbb{R}$, and let $f_k = f(x_k)$. Then if we take the Euclidean inner product of the left hand side of the above equation with the restriction vector $(f_k)_{k=1}^N$, we have

$$\begin{aligned}&\frac{1}{2}\sum_j \left(\sum_k c_k (A_k^N)^T A_k^N u\right)_j f_j \\ &= \frac{1}{2}\sum_k c_k \left[\sum_j p_{kj}u_j f_j \; - \left(\sum_j p_{kj} f_j\right)u_k - \left(\sum_l p_{kl}u_l\right)f_k + f_k u_k\right] \\ &= \frac{1}{2}\sum_k c_k \left[\left(\sum_j p_{kj}u_j f_j \; - u_k f_k\right) - \left(\sum_j p_{kj} f_j - f_k\right)u_k - \left(\sum_l p_{kl}u_l^* - u_k^*\right)f_k\right] \\ &= \frac{1}{2}\sum_k c_k \left[G_N((f|_{\mathcal{D}_N}\cdot u)_{k=1}^N)(x_k) - u_k \cdot G_N(f|_{\mathcal{D}_N})(x_k) - f_k \cdot G_N(u)(x_k)\right].\end{aligned}$$

As above, $G_N = \Pi_N - I$ denotes the generator of the random walk on $\mathcal{D}_N$ with transition probability matrix Π_n, and the vector u is viewed as the function $u : \mathcal{D}_N \mapsto \mathbb{R}$ such that $u(x_k) = u_k$. Then by taking the Euclidean inner product of the right hand size of the constraint equation with $(f_k)_{k=1}^N$, we see that the constraint equation $\sum_{k=1}^N c_k(A_k^N)^T A_K^N u = e_t - e_s$ is equivalent to the weak form

$$\frac{1}{2}\sum_k c_k \left[G_N((f|_{\mathcal{D}_N}\cdot u)_{k=1}^N)(x_k) - u_k \cdot G_N(f|_{\mathcal{D}_N})(x_k) - f_k \cdot G_N(u)(x_k)\right] = f_t - f_s \quad (3.4.10)$$

for all $(f_k)_{k=1}^N \in \mathbb{R}^N$.

Let us now investigate the limiting behavior of the left hand side of this constraint when $(f_k)_{k=1}^N = f|_{\mathcal{D}_N}$ for a sufficiently smooth $f : \mathcal{M} \mapsto \mathbb{R}$. By Theorem 3 of [THJ11], if $h_N \to 0$ and $\frac{N h_N^{d+2}}{\log N} \to \infty$, then for any $\phi \in \mathcal{C}^2(\mathcal{M})$

$$\|\frac{Z_{K,d}}{h_N^2} G_N(\phi|_{\mathcal{D}_N}) - \Delta_2\phi\|_\infty \to 0 \; a.s. \quad (3.4.11)$$

Here $\Delta_2 := \frac{1}{p^2}\text{div}(p^2\text{grad})$ denotes the 2-weighted Laplacian with respect to the density p. By Theorem 1 of [HAvL05], the constant $Z_{K,d} = \frac{2M_0}{M_2}$. Then the left hand side of the weak form of the dual constraint (3.4.10) evaluated at $(\bar{z}_k)_{k=1}^N$ is given by

$$\frac{1}{2}\sum_k \mu_k \left[\Delta_2(f(\cdot)d_{\mathcal{M}}(x_s,\cdot))(x_k) - d_{\mathcal{M}}(x_s,x_k)\cdot\Delta_2 f(x_k) - f(x_k)\cdot\Delta_2(d_{\mathcal{M}}(x_s,\cdot))(x_k)) + \epsilon_\Delta\right] \tag{3.4.12}$$

where ϵ_Δ is the error depending on N, h_N and f that was incurred by applying the estimate (3.4.11).

To bound the error ϵ_Δ, we define the operator

$$\left(\Delta_{h,2}f\right)(x) := \frac{1}{h^2}\left(\frac{1}{d_h(x)}\int_{\mathcal{M}}\frac{1}{h^m}K_h(x,y)f(y)p(y)d\text{vol}_{\mathcal{M}}(y) - f(x)\right)$$

where $d_h(x) = \int_{\mathcal{M}}\frac{1}{h^m}K_h(x,y)p(y)d\text{vol}_{\mathcal{M}}(y)$ is the continuous degree operator. From the triangle inequality,

$$\epsilon_\Delta \leq \frac{2M_0}{M_2}\left|\frac{1}{h_N^2}G_N(f|_{\mathcal{D}_N}) - \Delta_{h,2}f(x)\right| + \left|\frac{2M_0}{M_2}\Delta_{h,2}f(x) - \Delta_2 f\right|.$$

These two terms are known as the variance and the bias term respectively. By Theorem 1 of [HAvL05] we can bound the bias term by

$$\left|\frac{2M_0}{M_2}\Delta_{h,2}f(x) - \Delta_2 f\right| = \mathcal{O}\left(h^2\right),$$

where $\mathcal{O}(h^2)$ is a function depending on $\|f\|_{\mathcal{C}^3}$ and $\|p\|_{\mathcal{C}^3}$. For the variance term, we have by Theorem 2 of [HAvL05] that for any $\epsilon > 0$ the probabilistic bounds

$$P\left(\left|\frac{1}{h_N^2}G_N(f|_{\mathcal{D}_N}) - \Delta_{h,2}f(x)\right| \geq \epsilon\right) \leq CNe^{\frac{-Nh^{m+4}\epsilon^2}{C}}$$

which shows almost sure convergence when $Nh^{m+4}/\text{log}N \to \infty$. If we replace ϵ with $\epsilon \cdot h$, then this shows that $\frac{1}{h}\left|\frac{1}{h_N^2}G_N(f|_{\mathcal{D}_N}) - \Delta_{h,2}f(x)\right| \to 0$ a.s. under the stronger condition that $Nh^{m+6}/\text{log}N \to \infty$. In this case, we have $\frac{1}{h_N^2}G_N(f|_{\mathcal{D}_N}) = \Delta_{h,2}f(x) + \mathcal{O}(h)$). Hence, we have that $\epsilon_\Delta = \mathcal{O}(h)$, and we can improve this bound to $\mathcal{O}(h^2)$ if we allow $Nh^{m+8}/\text{log}N \to \infty$.

that we have shown the error ϵ_Δ goes to 0 under appropriate conditions, we now want to show the sum

$$\frac{1}{2}\sum_k \mu_k \left[\Delta_2(f(\cdot)d_{\mathcal{M}}(x_s,\cdot))(x_k) - d_{\mathcal{M}}(x_s,x_k)\cdot\Delta_2 f(x_k) - f(x_k)\cdot\Delta_2(d_{\mathcal{M}}(x_s,\cdot))(x_k))\right]$$

converges to $f(x_t) - f(x_s)$. To do this, in addition to (3.4.8), we require that measure μ given by

$$\mu := \sum_{k=1}^{N} \mu_k \delta_{x_k}$$

converges weakly to the arc-length measure on γ. That is,

$$\int_{\mathcal{M}} f d\mu = \sum_{k=1}^{N} f(x_k)\mu_k \to \int_{\gamma} f d\|\gamma\|$$

as $N \to \infty$ for any f in our suitably chosen family of test functions, $\mathcal{F}$. To find a measure μ approximating the arc length measure of γ, we apply Theorem 2.7 from [BCC$^+$10], which states that for all $\frac{1}{2} < \alpha < \frac{3}{2}$, there exists a $c > 0$ with the following property: If range$(\gamma) = \gamma_1 \cup \cdots \cup \gamma_K$ is a decomposition of range(γ) into disjoint pieces with length $|\gamma_j| = \omega_j$, then there exists a distribution of points $\{q_j\}_{j=1}^K$ with $q_j \in \gamma_j$ such that

$$\left| \sum_{j=1}^{K} \omega_j f(q_j) - \int_{\mathcal{M}} f d|\gamma| \right| \le c \max_{1\le j\le K} \left\{ |\gamma_j^{\alpha}| \right\} \|f\|_{\mathbb{W}^{\alpha,2}} \tag{3.4.13}$$

where $\|\cdot\|_{\mathbb{W}^{\alpha,2}}$ denotes the standard norm on the Sobolev space $\mathbb{W}^{\alpha,2}$ of all functions in $L^2(\mathcal{M})$ with weak derivatives up to order α with finite $L^2(\mathcal{M})$ norm.

We will say the δ-sampling condition holds on the manifold $\mathcal{M}$ for the data points $\{x_j\}_{j=1}^N$ if for every point $z \in \mathcal{M}$, there exists a data point x_i for which $d_{\mathcal{M}}(z, x_i) \le \delta$. For any $\delta > 0$, when N is large, the δ-sampling condition holds with high probability since the data point density $p(x)$ is bounded away from 0 by (2) of Assumption 3.2.2 (see the Sampling Lemma in Section 4 of [BSLT00] for details). If we assume the δ-sampling condition holds for some $\delta > 0$, then for each q_i $(1 \le i \le K)$, let $x_{\gamma(i)}$ denote a data point such that $d_{\mathcal{M}}(x_{\gamma(i)}, q_i) < \delta$. Then for any Holder continuous $f : \mathcal{M} \mapsto \mathbb{R}$ we have the estimate

$$\begin{aligned} \left| \sum_{j=1}^{K} \omega_j f(x_{\gamma(j)}) - \sum_{j=1}^{K} \omega_j f(q_j)) \right| &\le \|f\|_C^{0,1} \left| \sum_{j=1}^{K} \omega_j \|x_{\gamma(j)} - q_j\| \right| \\ \le \|f\|_C^{0,1} \cdot \sum_{j=1}^{K} d_{\mathcal{M}}(x_{\gamma(j)}, q_j)\omega_j &\le \|f\|_C^{0,1} \cdot \delta \cdot |\gamma|. \end{aligned} \tag{3.4.14}$$

Here $\|x_{\gamma(j)} - q_j\|$ denotes the Euclidean distance of the extrinsic coordinates in $\mathbb{R}^n$, and we have used the fact that $\|x - y\| \le d_{\mathcal{M}}(x, y)$. If we combine (3.4.14) with (3.4.13) from

Theorem 2.7 of [BCC+10], then we have the bounds

$$\begin{aligned}&\left|\sum_{j=1}^{K}\omega_j f(x_{\gamma(j)})-\int_{\mathcal{M}} fd|\gamma|\right| \le \left|\sum_{j=1}^{K}\omega_j \left(f(x_{\gamma(j)})-f(q_j)\right)\right|+\left|\sum_{j=1}^{K}\omega_j f(q_j)-\int_{\mathcal{M}} fd|\gamma|\right|\\ &\le \|f\|_{\mathcal{C}}^{0,1}\cdot|\gamma|\cdot\delta+c\max_{1\le j\le K}\{|\gamma_j|\}\,\|f\|_{\mathbb{W}^{1,2}}.\end{aligned} \tag{3.4.15}$$

Motivated by these estimates, let us define μ_k in (3.4.7) by $\mu_{\gamma_j}=\omega_j=|\gamma_j|$ for $j=1\ldots,K$, and $\mu_k=0$ otherwise, so that

$$\mu:=\sum_k \mu_k\delta_{x_k}=\sum_{j=1}^{K}\omega_j\delta_{x_{\gamma(j)}}.$$

Then, if our family of test functions $\mathcal{F}$ is uniformly bounded in the Holder semi-norm $\|\cdot\|_{\mathcal{C}}^{0,1}$ and the Sobolev norm $\|\cdot\|_{\mathcal{W}^{1,2}}$, and if the decomposition of γ above is into pieces of length $|\gamma_j|=\mathcal{O}(\delta)$, then μ weakly approximates the arc length measure $d|\gamma|$ in the sense that $\left|\int_{\mathcal{M}} fd\mu-\int_{\mathcal{M}} fd|\gamma|\right|=\mathcal{O}(\delta)$ for all $f\in\mathcal{F}$.

Now since the measure $\mu=\sum_{k=1}^{N}\mu_k\delta_{x_k}$ approximates the arc length on γ, the above sum is an approximation for the integral

$$\frac{1}{2}\int_{\gamma}\left[\Delta_2(f\cdot d_{\mathcal{M}}(x_s,\cdot))-d_{\mathcal{M}}(x_s,\cdot)\cdot\Delta_2 f-f\cdot\Delta_2 d_{\mathcal{M}}(x_s,\cdot)\right]d|\gamma| \tag{3.4.16}$$

where by (3.4.15), the error incurred from this estimate is $\mathcal{O}(\delta)$ provided we have uniform bounds on $\Delta_2(f\cdot d_{\mathcal{M}}(x_s,\cdot))$ and $\Delta_2 f$ in the semi norm $\|\cdot\|_{\mathcal{C}}^{0,1}$ and the norm $\|\cdot\|_{\mathbb{W}^{1,2}}$ for all $f\in\mathcal{F}$. Or equivalently, $\mathcal{F}$ should be uniformly bounded in $\|\cdot\|_{\mathcal{C}}^{0,1}$ and in $\|\cdot\|_{\mathbb{W}^{3,2}}$. To simplify this integral, we will make use of the following identities (see for example page 152 of [Cha06]):

$$\begin{aligned}\operatorname{div}(f\nabla g) &= f\cdot\Delta_{\mathcal{M}}g+\langle\nabla f,\nabla g\rangle\\ \Delta_{\mathcal{M}}(f\cdot g) &= f\cdot\Delta_{\mathcal{M}}g+2\langle\nabla f,\nabla g\rangle+g\cdot\Delta_{\mathcal{M}}f\end{aligned}$$

By the first equation, it follows that

$$\Delta_s f:=\frac{1}{p^s}\operatorname{div}(p^s\nabla f)=\Delta_{\mathcal{M}}f+\frac{1}{p^s}\langle\nabla p^s,\nabla f\rangle=\Delta_{\mathcal{M}}f+\frac{s}{p}\langle\nabla p,\nabla f\rangle.$$

Hence,

$$\begin{aligned}&\Delta_s(f\cdot g)\\ &=\Delta_{\mathcal{M}}(fg)+\frac{s}{p}\langle\nabla p,\nabla(fg)\rangle\\ &=g\Delta_{\mathcal{M}}f+f\Delta_{\mathcal{M}}g+2\langle\nabla f,\nabla g\rangle+g\frac{s\nabla p}{p}\cdot\nabla f+f\langle\frac{s\nabla p}{p},\nabla g\rangle\\ &=g\Delta_s f+f\Delta_s g+2\langle\nabla f,\nabla g\rangle.\end{aligned}$$

Applying this identity to the integral (3.4.16) shows it is equal to

$$\int_\gamma \langle \nabla f, \nabla d_{\mathcal{M}}(x_s, \cdot)\rangle d|\gamma|. \tag{3.4.17}$$

Without loss of generality, let us assume that γ is parameterized by arc length so that $\gamma(0) = x_s$, and $\gamma(d_{\mathcal{M}}(x_s, x_t)) = x_t$. To further simplify the integral (3.4.17), we will show that $\nabla d_{\mathcal{M}}(x_s, \gamma(l)) = \gamma'(l)$ by an application of Gauss' lemma. In order to prove this fact, we will need to make use of Theorems 14 and 15 in chapter 9 of [Spi79], which we reproduce here.

Theorem 3.4.1. *For every $p \in \mathcal{M}$ there is a neighborhood W and a number $\epsilon > 0$ such that*

1. *Any two points of W are joined by a unique geodesic in $\mathcal{M}$ of length $< \epsilon$.*
2. *Let $v(q, q')$ denote the unique vector $v \in T_q\mathcal{M}$ of length $< \epsilon$ such that $exp_q(v) = q'$. Then $(q, q') \mapsto v(q, q')$ is a $\mathcal{C}^\infty$ function from $W \times W \mapsto T\mathcal{M}$.*
3. *for each $q \in W$, the map exp_q maps the open ϵ-ball in M_q diffeomorphically onto an open set $U_q \supset W$.*

and

Theorem 3.4.2. *(Gauss' Lemma). In U_q, the geodesics through q are perpendicular to the hypersurfaces*

$$\{exp_q(v) : \|v\| = constant < \epsilon\}$$

Now let $\gamma(l), 0 \leq l \leq d_{\mathcal{M}}(x_s, x_t)$ denote an arbitrary point on the curve γ. Let W denote a neighborhood of $\gamma(l)$ satisfying the conditions of Theorem 3.4.1 for some $\epsilon > 0$. Now choose $l^- < l$ sufficiently close to l so that $\gamma(l^-) \in W$. By (1) of Theorem 3.4.1, $\gamma(l)$ and $\gamma(l^-)$ are connected by a unique geodesic. Clearly, this geodesic must be the segment of γ connecting $\gamma(l)$ and $\gamma(l^-)$. If we let $v\left(\gamma(l^-), \gamma(l)\right)$ be as in (2) of Theorem 3.4.1, then by Gauss' lemma, at the point $\gamma(l)$, the curve γ is perpendicular to the hypersurface

$$\begin{aligned}\{\exp_{\gamma(l^-)}(v) : \|v\| = \|v\left(\gamma(l^-), \gamma(l)\right)\|\} &= \{p \in \mathcal{M} : d_{\mathcal{M}}(\gamma(l^-), p) = d_{\mathcal{M}}(\gamma(l^-), \gamma(l))\} \\ &:= S_{\gamma(l^-)}(d_{\mathcal{M}}(\gamma(l^-), \gamma(l)), d_{\mathcal{M}}),\end{aligned}$$

where we let $S_p(r, d)$denote the sphere in $\mathcal{M}$ centered at p with radius r under the metric d. In other words, $\gamma'(l)$ is orthogonal to the level set of $d_{\mathcal{M}}(\gamma(l^-), \cdot)$ passing through $\gamma(l)$. Since this level set is a hypersurface, its orthogonal compliment has dimension one, which implies that $\gamma'(l)$ is parallel to $\nabla d_{\mathcal{M}}(\gamma(l^-), \cdot)|_{\gamma(l)}$, which is also orthogonal to the level set of $d_{\mathcal{M}}(\gamma(l^-), \cdot)$ passing through $\gamma(l)$. From this fact we can then reach our desired conclusion that $\gamma'(l)$ is parallel to $\nabla d_{\mathcal{M}}(x_s, \cdot)|_{\gamma(l)}$ by proving that the spheres $S_{x_s} := S_{x_s}(d_{\mathcal{M}}(x_s, \gamma(l); d_{\mathcal{M}})$ and

$S_{\gamma(l^-)} := S_{\gamma(l^-)}(d_{\mathcal{M}}(\gamma(l^-), \gamma(l)), d_{\mathcal{M}})$, which both pass through $\gamma(l^-)$, have the same tangent space at $\gamma(l^-)$.

Now suppose we fix a tangent vector $v \in T_{\gamma(I)}S_{x_s}$, and let $c : (-\epsilon, \epsilon) \mapsto S_{\gamma(l^-)}$ be a curve with $c'(0) = v$, and thus $c(0) = \gamma(l)$. Then $c'(0) \in T_{\gamma(l)}S_{\gamma(l^-)}$ if $d_{\mathcal{M}}(\gamma(l^-), c(\tau)) = d_{\mathcal{M}}(\gamma(l^-), c(0))$ for all τ in some neighborhood of 0. Given an orthonormal basis $\{e_1, \dots, e_m\}$ of $T_{\gamma(l^-)}\mathcal{M}$, by (3) of Theorem 3.4.1, we can define a chart $\mathbf{n} : W \mapsto \mathbb{R}^d$ by $\mathbf{n}^i(q) = \langle (exp|_W)^{-1}q, e_i\rangle$ which establishes Riemannian normal coordinates around $\gamma(l^-)$ (See page 90 of [Cha06]). From $\mathbf{n}$, we can naturally define a polar coordinate chart on W by $\phi = (r, \theta)$, where r is the radial parameter and $\theta = (\theta_1, \dots, \theta_{m-1})$ is a parameterization of the unit $(m-1)$-dimensional sphere. If we let $\phi \circ c(\tau) = (r(\tau), \theta(\tau))$ denote the representation of the curve in spherical coordinates, then an equivalent condition for $\mathcal{C}'(0)$ to be in $T_{\gamma(l)}S_{\gamma(l^-)}$ is $r'(0) = 0$. By the triangle inequality, we have

$$\begin{aligned} d_{\mathcal{M}}(x_s, \gamma(l)) &= d_{\mathcal{M}}(x_s, c(\tau)) \leq d_{\mathcal{M}}(x_s, \gamma(l^-)) + d_{\mathcal{M}}(\gamma(l^-), \gamma(l)) \\ &= d_{\mathcal{M}}(x_s, \gamma(l^-)) + r(\tau) \quad \forall \tau \in (-\epsilon, \epsilon). \end{aligned}$$

Also,

$$d_{\mathcal{M}}(x_s, r(0)) = d_{\mathcal{M}}(x_s, \gamma(l)) = d_{\mathcal{M}}(x_s, \gamma(l^-)) + d_{\mathcal{M}}(\gamma(l^-), \gamma(l)) = d_{\mathcal{M}}(x_s, \gamma(l^-) + r(0)$$

where the second equality holds because $\gamma(l^-)$ lies on the geodesic from x_s to $\gamma(l)$. From these two equations, it follows that r attains a local min at $\tau = 0$, and hence $r'(0) = 0$ as desired. Also, $v \in T_{\gamma(l)}S_{\gamma(l^-)}$, which implies $T_{\gamma(l)}S_{x_s)} \subseteq T_{\gamma(l)}S_{\gamma(l^-)}$. Furthermore, since both vector spaces have dimension m, we actually have that $T_{\gamma(l)}S_{x_s)} = T_{\gamma(l)}S_{\gamma(l^-)}$, which leads to our conclusion that $\gamma'(l)$ is parallel to $\nabla d_{\mathcal{M}}(x_s, \cdot)|_{\gamma(l)}$ because we showed above that it is parallel to $\nabla d_{\mathcal{M}}(\gamma(l^-, \cdot)|_{\gamma(l)}$. So there exists a real valued function κ such that $\nabla d_{\mathcal{M}}(x_s, \cdot)|_{\gamma(l)} = \kappa(l)\gamma'(l)$ for all $l \in [0, d_{\mathcal{M}}(x_s, x_t)]$. Now, since γ is parameterized by arc length, it follows that

$$1 = \frac{\partial}{\partial l} d_{\mathcal{M}}(x_s, \gamma(l)) = \langle \nabla d_{\mathcal{M}}(x_s, \cdot)|_{\gamma(l)}, \gamma'(l)\rangle = \langle \kappa(l)\gamma'(l), \gamma'(l)\rangle = \kappa(l)\|\gamma'(l)\|^2 = \kappa(l).$$

Thus, $\nabla d_{\mathcal{M}}(x_s, \cdot)|_{\gamma(l)} = \gamma'(l)$. Finally, we can use this fact to compute the integral (3.4.17)

$$\begin{aligned} &\int_\gamma \langle \nabla f, \nabla d_{\mathcal{M}}(x_s, \cdot)\rangle d|\gamma| \\ &= \int_0^{d_{\mathcal{M}}(x_s, x_t)} \langle \nabla f(\gamma(l)), \gamma'(l)\rangle \|\gamma'(l)\| dl = \int_0^{d_{\mathcal{M}}(x_s, x_t)} \frac{\partial}{\partial l} f\left(\gamma(l)\right) dl \\ &= f(\gamma(d_{\mathcal{M}}(x_s, x_t))) - f(\gamma(0)) = f(x_t) - f(x_s). \end{aligned}$$

By the definition of ν_N in (3.4.4), it follows that $|\langle \nu_N, f\rangle| \to 0$ for any $f \in \mathcal{C}^2(\mathcal{M})$. To attain bounds uniform for $f \in \mathcal{F}$, we need $\mathcal{F}$ to be uniformly bounded in $\|\cdot\|_{\mathcal{C}}^{0,1}$ and $\|\cdot\|_{\mathbb{W}^{3/2}}$. Heuristically, this shows that the perturbation ν_N is in some sense small for large N. It remains to show rigorously that these estimates can be used to show (3.4.6) holds. This question is open for further study.

Bibliography

[BCC+10] Luca Brandolini, Christine Choirat, Leondardo Colzani, Giacomo Gigante, Raffaello Seri, and Giancarlo Travaglini, *Quadrature rules and distribution of points on manifolds* (2010). ↑3.4, 3.4

[Bil99] P. Billingsley, *Convergence of probability measures, joh wiley & sons*, INC, New York (1999). ↑2.3, 2.4

[BN06] F. Baudoin and D. Nualart, *Notes on the two-dimensional fractional Brownian motion*, Ann. Probab **34** (2006), no. 1, 159–180. ↑1.1

[BPL76] A. Bensoussan, G. C. Papanicolaou, and J.-L. Lions, *Boundary layers and homogenization of transport processes*, University of Utah, 1976. ↑2.1

[BR14] K. Burdzy and D. Rizzolo, *A Lorentz model with variable density in a gravitational field: limit theorems and recurrence*, ArXiv e-prints (2014). Available at `http://arxiv.org/abs/1310.7312`. ↑(document), 2.1, 2.1.2, 2.2, 2.4.5

[BSLT00] Mira Bernstein, Vin De Silva, John C. Langford, and Joshua B. Tenenbaum, *Graph approximations to geodesics on embedded manifolds*, 2000. ↑3.4

[CD09] N. Chernov and D. Dolgopyat, *The galton board: limit theorems and recurrence*, Journal of the American Mathematical Society **22** (2009), no. 3, 821–858. ↑2.1

[Cha06] I. Chavel, *Riemannian geometry: A modern introduction*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2006. ↑3.4, 3.4

[Chy08] O. Chybiryakov, *Skew-product representations of multidimensional Dunkl Markov processes*, Ann. Inst. Henri Poincaré Probab. Stat. **44** (2008), 593–611. ↑1.1

[CK06] C. Costantini and T. G. Kurtz, *Diffusion approximation for transport processes with general reflection boundary conditions*, Mathematical Models and Methods in Applied Sciences **16** (2006), no. 05, 717–762. ↑2.1

[CL06] Ronald R. Coifman and Stéphane Lafon, *Diffusion maps*, Appl. Comput. Harmon. Anal. **21** (2006), no. 1, 5–30. MR2238665 (2008a:60210) ↑3.3.4

[Cos91] C. Costantini, *Diffusion approximation for a class of transport processes with physical reflection boundary conditions*, The Annals of Probability (1991), 1071–1101. ↑2.1, 2.1.2

[Dav93] E. B. Davies, *Analysis on graphs and noncommutative geometry*, J. Funct. Anal. **111** (1993), no. 2, 398–430. MR1203460 (93m:58110) ↑(document), 3.1

[EHW14] S. N. Evans, A. Hening, and E. S. Wayman, *A skew-product decomposition counterexample*, ArXiv e-prints (August 2014), available at `1408.6507`. ↑(document)

[ELJL10] K. D. Elworthy, Y. Le Jan, and X.-M. Li, *The geometry of filtering*, Birkhäuser Verlag, Basel, 2010. ↑1.1

[EM91] A. M. Etheridge and P. March, *A note on superprocesses*, Probab. Theory Related Fields **89** (1991), no. 2, 141–147. ↑1.1

[Gal63] A. R. Galmarino, *Representation of an isotropic diffusion as a skew product*, Probab. Theory Related Fields **1** (1963), no. 4, 359–378. ↑1.1

[HAvL05] Matthias Hein, Jean-Yves Audibert, and Ulrike von Luxburg, *From graphs to manifolds—weak and strong pointwise consistency of graph Laplacians*, Learning theory, 2005, pp. 470–485. MR2203281 (2006i:68108) ↑3.2.3, 3.2.1, 3.2.1, 3.3, 3.4, 3.4

[Her74] R. Hersh, *Random evolutions: A survey of results and problems*, Rocky Mountain J. Math **4** (1974), no. 3. ↑2.1

[Hir00] S. Hiraba, *Jump-type Fleming-Viot processes*, Adv. in Appl. Probab. **32** (2000), no. 1, 140–158. ↑1.1

[HJ13] R. A. Horn and C. R. Johnson, *Matrix analysis*, Second, Cambridge University Press, Cambridge, 2013. ↑1.2

[HRW14] A. Hening, D. Rizzolo, and E. S. Wayman, *A Lorentz model with variable density in a conservative force field* (2014). preprint. ↑(document)

[JS03] J. Jacod and A. N. Shiryaev, *Limit theorems for stochastic processes*, Second, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 288, Springer-Verlag, Berlin, 2003. MR1943877 (2003j:60001) ↑2.3, 2.3.3

[Kal02] O. Kallenberg, *Foundations of modern probability*, Springer-Verlag, New York, 2002. ↑1.1, 1.2

[Øks03] B. Øksendal, *Stochastic differential equations*, Springer, 2003. ↑2.3, 2.4

[KT81] S. Karlin and H. M. Taylor, *A second course in stochastic processes.*, New York etc.: Academic Press, A Subsidiary of Harcourt Brace Jovanovich, Publishers. XVIII, 542 p. $ 35.00, 1981. ↑2.5.1

[LCO09] J.-A. Lázaro-Camí and J.-P. Ortega, *Reduction, reconstruction, and skew-product decomposition of symmetric stochastic differential equations*, Stoch. Dyn. **9** (2009), no. 01, 1–46. ↑1.1

[Lia09] M. Liao, *A decomposition of Markov processes via group actions*, J. Theoret. Probab. **22** (2009), no. 1, 164–185. ↑(document), 1.1, 1.2, 4, 1.2, 2, 3, 4, 5, 1.2, 1.2.3

[LVBL98] Miguel Sousa Lobo, Lieven Vandenberghe, Stephen Boyd, and Hervé Lebret, *Applications of second-order cone programming*, Linear Algebra Appl. **284** (1998), no. 1-3, 193–228. ILAS Symposium on Fast Algorithms for Control, Signals and Image Processing (Winnipeg, MB, 1997). MR1655138 (99k:90103) ↑3.2, 3.2, 3.2

[Pap75] G. C. Papanicolaou, *Asymptotic analysis of transport processes*, Bulletin of the American Mathematical Society **81** (1975), no. 2, 330–392. ↑2.1

[Per92] E. A. Perkins, *Conditional Dawson-Watanabe processes and Fleming-Viot processes*, Seminar on Stochastic Processes, 1991 (Los Angeles, CA, 1991), 1992, pp. 143–156. ↑1.1

[PR88] E. J. Pauwels and L. C. G. Rogers, *Skew-product decompositions of Brownian motions*, Contemp. Math. **73** (1988), 237–262. ↑1.1

[RT99] K. Ravishankar and L. T., *Diffusive limit of the Lorentz model with a uniform field starting from the Markov approximation*, Markov Processes and Related Fields **5** (1999), no. 4, 385–421. ↑2.1

[RW00] L. C. G. Rogers and D. Williams, *Diffusions, Markov processes, and martingales. Vol. 2: Itô calculus*, Vol. 2, Cambridge University Press, Cambridge, 2000. ↑1.2

[Spi79] Michael Spivak, *A comprehensive introduction to differential geometry. Vol. I*, Second, Publish or Perish Inc., Wilmington, Del., 1979. MR532830 (82g:53003a) ↑3.4

[Stu94] Karl-Theodor Sturm, *Analysis on local Dirichlet spaces. I. Recurrence, conservativeness and L^p-Liouville properties*, J. Reine Angew. Math. **456** (1994), 173–196. MR1301456 (95i:31003) ↑3.1, 3.2

[Tay92] J. C. Taylor, *Skew products, regular conditional probabilities and stochastic differential equations: a technical remark*, Séminaire de probabilités xxvi, 1992, pp. 113–126. ↑1.1

[THJ11] D. Ting, L. Huang, and M. Jordan, *An Analysis of the Convergence of Graph Laplacians*, ArXiv e-prints (January 2011), available at `1101.5435`. ↑3.1, 3.2.1, 3.3, 3.4