Comparison Properties of Diffusion Semigroups on Spaces with Lower Curvature Bounds

Dissertation
zur Erlangung des Doktorgrades (Dr. rer. nat.)
der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von Max-K. v. Renesse
aus Bochum

Bonn, Dezember 2001

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Referent: Prof. Dr. Karl-Theodor Sturm
2. Referent: Prof. Dr. Werner Ballmann

Tag der Promotion:

Contents

1 Introduction

The interaction between geometry and probability for stochastic processes on Riemannian manifolds has been extensively studied since the 1950's when Itô himself initiated the theory of diffusion processes on manifolds*. Due to the close relationship between diffusion processes, operator semigroups and their generators it is often possible and convenient to switch between the analytical and stochastic picture in order to develop a better understanding of curvature effects in diffusion semigroup theory.

In the study of stochastic processes on manifolds many authors prefer to employ the language of stochastic differential equations including such important tools as the geometric Itô formula and which has lead to a variety of beautiful results. However, usually the requirements on the smoothness of the data are quite rigid and therefore SDE do not seem to be the appropriate framework for stochastic analysis in particular on such non-smooth geometries which have gained a lot of interest in recent years. On the other hand, since the natural category of state spaces for Markov processes is formed by the measurable, measured or topological spaces, it seems fairly plausible to use the equivalence of many stochastic and analytical notions on smooth spaces as a basis for the definition and investigation of analytical objects even on spaces with low regularity by 'doing analysis in terms of Markov processes'. These remarks motivate the *intrinsic* point of view that is taken in the present work, where certain classical results on the heat semigroup on manifolds with lower curvature bounds are reproved or extended to either nonsmooth Riemannian (M, g) or measured metric spaces (X, d, m) by using mainly the properties of d and m and as little other 'extrinsic' structure for X as possible. In particular we refrain completely from the usage of SDE concepts whenever we deal with stochastic processes.

A first example of this approach is given in section two by a new version of Kendall's probabilistic proof of Yau's gradient estimates for harmonic functions on smooth Riemannian manifolds with lower (Ricci) curvature bounds. We obtain here the crucial coupling process and coupling probability estimate from a direct by-hand-argumentation which involves a central limit theorem for coupled geodesic random walks and an asymptotic quadruple inequality, bypassing the sophisticated and restricted machinery of horizontal lifts of processes to the principal frame bundle etc. Except its simplicity our method has another major advantage to the SDE approach because it yields the coupling time estimate irrespective if the manifold has a nonempty cut locus or not. Additionally our proof readily suggests to be extended to more general and non-smooth spaces which is illustrated by the example of certain Riemannian polyhedra.

*Detailed expositions of the basic theory of stochastic analysis on manifolds can be found in the monographs [Éme89, IW89, HT94].

The subsequent sections deal with comparison properties of the heat semigroup on Alexandrov spaces. These are geodesic metric spaces (X,d) whose lower curvature bound is formulated in terms of the convexity of geodesic triangles and which can be considered as a straightforward generalization of Riemannian manifolds with lower sectional curvature bounds because many examples appear most naturally as Gromov-Hausdorff limits of sequences of manifolds with uniform lower sectional curvature bound. In section three we give a short review of the most important geometric concepts and regularity properties of Alexandrov spaces.

In section four we present a purely analytic approach to the extension of Yau's gradient estimate for harmonic functions onto two-dimensional Alexandrov spaces. Since Chen and Hsu have shown that distributional lower curvature bounds are sufficient for gradient estimates to hold we present an argument how to derive a distributional inequality for the Gaussian curvature from Alexandrov's geometric curvature bound, the main tools being an integral version of the Gauss-Bonnet theorem and the lower bound for the excess measure on (X,d). A similar problem was treated in a different way by I. Nikolaev in the case of metric spaces with two-sided curvature bounds and our discussion in this section is inspired by his works. However, since we impose a priori a certain regularity on the surfaces (X,d) it is not clear to which extent our results can be applied in more general situations. It is worth mentioning in this context that our analytic proof is based upon triangle comparison, whereas the stochastic proof of section two relies on an asymptotic quadruple comparison property of the underlying space and that both conditions are not equivalent in the lower curvature bound case.

The topic of section five is an investigation of the heat kernel and what we call canonical diffusion process on Alexandrov spaces (X,d). Starting from Sturm's construction of a diffusion process on measured metric spaces we show that a number of classical results about Brownian motion and the heat semigroup carry over to Alexandrov spaces, such as the lower bound for the heat kernel by Cheeger and Yau, the short time asymptotic formula of Varadhan and the first eigenvalue estimate for the Laplacian by Cheng. Finally we give a lower estimate for the escape speed of Brownian motion, generalizing the well known result by Debiard, Gaveau and Mazet. Concerning the regularity of the space (X,d) we are very general in this section, because except Alexandrov's curvature bound we only require an asymptotic growth condition for the volume of small geodesic balls in (X,d) which is satisfied even in very irregular Alexandrov spaces.

As a final introductory remark we would like to recall that all comparison properties which we have dealt with so far are usually obtained from lower Ricci curvature bounds as long as the spaces are sufficiently smooth. On the other hand it is well justified to think of

Alexandrov spaces as generalized Riemannian manifolds with lower sectional curvature bounds and hence one might expect another class of metric spaces to be a suitable substitute for Riemannian manifolds with lower Ricci curvature bounds. Unfortunately, the currently most promising candidates and proposals for such spaces exhibit such a little amount of regularity or other familiar properties that a satisfactory stochastic geometric analysis is yet to be developed before aiming at a further extension of classical comparison results.

Acknowledgements:

I would like to express my sincerest gratitude to my scientific advisor Karl-Theodor Sturm for giving me the opportunity to study the exciting area of stochastic differential geometry with him, for his confidence and support. Furthermore, my thanks go to Sergio Albeverio for creating this inspiring and cordial atmosphere in the stochastics group in Bonn and to all its other members for the very enjoyable time, in particular to Martin Hesse and Gustav Paulik. I am indebted also to Werner Ballmann and Anton Thalmaier for numerous enlightening discussions, as well as to Takashi Shioya for his generous hospitality during my stay at Tôhoku University. Finally, I am grateful to my family and friends for their long-lasting patience but most of all to Mignon for her empathy and understanding.

A remark on notation:

Throughout this work we make frequent use of Landau's symbols, which we understand in the following way. $O(s)$ (or sometimes $O_{\alpha,\beta,\cdots}(s)$) is a function depending on s (and parameters $\alpha, \beta, \cdots$) such that $O(s)/s \leq C$ for s in some neighborhood of zero, $o(s)$ has the property that $o(s)/s \to 0$ for $s \to 0$ and $\vartheta(s)$ is a function for which $\vartheta(s) \to 0$ for $s \to 0$ and which replaces in this text the usually more common notation $O(1)$. In order to distinguish from the case of a Riemannian manifold (M, g) we allocate the letters X to a metric space and Ξ to a diffusion process. Parallel translation along a curve γ on (M, g) is denoted by $/\!\!/_\gamma$ and $\mathbb{M}_{n,k}$ is the simply connected n-dimensional Riemannian manifold of constant curvature $k \in \mathbb{R}$.

2 Intrinsic Coupling on Manifolds

The coupling method by Kendall [Ken86] and Cranston [Cra91] provides an elegant stochastic proof of Yau's famous gradient estimates for harmonic functions on Riemannian manifolds with lower Ricci curvature bounds [Yau75]. Nevertheless from our point of view their construction of the coupling process has two weaknesses, namely the very 'extrinsic' concepts involved (like horizontal lifts of processes on a Riemannian manifold onto the orthonormal frame bundle and stochastic anti-development, for instance) and the somewhat nebulous arguments concerning the behaviour of the coupling process on the cut locus.

In the following paragraphs we aim at a more intrinsic construction of the coupling process on $M \times M$ which is motivated by a view towards non-smooth geometries as explained in the introduction. Also, the construction given below easily yields the crucial coupling probability estimate irrespective if the manifold has a nonempty cut locus or not.

2.1 Construction

The idea about coupling of Brownian Motion on a Riemannian manifold (M, g) is to construct a stochastic process $\overline{\Xi}$ on the product $M \times M$ such that

i) each factor $\Xi_1 = \pi_1(\overline{\Xi})$ and $\Xi_2 = \pi_2(\overline{\Xi})$ is a Brownian motion on (M, g)

ii) the compound process $d(\overline{\Xi})$ of $\overline{\Xi}$ with the intrinsic distance function d on M is dominated by a real semi-martingale ξ whose hitting time at zero $T_N(\xi)$ can be estimated from above.

Instead of using the SDE approach we follow the lines of the Markov chain approximation scheme for solutions to martingale problems for degenerate diffusion operators (cf. chapter 11 in [SV79]). Throughout this chapter we assume that (M, g) is smooth and complete.

As a preparation we recall Jørgensen's central limit theorem [Jør75] for geodesic random walks: Let (M, g) be a smooth Riemannian manifold of dimension d and fix for every $x \in M$ an isometry $\Phi_x : \mathbb{R}^d \xrightarrow{\simeq} T_x M$ such that the resulting function

$$\Phi(.) : M \to O(M), \Phi(x) : \mathbb{R}^d \xrightarrow{\simeq} T_x M \quad \forall x \in M$$

is measurable. Let $(\xi_k)_{k \in \mathbb{N}}$ be a sequence of $\mathbb{R}^d$-valued and independent random variables defined on some probability space $(\Omega, \mathcal{O}, P)$ whose distribution equals the normalized uniform distribution on S^{d-1}. A geodesic random walk $(\Xi_k^{\epsilon,x})_{k\in\mathbb{N}}$ with step size $\epsilon > 0$ and starting point $x \in M$ is given inductively by

$$\begin{aligned} \Xi_0^{\epsilon,x} &= x \\ \Xi_{k+1}^{\epsilon,x} &= \exp_{\Xi_k^{\epsilon,x}}(\epsilon \Phi_{\Xi_k^{\epsilon,x}} \xi_{k+1}), \end{aligned}$$

where exp is the exponential map of (M, g). By geodesic interpolation we can extend $(\Xi_k^{\epsilon,x})_{k\in\mathbb{N}}$ to a process $(\hat{\Xi}_t^{\epsilon,x})_{t\geq 0}$ with continuous time parameter, i.e.

$$\hat{\Xi}_t^{\epsilon,x} = \exp_{\Xi_{\lfloor t\rfloor}^{\epsilon,x}}\left((t - \lfloor t\rfloor)\epsilon \Phi_{\Xi_{\lfloor t\rfloor+1}^{\epsilon,x}} \xi_{\lfloor t\rfloor+1}\right)$$

with $\Xi_{-1}^{\epsilon,x} := x$, $\xi_0 := 0$ and $\lfloor t\rfloor := \sup\{k \in \mathbb{Z} | k < t\}$ for $t \in \mathbb{R}$. Then one may consider two different sequences of continuous time processes obtained from rescaling $\Xi_\cdot^{\epsilon,x}$, namely either by scaling the geodesic interpolation $\hat{\Xi}_k^x(t) := \hat{\Xi}_{kt}^{\sqrt{\frac{1}{k}},x}$ or by a Poisson subordination $\tilde{\Xi}_k^x(t) := \Xi_{\tau_k(t)}^{\sqrt{\frac{1}{k}},x}$ for $k \in \mathbb{N}_0$, where τ_k is a Poisson jump process on $\mathbb{N}$ with parameter k. Even if $\hat{\Xi}_k^x(.)$ is continuous by construction the cadlag process $\tilde{\Xi}_k^x(.)$ is easier to handle because it is a (time homogeneous) Markov process with transition function

$$P(\tilde{\Xi}_k^x(t) \in A \,|\, \tilde{\Xi}_k^x(s) = y) = e^{-(t-s)k} \sum_{i\geq 0} \frac{((t-s)k)^i}{i!} \mu_{1/\sqrt{k}}^i(y, A) =: (\tilde{P}_{t-s}^k \mathbb{1}_A)(y)$$

with $\mu_\epsilon(z, A) = ⨍_{S_z^{d-1}\subset T_zM} \mathbb{1}_A(\exp_z(\epsilon\theta))d\theta$ and $\mu^i = \mu \circ \mu \cdots \circ \mu$ (i times). The generator of $\tilde{\Xi}^k$, or equivalently of the semigroup $(\tilde{P}_t^k)_{t\geq 0}$, is therefore given by

$$A_k f(x) = k(⨍_{S_x^{d-1}} f(\exp_x(\frac{1}{\sqrt{k}}\theta)d\theta - f(x))) \stackrel{k\to\infty}{\longrightarrow} \Delta f(x) \quad \forall\, x \in M \tag{2.1}$$

where $\Delta^M f(x) = \frac{1}{2d}\,\mathrm{trace}(\mathrm{Hess} f)(x)$, see lemma 5.1[†]. Using (2.1) and Kurtz' semigroup approximation theorem it is easy to show that

$$\tilde{P}_t^k \longrightarrow P_t^M \text{ for } k \longrightarrow \infty$$

in the strong operator sense where $P_t^M = e^{t\Delta^M}$ is the heat semigroup on (M, g), and thus the weak convergence for the family $\tilde{\Xi}_\cdot^k$ to a Brownian motion $\Xi_\cdot^x$ starting in x is established by showing tightness of the distributions of $\tilde{\Xi}_\cdot^k$ on the Skorokhod path space $D_{\mathbb{R}_+}(M)$. Finally, the convergence of the sequence $\hat{\Xi}_\cdot^k$ to the same limit $\Xi_\cdot^x$ is proved by showing that the distance between $\hat{\Xi}_t^k$ and $\tilde{\Xi}_t^k$ tends to zero in probability locally uniformly with respect to t for k tending to infinity.

[†]Since we approximate the Laplacian on M by mean values on tangent spaces this definition of Δ is natural. Accordingly in this section a Brownian motion Ξ on (M, g) is a Markov process generated by Δ. Of course, the normalization of Δ is just a matter of linear time change for Ξ.

For a similar construction for two coupled Brownian motions with different starting points $x,\, y \in M \times M$ let $D(M) = \{(x,x) \in M \times M \,|\, x \in M\}$ be the diagonal in $M \times M$. Then for all $x,\, y \in M \times M \setminus D(M)$ choose some minimal geodesic $\gamma_{xy} : [0,1] \to M$ connecting x and y. Fix a function

$$\begin{aligned} \Phi(.,.) &: M \times M \setminus D(M) \to O(M) \times O(M) \\ \Phi_1(x,y) &:= \pi_1 \circ \Phi(x,y) : \mathbb{R}^d \overset{\simeq}{\to} T_x M \\ \Phi_2(x,y) &:= \pi_2 \circ \Phi(x,y) : \mathbb{R}^d \overset{\simeq}{\to} T_y M \end{aligned}$$

with the additional property that

$$\Phi_1(x,y)e_1 = \frac{\dot\gamma_{xy}(0)}{\|\dot\gamma_{xy}(0)\|}, \Phi_2(x,y)e_1 = \frac{\dot\gamma_{yx}(0)}{\|\dot\gamma_{yx}(0)\|} \text{ if } x \neq y \tag{$*$}$$

where e_i is the i-th unit vector in $\mathbb{R}^d$. On the diagonal $D(M)$ we set

$$\Phi(x,x) := (\phi(x), \phi(x)) \in O_x(M) \times O_x(M) \tag{$*_D$}$$

where $\phi : M \to O(M)$ is some choice of bases as in the previous paragraph.

In the existence and regularity statement for a possible choice of Φ below $\mathrm{Cut}(M) \subset M \times M$ is defined as the set of all pairs of points (x,y) which can be joined by at least two distinct minimal geodesics, hence $\mathrm{Cut}(M)$ itself is symmetric and measurable.

Lemma 2.1. *There is some choice of a minimal geodesic γ_{xy} (parameterized on $[0,1]$) for each $(x,y) \in M \times M$ such that the resulting map $\gamma : M \times M \to C^1([0,1],M)$, $(x,y) \mapsto \gamma_{xy}$ is measurable, symmetric, i.e. $\gamma_{xy}(t) = \gamma_{yx}(1-t)$ for all $t \in [0,1]$, and continuous on $M \times M \setminus (D(M) \cup \mathrm{Cut}(M))$. Furthermore, for any measurable frame map $\phi : M \to \Gamma(O(M))$ it is possible to find a measurable function $\Phi : M \times M \to O(M) \times O(M)$ satisfying the conditions $(*)$ and $(*_D)$ above and which is continuous on $M \times M \setminus (D(M) \cup \mathrm{Cut}(M))$.*

Proof. Suppose first that we found a measurable symmetric function $\gamma : M \times M \to C^1([0,1],M)$ as above and let $\psi_i \in \Gamma(O(M))$, $i = 1,2$ be two arbitrary continuous frame maps on M. For $(x,y) \in M \times M \setminus D(M)$ we construct a new orthonormal frame on $T_xM \oplus T_yM$ by $\Phi(x,y) = \{\dot\gamma_{xy}/\|\dot\gamma_{xy}\|, \tilde\psi_1^2, \ldots, \tilde\psi_1^d, \dot\gamma_{yx}/\|\dot\gamma_{yx}\|, \tilde\psi_2^2, \ldots, \tilde\psi_2^d\}$ out of the frame $\{\psi_1(x), \psi_2(y)\}$ via Schmidt's orthogonalization procedure applied to the vectors $\psi_i(e_k)$, $k = 1, \ldots, d$ in the orthogonal complements of $\dot\gamma_{xy}$ and $\dot\gamma_{yx}$ in T_xM and T_yM respectively. Since the maps $\partial_{s_{|s=0}}$ and $\partial_{s_{|s=1}} : C^1([0,1],M) \to TM$ are continuous and the construction of the basis Φ in $T_xM \oplus T_yM$ depends continuously on the data $\{\psi_1, \psi_2\}$, $\dot\gamma_{xy}$ and $\dot\gamma_{yx}$ it is clear that the map Φ inherits the regularity properties of the function γ on $M \times M \setminus D(M)$. Since $D(M)$ is closed in $M \times M$ and hence measurable any extension of

Φ by a measurable $(\phi(.), \phi(.))$ as above on $D(M)$ yields a measurable map on the whole $M \times M$. This proves the second part of the lemma.

Thus it remains to find a map γ as desired. In order to deal with the symmetry condition we first introduce a continuous complete ordering $\geq$ on M (which can be obtained as an induced ordering from an embedding of M into a high dimensional Euclidean space $\mathbb{R}^l$ and some complete ordering on $\mathbb{R}^l$) and restrict the discussion to the closed subset $D_-(M) = \{(x, y) \,|\, x \geq y\} \subset M \times M$ endowed with its Borel σ-algebra which is the trace of $\mathcal{B}(M \times M)$ on $D_-(M)$. We define a measurable set-valued map $\Gamma : D_-(M) \to 2^{C^1([0,1],M)}$ as follows: for each $\epsilon > 0$ choose some ϵ-net $P^\epsilon = \{p_i^\epsilon \,|\, i \in \mathbb{N}\}$ in $D_-(M)$ and choose some minimal geodesic $\gamma_{p_i^\epsilon, p_j^\epsilon}$ for each pair of points p_i^ϵ, $p_j^\epsilon \in P^\epsilon$. Arrange the set of pairs $(p_i^\epsilon, p_j^\epsilon)$ into a common sequence $\{(p_{i_k}^\epsilon, p_{j_k}^\epsilon) \,|\, k \in \mathbb{N}\}$ and let $\gamma^\epsilon : D_-(M) \to C^1([0,1], M)$ be the map defined inductively by

$$\gamma^\epsilon(x, y) = \gamma_{p_{i_0}^\epsilon p_{j_0}^\epsilon} \text{ for } (x, y) \in B_{2\epsilon}(p_{i_0}^\epsilon, p_{j_0}^\epsilon)$$

$$\gamma^\epsilon(x, y) = \gamma_{p_{i_{k+1}}^\epsilon p_{j_{k+1}}^\epsilon} \text{ for } (x, y) \in B_{2\epsilon}(p_{i_{k+1}}^\epsilon, p_{j_{k+1}}^\epsilon) \setminus \bigcup_{l=0}^{k} B_{2\epsilon}(p_{i_l}^\epsilon, p_{j_l}^\epsilon)$$

It is clear from the definition that the functions γ^ϵ are measurable and, moreover, using the geodesic equation in (M, g) together with the Arzela-Ascoli-theorem it is easy to see that for each $(x, y) \in D_-(M)$ the set of curves $\{\gamma_{xy}^\epsilon\}_{\epsilon > 0}$ are relatively compact in $C^1([0,1], M)$. Trivially any limit point of $\{\gamma_{xy}^\epsilon\}_{\epsilon > 0}$ for ϵ tending to zero will be a minimal geodesic from x to y. Let us choose a priori some sequence $\epsilon_k \to 0$ for $k \to \infty$ then we define the set valued function $\Gamma : D_-(M) \to 2^{C^1([0,1],M)}$ for $(x, y) \in D_-(M)$ as the collection of all possible limit points of $\Gamma^{\epsilon_k}(x, y)$, i.e.

$$\Gamma(x, y) := \left\{ \gamma_{xy} \;\middle|\; \begin{array}{c} \exists \text{ subsequence } \epsilon_{k'} \text{ and } \gamma_{xy}^{\epsilon_{k'}} \in \Gamma^{\epsilon_{k'}}(x, y) : \\ \gamma_{xy}^{\epsilon_{k'}} \to \gamma_{xy} \text{ in } C^1([0,1], M) \text{ for } k' \to \infty \end{array} \right\} \subset C^1([0,1], M).$$

The fact that we can find a measurable 'selector', i.e. a measurable map $\gamma : D_-(M) \to C^1([0,1], M)$ with $\gamma(x, y) \in \Gamma(x, y)$ follows from a measurable selection theorem as formulated in the subsequent lemma. Furthermore, the uniqueness of γ_{xy} and compactness arguments imply that any such selector obtained from the map Γ above must be continuous on $D_-(M) \cap (\mathrm{Cut}(M) \cup (D(M))^c$. It is also clear that γ_{xx} is the constant curve in x for all $x \in M$ and hence we may extend our chosen γ from $D_-(M)$ continuously onto the whole $M \times M$ by putting $\gamma_{yx}(t) := \gamma_{xy}(1 - t)$ if $(x, y) \in D_-(M)$. This proves the first assertion of the lemma and the proof is completed. $\square$

Lemma 2.2. *Let $(X, \mathcal{S})$ be a measurable and (Y, d) be a complete separable metric space endowed with its Borel σ-algebra $\mathcal{B}(Y)$. Let furthermore $f_k : X \to Y$ be a sequence of measurable functions which are pointwise relatively compact, i.e. for all x in X the set*

$\{f_k(x)\}_{k\in\mathbb{N}}$ is relatively compact in Y. Let the set valued map $F : X \to 2^Y$ be defined by pointwise collecting all possible limit points of the sequence f_k. Then there is a measurable function $f : X \to Y$ with $f(x) \in F(x)$ for all $x \in X$.

Proof. Since the set $F(x)$ is obviously closed for any x in X it remains to check the measurability of F, i.e. we need to show that $F^{-1}(O) := \{x|F(x) \cap O \neq \emptyset\}$ is measurable in X for any $O \subset Y$ open. Since any open $O \subset Y$ can be exhausted by countably many set of the type $\overline{B_\delta(y)}$ with $\delta > 0$, $y \in Y$ we may replace O by $\overline{B_\delta(y)}$ in the condition above. But using the pointwise compactness of the sequence f_k and a diagonal sequence argument it is easy to show that

$$F^{-1}\left(\overline{B_\delta(y)}\right) = \bigcap_{\delta'>\delta} \limsup_{k\to\infty} f_k^{-1}\left(B_{\delta'}(y)\right).$$

Choosing some sequence $\delta'_l \searrow \delta$ we see that in fact $F^{-1}\left(\overline{B_\delta(y)}\right)$ is measurable. Hence we may apply the measurable selection theorem of Kuratowksi and Ryll-Nardzewski to the function F which yields the claim. □

We now take two independent sequences $(\xi_k)_{k\in\mathbb{N}}$ and $(\eta_k)_{k\in\mathbb{N}}$ of $\mathbb{R}^d$-valued i.i.d. random variables with normalized uniform distribution on S^{d-1} and define a coupled geodesic random walk $\Xi_k^{\epsilon,(x,y)} = (\Xi_{1,k}^{\epsilon,(x,y)}, \Xi_{2,k}^{\epsilon,(x,y)})$ with step size ϵ and starting point (x, y) in $M \times M$ inductively by

$$\Xi_0^{\epsilon,(x,y)} = (x, y)$$

and if $\Xi_k^{\epsilon,(x,y)} \in M \times M \setminus D(M)$:

$$\Xi_{k+1}^{\epsilon,(x,y)} = \left(\exp_{\pi_1(\Xi_k^{\epsilon,(x,y)})}[\epsilon\Phi_1(\Xi_k^{\epsilon,(x,y)})\xi_{k+1}], \exp_{\pi_2(\Xi_k^{\epsilon,(x,y)})}[\epsilon\Phi_2(\Xi_k^{\epsilon,(x,y)})\xi_{k+1}]\right) \tag{2.2}$$

if $\Xi_k^{\epsilon,(x,y)} \in D(M)$:

$$\Xi_{k+1}^{\epsilon,(x,y)} = \left(\exp_{\pi_1(\Xi_k^{\epsilon,(x,y)})}[\epsilon\phi(\pi_1(\Xi_k^{\epsilon,(x,y)}))\xi_{k+1}], \exp_{\pi_1(\Xi_k^{\epsilon,(x,y)})}[\epsilon\phi(\pi_1(\Xi_k^{\epsilon,(x,y)}))\eta_{k+1}]\right) \tag{2.3}$$

where π_i, $i = 1, 2$ are the projections of $M \times M$ on the first and second factor respectively.

As before we have at least two possibilities to extend $\Xi_k^{\epsilon,(x,y)}$ to a process with continuous time parameter $t \in \mathbb{R}_+$, namely

i) by geodesic interpolation $\hat{\Xi}_t^{\epsilon,(x,y)}$,

ii) by Poisson subordination $\tilde{\Xi}_{\tau_\lambda(t)}^{\epsilon,(x,y)}$.

In particular choosing $\epsilon = 1/\sqrt{k}$ and $\lambda = k$ in ii) for $k \in \mathbb{N}$ one obtains a sequence of Markov processes $\tilde{\overline{\Xi}}_t^{k,(x,y)} = \Xi_{\tau_k(t)}^{1/\sqrt{k},(x,y)}$ on $M \times M$ with transition function

$$P\big(\tilde{\overline{\Xi}}^{k,(x,y)}(t) \in A \times B \big| \tilde{\overline{\Xi}}^{k,(x,y)}(s) = (u,v)\big)$$
$$= e^{-(t-s)k} \sum_{i \geq 0} \frac{((t-s)k)^i}{i!} \overline{\mu}^i_{1/\sqrt{k}}((u,v), A \times B)$$

where the kernel $\overline{\mu}_\epsilon : M^2 \times \mathcal{B}(M^2) \to \mathbb{R}$ is given by

$$\overline{\mu}_\epsilon((u,v), A \times B) = \begin{cases} \displaystyle\fint_{S_0^{d-1} \subset \mathbb{R}^d} \mathbb{1}_A(\exp_u(\epsilon \Phi^1_{(u,v)}\theta)) \mathbb{1}_B(\exp_u(\epsilon\Phi^2_{(u,v)}\theta))d\theta \\ \qquad\qquad\qquad\qquad\qquad \text{if } (u,v) \in D(M)^c \\ \displaystyle\fint_{S_u^{d-1} \subset T_u M} \mathbb{1}_A(\exp_u(\epsilon\theta))d\theta \cdot \fint_{S_v^{d-1} \subset T_v M} \mathbb{1}_B(\exp_v(\epsilon\theta))d\theta \\ \qquad\qquad\qquad\qquad\qquad \text{else.} \end{cases}$$

Trivially the generator of the semigroup $(\tilde{\overline{P}}_t^{k,(x,y)})_{t \geq 0}$ induced by $\tilde{\overline{\Xi}}^{k,(x,y)}$ is

$$\overline{L}_k = k(\overline{\mu}_{1/\sqrt{k}} - Id).$$

Lemma 2.3. *Let $F : M \times M \to \mathbb{R}$ be a smooth function. Then for $k \to \infty$*

$$\overline{L}_k F(u,v) \longrightarrow L_c F(u,v) \quad \forall (u,v) \in M \times M, \textit{locally uniformly on } D(M)^c$$

where the operator $L_c = L_c^{M,\phi}$ is defined by

$$L_c(f \otimes g) = \Delta f \otimes g + f \otimes \Delta g + \mathbb{1}_{D(M)^c} \langle \nabla f, \nabla g \rangle_\Phi \tag{2.4}$$
$$\langle \nabla f, \nabla g \rangle_\Phi(x,y) := \frac{1}{d} \langle \Phi_1^{-1}(x,y) \nabla f(x), \Phi_2^{-1}(x,y) \nabla g(y) \rangle_{\mathbb{R}^d}$$

whenever $F : M \times M \to \mathbb{R}$ is of the form $F = f \otimes g$ with smooth f, $g : M \to \mathbb{R}$. Moreover, for the case $F = f \otimes 1$ or $F = 1 \otimes g$ one finds $\overline{L}_k f \otimes 1 \to \Delta f \otimes 1$ and $\overline{L}_k 1 \otimes g \to 1 \otimes \Delta g$ locally uniformly on $M \times M$ for k tending to infinity.

Proof. Suppose first that $(u,v) \in D(M)^c$ and let U be some neighborhood with $(u,v) \in U \subset D(M)^c$. Now for any $(u',v') \in U$ the Taylor expansion of $F = f \otimes g$ about (u',v') and the definition of the exponential map yield

$$f\big(\exp_{u'}(\frac{1}{\sqrt{k}}\Phi^1_{(u',v')}\theta)\big) g\big(\exp_{v'}(\frac{1}{\sqrt{k}}\Phi^2_{(u',v')}\theta)\big)$$
$$= f(u')g(v') + \frac{1}{\sqrt{k}} f(u') \langle \nabla g(v'), \Phi^2_{(u',v')}\theta \rangle_{T_{v'}M} + \frac{1}{\sqrt{k}} g(v') \langle \nabla f(u'), \Phi^1_{(u',v')}\theta \rangle_{T_{u'}M}$$
$$+ \frac{1}{k} \langle \nabla f(u'), \Phi^1_{(u',v')}\theta \rangle_{T_{u'}M} \cdot \langle \nabla g(v'), \Phi^2_{(u',v')}\theta \rangle_{T_{v'}M}$$
$$+ \frac{1}{2k} f(u') \mathrm{Hess} g_{v'}(\Phi^2_{(u',v')}\theta, \Phi^2_{(u',v')}\theta) +$$
$$\frac{1}{2k} g(v') \mathrm{Hess} f_{u'}(\Phi^1_{(u',v')}\theta, \Phi^2_{(u',v')}\theta) + o_{u',v'}(\frac{1}{k})$$

where $o(.)$ is a "little o" Landau function. In fact, $o_{u',v'}(\frac{1}{k})$ can be replaced by $o_U(\frac{1}{k})$ due to the smoothness of the data (M,g) and the function F. Inserting this into $\overline{L}_k$ gives

$$\begin{aligned}\overline{L}_k(F)(u',v') &= \Delta f(u')g(v') + f(u')\Delta g(v') \\ &+\frac{1}{d}\langle \Phi_1^{-1}(u',v')\nabla f(u'), \Phi_2^{-1}(u',v')\nabla g(v')\rangle_{\mathbb{R}^d} + \vartheta_U(\frac{1}{k})\end{aligned} \tag{2.5}$$

with $\vartheta_U(\frac{1}{k}) \to 0$ for $k \to \infty$ because

$$\begin{aligned}\fint_{S^{d-1}} \langle \nabla f(u'), \Phi^1_{(u',v')}\theta\rangle_{T_{v'}M} d\theta &= \fint_{S^{d-1}} \langle \nabla g(v'), \Phi^2_{(u',v')}\theta\rangle_{T_{v'}M} d\theta = 0 \\ \frac{1}{2}\fint_{S^{d-1}} \mathrm{Hess} f_{u'}(\Phi^1_{(u',v')}\theta, \Phi^2_{(u',v')}\theta)d\theta &= \Delta f(u') \\ \frac{1}{2}\fint_{S^{d-1}} \mathrm{Hess} g_{v'}(\Phi^2_{(u',v')}\theta, \Phi^2_{(u',v')}\theta) &= \Delta g(v') \\ \fint_{S^{d-1}} \langle \nabla f(u'), \Phi^1_{(u',v')}\theta\rangle_{T_{u'}M} \cdot \langle \nabla g(v'), \Phi^2_{(u',v')}\theta\rangle_{T_{v'}M} & \\ = \frac{1}{d}\langle \Phi_1^{-1}(u',v')\nabla f(u'), \Phi_2^{-1}(u',v')\nabla g(v')\rangle_{\mathbb{R}^d}.\end{aligned}$$

Now if $(u,v) \in D(M)$ by definition of $\overline{L}_k$ the coupling term $\langle \nabla f, \nabla g\rangle_\Phi$ does not appear and thus the claim is proved. □

Remark 2.1. The previous proof remains the same for general smooth functions $F : M \times M \to \mathbb{R}$. The characterization of L_c as above is just more instructive.

2.2 The Central Limit Theorem for Coupled Random Walks

The operator L_c has two irregular properties, one being its degeneracy, i.e. the second order part acts only in d of the $2d$ directions, and the other one being the discontinuity of the coefficients on $D(M) \cup \mathrm{Cut}(M)$. Both features together cause problems for the definition of a semigroup e^{tL_c} via the Hille-Yosida theorem. Therefore we confine ourselves to the construction of a solution $\overline{\Xi}$ to the martingale problem for L_c in a restricted sense by showing compactness of the (laws of the) sequence of processes $\big(\tilde{\Xi}^{k,(x,y)}_{\cdot}\big)_k$ on the space $D_{\mathbb{R}_+}(M \times M)$ of cadlag paths equipped with the Skorokhod topoply.

Theorem 2.1 (Coupling Central Limit Theorem). *The sequence* $\big(\tilde{\overline{\Xi}}^{k,(x,y)}_{\cdot}\big)_{k\geq 0}$ *is tight on* $D_{\mathbb{R}_+}(M \times M)$ *and any weak limit of a converging subsequence* $(\tilde{\overline{\Xi}}^{k',(x,y)}_{\cdot})_{k'}$ *is a solution to the martingale problem for* L_c *in the following restricted sense: let*

$$(\Omega^\infty, P^{\infty,(x,y)}, (\overline{\Xi}^{\infty,(x,y)}_s)_{s\geq 0}) = (D_{\mathbb{R}_+}(M \times M), \underset{k'\to\infty}{w\text{-}\lim}(\overline{\Xi}^{k',(x,y)}_{\cdot})_* P, (\pi_s)_{s\geq 0})$$

denote the canonical process on $M \times M$ *induced from a limit measure* $w\text{-}\lim_{k' \to \infty}(\Xi_{\cdot}^{k',(x,y)})_* P$ *on* $D_{\mathbb{R}_+}(M \times M)$ *and the natural coordinate projections* $\pi_s : D_{\mathbb{R}_+}(M \times M) \to M \times M$, *then for all* $F \in C_0^\infty(M \times M \setminus (D(M) \cup \mathrm{Cut}(M)))$, $F = f \otimes 1$ *or* $F = 1 \otimes g$ *with smooth* $f, g : M \to \mathbb{R}$ *the process*

$$F(\Xi_t^{(x,y)}) - F(x,y) - \int_0^t L_c F(\Xi_s^{(x,y)}) ds$$

is a $P^{\infty,(x,y)}$ *martingale with initial value* 0. *In particular, under* $P^{\infty,(x,y)}$ *both marginal processes* $(\Xi_s^1 := \pi_s^1)_{s \geq 0}$ *and* $(\Xi_s^2 := \pi_s^2)_{s \geq 0}$ *are standard Brownian motions on* (M, g) *starting in* x *and* y *respectively.*

Remark 2.2. We do not claim uniqueness here for the solution to the martingale problem in the form stated above nor a Markov property. Note also that we cautiously circumvented the problem of the cut locus by the choice of admissible test functions F.

We will call for short any probability measure on $D_{\mathbb{R}_+}(M \times M)$ with the properties above a solution to the (restricted) coupling martingale problem.

For the proof of theorem 2.1 we show that the corresponding arguments for Euclidean diffusions carry over to the present situation with only few changes. We follow along the lines of chapter 8 in [Dur96], extending and simplifying the results in [Jør75].

Remember that a family $(P_i)_{i \in I} \subset \mathcal{P}(X)$ of Borel probability measures on a topological space (X, τ) is called tight iff

$$\forall \epsilon > 0 \, \exists K \subset X \text{ compact: } \inf_i P_i(K) \geq 1 - \epsilon.$$

If (X, τ) is metrizable, complete and separable then Prohorov's theorem states that the family $(P_i)_{i \in I}$ is tight if and only if it is relatively compact with respect to the weak-$*$-topology induced on $\mathcal{P}(X)$ from the uniformly continuous functions on (X, τ). - When talking about tightness of stochastic process $(\Xi_{\cdot}^k)_k$ defined on probability spaces $(\Omega_k, \mathcal{O}_k, P_k)$ and with a common state space X one actually means tightness of the corresponding image measures $(\Xi_{\cdot}^k)_* P_k$ on (a properly chosen and topologized subset) of the path space $X^{\mathbb{R}_+}$, which in the present context is $D_{\mathbb{R}_+}(M \times M)$ equipped with the Skorokhod topology, where $M \times M$ is endowed with the standard product metric $d((x_1, y_1), (x_2, y_2)) = \sqrt{d^2(x_1, x_2) + d^2(y_1, y_2)}$.

The tightness of the sequence $(\tilde{\Xi}_{\cdot}^{k,(x,y)})_k$ is shown by verifying that it satisfies the conditions i) and ii) of the following theorem:

Theorem 2.2 (Tightness criterion on $D_{\mathbb{R}_+}(X)$). *Let (X,d) be a complete and separable metric space and let $(\Omega_n, P_n, (\Xi^n_t)_{t\geq 0})_{n\in\mathbb{N}}$ be a sequence of cadlag processes on X. Then the following condition is sufficient for tightness of $(\Xi^n_\cdot)_*(P_N)$ on $D_{\mathbb{R}_+}(X)$: For all $N\in\mathbb{N}$, and η, $\epsilon>0$ there are $x_0\in X$, $n_0\in\mathbb{N}$ and M, $\delta>0$ such that*

i) $P_n(d(x_0,\Xi^n_0)>M)\leq\epsilon$ *for all* $n\geq n_0$

ii) $P_n(w(\Xi^n,\delta,N)\geq\eta)\leq\epsilon$ *for all* $n\geq n_0$

with the modulus of continuity $w(\Xi^n,\delta,N)(\omega):=\sup_{\substack{0\leq s,\,t\leq N\\|s-t|\leq\delta}} d(\Xi^n_s(\omega),\Xi^n_t(\omega))$.

Proof. This follows just as in the case $D_{[0,1]}(X)$ essentially from the inequality $w'(x,\delta,N)\leq w(x,2\delta,N)$ for $x\in D_{\mathbb{R}_+}(X)$ and $\delta\leq\frac{1}{2}N$ where

$$w'(x,\delta,N)=\inf_{\substack{0=t_0<t_1\cdots<t_k=N\\|t_i-t_{i-1}|\geq\delta}}\max_{i\leq k}\sup_{s,\,t\in[t_i,t_{i+1})} d(x(s),x(t))$$

and the characterization of compact subsets in $D_{\mathbb{R}_+}(X)$ by the functionals $w'(x,\delta,N)$, cf. theorem 15.5 in [Bil68] and theorem VI.1.5 in [JS87]. □

First we state an auxiliary result concerning the stopping times $\tau^{k,(x,y)}_\eta$ on $(\Omega,\mathcal{O},P)$[‡] given by

$$\tau^{k,(x,y)}_\eta=\inf\big\{s\geq 0\,|\,d(\tilde{\overline{\Xi}}^{k,(x,y)}_s,\tilde{\overline{\Xi}}^{k,(x,y)}_0)\geq\eta\big\}$$

Lemma 2.4. *For all compact $K\subset M\times M$ and η there is a $C_\eta\geq 0$ such that $\forall\delta>0$*

$$\sup_{k\geq 0}\sup_{(x,y)\in K} P(\tau^{k,(x,y)}_\eta<\delta)\leq C_\eta\delta.$$

Proof. Since $\big(\tilde{\overline{\Xi}}^{k,(x,y)}_t\big)^{(x,y)\in M\times M}_{t\geq 0}$ is a continuous time Markov process on $M\times M$ with generator $\overline{L}_k$ for all $k\in\mathbb{N}$ and $F\in C_0(M\times M)$ the process

$$F\big(\tilde{\overline{\Xi}}^{k,(x,y)}_t\big)-\int_0^t \overline{L}_kF\big(\tilde{\overline{\Xi}}^{k,(x,y)}_s\big)ds$$

is a P-martingale with respect to the parameter $t\in\mathbb{R}$. Furthermore, from lemma 2.3 and the smoothness of M it follows that for $K\subset M\times M$ compact one can find a family of smooth cut of functions $\big(F_{(x,y),\eta}(.,.):M\times M\to\mathbb{R}\big)_{(x,y)\in K}$ with $F_{(x,y),\eta}=1$ on $B_\eta(x,y)$, vanishing on $B_{2\eta}((x,y))^c$ and such that

$$\sup_{k\geq 0}\sup_{(x,y)\in K}\big\|\overline{L}_kF_{(x,y),\eta}\big\|_{\infty,M\times M}\leq C_\eta$$

‡The terms stopping time, martingale etc. are used with respect to the filtrations generated by $\tilde{\overline{\Xi}}^{k,(x,y)}_{\tau_k(t)}$.

(this follows from (2.5) for a function like $F_{(x,y),\eta}((u,v)) = h_\eta(d(x,u))h_\eta(d(y,v))$ with a smooth real cut-off function $h_\eta \in C^3(\mathbb{R})$) and which implies that the process

$$\big(F_{(x,y),\eta}(\tilde{\Xi}_t^{k,(x,y)}) + tC_\eta\big)_{t\geq 0}$$

is a submartingale starting with the initial value 1. The optional stopping theorem applied to $\tau_\eta^{k,(x,y)} \wedge \delta$ yields

$$\mathbb{E}_P\Big[F_{(x,y),\eta}\big(\tilde{\Xi}^{k,(x,y)}_{\tau_\eta^{k,(x,y)}\wedge\delta}\big) + C_\eta(\tau_\eta^{k,(x,y)} \wedge \delta)\Big] \geq 1$$

from which one obtains by rearrangement

$$C_\eta\delta \geq C_\eta\mathbb{E}_P(\tau_\eta^{k,(x,y)} \wedge \delta) \geq \mathbb{E}_P\big[1 - F_{(x,y),\eta}\big(\tilde{\Xi}^{k,(x,y)}_{\tau_\eta^{k,(x,y)}\wedge\delta}\big)\big] \geq P(\tau_\eta^{k,(x,y)} \leq \delta).$$

Here the last inequality follows from the right continuity of $\tilde{\Xi}^{k,(x,y)}$ and the definition of $F_{(x,y),\eta}$. □

For N, $k \in \mathbb{N}$ and $\eta > 0$ let us introduce the following stopping times $\tau_n = \tau_{n,\eta}^{k,(x,y)}$ and functionals $\sigma = \sigma_{N,\eta}^{k,(x,y)}$, $\theta = \theta_N^{k,(x,y)}$ of the process $\big(\tilde{\Xi}_t^{k,(x,y)}\big)$ by

$$\tau_0 := 0, \quad \tau_1 := \tau_\eta^{k,(x,y)}$$
$$\tau_n = \inf\big\{s > \tau_{n-1} \,|\, d(\tilde{\Xi}^{k,(x,y)}_{\tau_{n-1}}, \tilde{\Xi}^{k,(x,y)}_s) \geq \eta\big\}$$
$$\sigma = \min\{\tau_n - \tau_{n-1} \,\big|\, \tau_{n-1} \leq N\}$$
$$\theta = \sup\big\{d(\tilde{\Xi}^{k,(x,y)}_s, \tilde{\Xi}^{k,(x,y)}_{s-}) \,\big|\, s \leq N\big\}$$

Lemma 2.5. *If $\sigma_{N,\eta}^{k,(x,y)} > \delta$ and $\theta_N^{k,(x,y)} \leq \eta$ then $w(\Xi^n, \delta, N) \leq 4\eta$.*

Proof. Given $0 \leq s < t \leq N$ with $t - s < \delta$, we have either $\tau_{n-1} \leq s < \tau_n \leq t < \tau_{n+1}$, in which case

$$\begin{aligned} d(\tilde{\Xi}^{k,(x,y)}_s, \tilde{\Xi}^{k,(x,y)}_t) &\leq d(\tilde{\Xi}^{k,(x,y)}_s, \tilde{\Xi}^{k,(x,y)}_{\tau_{n-1}}) + d(\tilde{\Xi}^{k,(x,y)}_{\tau_{n-1}}, \tilde{\Xi}^{k,(x,y)}_{\tau_n-}) \\ &\quad + d(\tilde{\Xi}^{k,(x,y)}_{\tau_n-}, \tilde{\Xi}^{k,(x,y)}_{\tau_n}) + d(\tilde{\Xi}^{k,(x,y)}_{\tau_n}, \tilde{\Xi}^{k,(x,y)}_t) \leq 4\eta \end{aligned}$$

or $\tau_n \leq s < t \leq \tau_{n+1}$ which yields

$$d(\tilde{\Xi}^{k,(x,y)}_s, \tilde{\Xi}^{k,(x,y)}_t) \leq d(\tilde{\Xi}^{k,(x,y)}_s, \tilde{\Xi}^{k,(x,y)}_{\tau_n}) + d(\tilde{\Xi}^{k,(x,y)}_{\tau_n}, \tilde{\Xi}^{k,(x,y)}_t) \leq 2\eta$$ □

Proof of theorem 2.1. Tightness. For given compact $K \subset M \times M$, $(x, y,) \in K$, we may assume w.l.o.g. that the sample paths of all $\big(\tilde{\Xi}^{k,(x,y)}_\cdot\big)_{k\in\mathbb{N}}$ lie entirely in $B_{1/\sqrt{k}}(K)$, because otherwise we stop each process when it leaves K and let $K \nearrow M \times M$ in a final step.

Since the condition i) of theorem 2.2 is trivially satisfied, by lemma 2.5 it is sufficient to show that for all $N \in \mathbb{N}$, and η, $\epsilon > 0$ one can find $\delta > 0$ and $k_0 \in \mathbb{N}$ such that $P(\sigma_{N,\eta}^{k,(x,y)} > \delta; \theta_N^{k,(x,y)} \leq \eta/4) \geq 1 - \epsilon$ for all $k \geq k_0$.

First note that trivially $P(\theta_N^{k,(x,y)} > \eta) = 0$ for $k \geq 1/\eta^2$ because

$$d(\tilde{\overline{\Xi}}_s^{k,(x,y)}, \tilde{\overline{\Xi}}_{s-}^{k,(x,y)}) \leq \sqrt{\frac{1}{k}(\tau_k(s) - \tau_k(s-))} \leq \sqrt{\frac{1}{k}} \quad P\text{-a.s.} \tag{2.6}$$

and we may assume w.l.o.g. that τ_k is a regular Poisson process with jumps of size 1. Thus it remains to verify that it is possible to find $\delta > 0$ and k_0 such that $P(\sigma_{N,\eta}^{k,(x,y)} \leq \delta) \leq \epsilon$ for all $k \geq k_0$. For this introduce the functional $L_{N,\eta}^{k,(x,y)} = \max\{l | \tau_{l,\eta}^{k,(x,y)} \leq N\}$. Lemma 2.4 and Fubini's theorem yield

$$\begin{aligned}
\mathbb{E}_P(\exp(-\tau_\eta^{k,(x,y)})) =& 1 - \int_{\mathbb{R}_+} (1 - P(\tau \leq s))e^{-s} ds \leq 1 - \int_0^{\frac{1}{2C_\eta}} (1 - C_\eta s)e^{-s} ds \\
&\leq 1 - \frac{1}{2} \int_0^{\frac{1}{2C_\eta}} e^{-s} ds =: \lambda_\eta < 1 \quad \forall\, (x,y) \in K,\, k \in \mathbb{N}
\end{aligned}$$

and by the strong Markov property $\sup_k \sup_{(x,y)\in K} \mathbb{E}_P(\exp(-\tau_{l,\eta}^{k,(x,y)})) \leq \lambda_\eta^l$. Consequently for all $k \in \mathbb{N}$, $(x,y) \in K$

$$P(L_{N,\eta}^{k,(x,y)} \geq l) = P(\tau_{l,\eta}^{k,(x,y)} \leq N) \leq e^N \mathbb{E}_P(\exp(-\tau_{l,\eta}^{k,(x,y)})) \leq e^N \lambda_\eta^l$$

and thus finally for any $\delta > 0$, $l \in \mathbb{N}$

$$\begin{aligned}
P(\sigma_{N,\eta}^{k,(x,y)} \leq \delta) &\leq P(\sigma_{N,\eta}^{k,(x,y)} \leq \delta; L_{N,\eta}^{k,(x,y)} \leq l) + P(L_{N,\eta}^{k,(x,y)} \geq l) \\
&\leq l \sup_{(u,v)\in K} P(\tau_\eta^{k,(u,v)} \leq \delta) + e^N \lambda_\eta^l \\
&\leq l C_\eta \delta + e^N \lambda_\eta^l
\end{aligned}$$

by lemma 2.4. Hence for N, η, ϵ given choose $l_0 \in \mathbb{N}$ such that $e^N \lambda_\eta^{l_0} \leq \epsilon/2$ and $\delta > 0$ such that $l_0 C_\eta \delta \leq \epsilon/2$. Then choosing $k_0 \geq 2/\eta^2$ establishes condition ii) and the tightness assertion is proved.

Martingale Problem for L_c. Since $\overline{L}_k$ generates the process $\left(\tilde{\overline{\Xi}}_\cdot^{k,(x,y)}\right)$ this is also true for its realization on the path space $(D_{\mathbb{R}_+}(M \times M), (\overline{\Xi}_\cdot^{k,(x,y)})_* P, (\pi_s)_{s\geq 0})$ and which is in this case equivalent to

$$F(\pi_t) - F(x,y) - \int_0^t \overline{L}_k F(\pi_s) ds \text{ is a nomalized } (\overline{\Xi}_\cdot^{k,(x,y)})_* P\text{-Martingale} \tag{2.7}$$

for all $F \in \mathrm{Dom}(\overline{L}_k) \supset C_0^3(M \times M)$. If $F \in C_0^\infty(M \times M \setminus (D(M) \cup \mathrm{Cut}(M)))$ by lemma 2.3 $\left\|\overline{L}_k F - L_c F\right\|_\infty \to 0$ for $k \to \infty$ and thus by a general continuity argument (cf. lemma 5.5.1. in [EK86]) we may pass to the limit in the statement above provided k' is a subsequence such that $\text{w-}\lim_{k' \to \infty}(\overline{\Xi}^{k',(x,y)}_{\cdot})_* P$ exists. Finally, the assertion concerning the marginal processes follows either from the results [Jør75] or from putting $F = 1 \otimes f$ and $F = f \otimes 1$ in (2.7) respectively, in which case one may pass to the limit for k' tending to infinity without further restriction on the support of f. $\square$

For the derivation of the coupling time estimate it is easier to work with the (continuous) interpolated processes $(\hat{\overline{\Xi}}^{k,(x,y)}_{\cdot})_k$ as approximation of a suitable limit $\overline{\Xi}^{(x,y)}_{\cdot}$. Therefore we need the following

Corollary 2.1. *The sequence of processes* $\hat{\overline{\Xi}}^{k,(x,y)}_{\cdot}$ *is tight. For any subsequence* k' *the sequence of measures* $(\hat{\overline{\Xi}}^{k',(x,y)}_{\cdot})_* P$ *on* $D_{\mathbb{R}_+}(M \times M)$ *is weakly convergent if and only if* $(\tilde{\overline{\Xi}}^{k',(x,y)}_{\cdot})_* P$ *is, in which case the limits coincide. In particular the family* $((\hat{\overline{\Xi}}^{k,(x,y)}_{\cdot})_* P)_k$ *is weakly precompact and any weak accumulation point is a solution of the (restricted) coupling martingale problem, which is supported by* $C_{\mathbb{R}_+}(M \times M)$.

Proof. By construction of $(\hat{\overline{\Xi}}^{k,(x,y)}_{\cdot})$ and $(\tilde{\overline{\Xi}}^{k,(x,y)}_{\cdot})$ we have $\hat{\overline{\Xi}}^{k,(x,y)}_t = \tilde{\overline{\Xi}}^{k,(x,y)}_{\frac{1}{k}\tau_k(s)}$ for all $t \geq 0$, $k \in \mathbb{N}$, i.e.

$$(\tilde{\overline{\Xi}}^{k,(x,y)}_{\cdot}) = (\hat{\overline{\Xi}}^{k,(x,y)}_{\cdot}) \circ \Theta_k$$

with the random time transformation $\Theta_k(s,\omega) = \frac{1}{k}\tau_k(s,\omega)$. Moreover, every process $(\hat{\overline{\Xi}}^{k,(x,y)}_{\cdot})$ has continuous paths, so that

$$\mathrm{supp}\left(\underset{k' \to \infty}{\text{w-}\lim}(\hat{\overline{\Xi}}^{k',(x,y)}_{\cdot})_* P\right) \subset C_{\mathbb{R}_+}(M \times M)$$

for every possible weak limit of a converging subsequence $(\hat{\overline{\Xi}}^{k',(x,y)}_{\cdot})_* P$ and since the sequence Θ_k converges to $\mathrm{Id}_{\mathbb{R}_+}$ weakly, the continuity argument in section 17. of [Bil68] can be applied, giving

$$\underset{k' \to \infty}{\text{w-}\lim}(\tilde{\overline{\Xi}}^{k',(x,y)}_{\cdot})_* P = \underset{k' \to \infty}{\text{w-}\lim}(\hat{\overline{\Xi}}^{k',(x,y)}_{\cdot})_* P \tag{2.8}$$

for that specific subsequence, i.e. we have shown thatfor any subsequence $k' \to \infty$

$$\left((\hat{\overline{\Xi}}^{k',(x,y)}_{\cdot}) \Rightarrow \nu\right) \Longrightarrow \left((\tilde{\overline{\Xi}}^{k',(x,y)}_{\cdot}) \Rightarrow \nu\right).$$

To prove the other implication note first that (2.6) implies the almost sure continuity of the coordinate process $\pi_{\cdot}$ w.r.t. any weak limit of $(\hat{\overline{\Xi}}^{k,(x,y)}_{\cdot})_* P$. We may also write

$$(\tilde{\overline{\Xi}}^{k,(x,y)}_{\cdot}) \circ \overline{\Theta}_k = (\hat{\overline{\Xi}}^{k,(x,y)}_{\cdot}) \circ \Theta_k(\overline{\Theta}_k)$$

where $\overline{\Theta}_k(s,\omega) = \inf\{t \geq 0 \,|\, t > \Theta_k(s,\omega)\}$ is the (right continuous) generalized upper inverse of Θ_k which converges to $\mathrm{Id}_{\mathbb{R}_+}$ weakly, too. Since $\left\|\Theta_k(\overline{\Theta}_k) - \mathrm{Id}_{\mathbb{R}_+}\right\|_\infty \leq \frac{1}{k}$ and $(\hat{\Xi}^{k,(x,y)}_{\cdot})$ is a continuous process, it is easy to see that $(\hat{\Xi}^{k,(x,y)}_{\cdot})$ is tight on $C_{\mathbb{R}_+}(M \times M)$ if and only if $(\hat{\Xi}^{k,(x,y)}_{\cdot}) \circ \Theta_k(\overline{\Theta}_k)$ is tight on $D_{\mathbb{R}_+}(M \times M)$ and thus we may argue as before. Finally, the compactness itself comes from theorem 2.1 as well as as the characterization of any limit as a solution to the (restricted) coupling martingale problem via equation (2.8). □

2.3 The Coupling Probability and Gradient Estimates

The curvature condition on the Riemannian manifold (M,g) enters our probabilistic proof of gradient estimates through the following lemma, in which we confine ourselves to the only nontrivial case of strictly negative lower (sectional) curvature bounds.

Lemma 2.6. *Let (M^d,g) be a smooth Riemannian manifold with $\mathrm{Sec}\,M \geq -k$, $k > 0$ and let x, $y \in M$, $x \neq y$, be joined by a unit speed geodesic $\gamma = \gamma_{xy}$. Then for any ξ, $\eta \in S^{d-1}_x \subset T_xM$*

$$d(\exp_x(t\xi), \exp_y(t/\!\!/_\gamma \eta)) \leq |xy| + t\langle \eta - \xi, \dot\gamma(0)\rangle_{T_xM}$$
$$+\frac{1}{2}t^2\frac{\sqrt{k}}{s_k(|xy|])}\big((|\xi^\perp|^2 + |\eta^\perp|^2)c_k(|xy|) - 2\langle \eta^\perp, \xi^\perp\rangle_{T_xM}\big) + o_\gamma(t^2) \tag{2.9}$$

where $/\!\!/_\gamma$ denotes parallel translation on (M,g) along γ and $\xi^\perp$, $\eta^\perp$ denote the normal (w.r.t. $\dot\gamma(0)$) part of ξ and η respectively. In particular if $\xi^\perp = \eta^\perp$ and $\xi^\parallel = -\eta^\parallel$

$$d(\exp_x(t\xi), \exp_y(t/\!\!/_\gamma \eta)) \leq |xy| - 2t\langle \xi, \dot\gamma(0)\rangle_{T_xM} + t^2\sqrt{k}|\xi^\perp|^2 + o_\gamma(t^2). \tag{2.10}$$

Proof. If y is not conjugate to x along γ then the assertion of the lemma is derived by the standard argument: for sufficiently small t there is a unique geodesic γ_t close to γ connecting $\exp_x(t\xi)$ with $\exp_y(t/\!\!/_\gamma \eta)$ and which depends smoothly on t in a neighborhood of zero. In particular $\gamma_t \to \gamma$ for $t \to 0$ and the first and second variation formula for the arclength of γ_t yield

$$L(\gamma_t) = |xy| + t\langle \eta - \xi, \dot\gamma(0)\rangle_{T_xM} + \frac{1}{2}\mathrm{I}_\gamma(J^\perp_{\xi,\eta}, J^\perp_{\xi,\eta})t^2 + o_\gamma(t^2).$$

Here $J^\perp_{\xi,\eta}$ is the (w.r.t. $\dot\gamma$) normal part of the Jacobi-field induced from the geodesic variation $c(t,s) = \gamma_t(s)$ and I_γ is the index form of (M,g) along γ. Choose some parallel o.n. basis $(e_i)_{i=1,\dots,d}$, $e_1 = \dot\gamma$ along γ and define the vector field $V^\perp_{\xi,\eta}$ along γ by

$$V^\perp_{\xi,\eta}(s) = \sum_{i=2}^d \left(\xi_i c_k(t) + \frac{\eta_i - \xi_i c_k(|xy|)}{s_k(|xy|)} s_k(t)\right) e_i(s) \tag{2.11}$$

then we have that $J^\perp_{\xi,\eta}(0) = J^\perp_{\xi,\eta}(0)$ and $J^\perp_{\xi,\eta}(|xy|) = J^\perp_{\xi,\eta}(|xy|)$ such that the index lemma gives $\mathrm{I}_\gamma(J^\perp_{\xi,\eta}, J^\perp_{\xi,\eta}) \le \mathrm{I}_\gamma(V^\perp_{\xi,\eta}, V^\perp_{\xi,\eta})$. By the definition of the index form and the curvature assumption one finds that $\mathrm{I}_\gamma(V^\perp_{\xi,\eta}, V^\perp_{\xi,\eta}) \le \mathrm{I}_{\overline{\gamma}}(\overline{V}^\perp_{\xi,\eta}, \overline{V}^\perp_{\xi,\eta})$, where $\overline{\gamma}$ is a unit speed geodesic of arclength $|xy|$ in in the model space $\mathbb{M}_{d,k}$ and $\overline{V}^\perp_{\xi,\eta}$ is a vector field over $\overline{\gamma}$ defined by (2.11) with e_i replaced by an analogous parallel o.n. basis $\overline{e}_i$ along $\overline{\gamma}$. But since in this case $\overline{V}^\perp_{\xi,\eta}$ is a Jacobi field in $\mathbb{M}_{d,k}$ over $\overline{\gamma}$ one finds

$$
\begin{aligned}
\mathrm{I}_{\overline{\gamma}}(\overline{V}^\perp_{\xi,\eta}, \overline{V}^\perp_{\xi,\eta}) &= \langle \overline{V}^\perp_{\xi,\eta}, \nabla_{\overline{\gamma}} \overline{V}^\perp_{\xi,\eta} \rangle \Big|_0^{|xy|} \\
&= \frac{\sqrt{k}}{s_k(|xy|)} \sum_{i=2}^d \left[(\eta_i^2 + \xi_1^2) c_k(|xy|) + \eta_i \xi_i \big(s_k^2(|xy|) - c_k^2(|xy|) - 1 \big) \right] \\
&= \frac{\sqrt{k}}{s_k(|xy|)} \big((|\xi^\perp|^2 + |\eta^\perp|^2) c_k(|xy|) - 2\langle \eta^\perp, \xi^\perp \rangle_{T_xM} \big)
\end{aligned}
$$

In the case that y is conjugate to x along γ we may find a partition $0 = s_0 < s_1 < \cdots < s_n = |xy|$ such that $x_{i+1} = \gamma(s_{i+1})$ is not conjugate to $x_i = \gamma(s_i)$ along $\gamma_i = \gamma_{|[s_i, s_{i+1}]}$. Let $V^\perp_{\xi,\eta}$ be the vector field along γ be defined by (2.11) as before and for each i define $x^i_t = \exp_{x_i}(tV^\perp_{\xi,\eta}(s_i))$ and for sufficiently small t let γ^i_t be the unique geodesic close to γ_i connecting x^i_t with x^{i+1}_t which give rise to a Jacobi-field $J^i_{\xi,\eta}$ on each γ_i. Since by construction we have that $J^{i,\perp}_{\xi,\eta}$ and $V^\perp_{\xi,\eta}$ coincide at the endpoints of each interval $[s_i, s_{i+1}]$ we may apply the same arguments as before to each of the segments γ_i from which we deduce

$$
\begin{aligned}
d(\exp_x(t\xi), \exp_y(t/\!\!/_\gamma \eta)) &\le \sum_{i=0}^{n-1} d(x^i_t, x^{i+1}_t) \\
&\le \sum_{i=0}^{n-1} d(x^i, x^{i+1}) + t\langle \eta - \xi, \dot\gamma(0) \rangle_{T_xM} + \frac{1}{2} t^2 \sum_{i=0}^{n-1} \mathrm{I}_{\gamma_i}(J^{i,\perp}_{\xi,\eta}, J^{i,\perp}_{\xi,\eta}) + o_{n,\gamma}(t^2) \\
&\le |xy| + t\langle \eta - \xi, \dot\gamma(0) \rangle_{T_xM} + \frac{1}{2} t^2 \sum_{i=0}^{n-1} \mathrm{I}_{\overline{\gamma}_i}(\overline{V}^{i,\perp}_{\xi,\eta}, \overline{V}^{i,\perp}_{\xi,\eta}) + o_{n,\gamma}(t^2)
\end{aligned}
$$

with $\overline{\gamma}$ and $\overline{V}^\perp_{\xi,\eta}$ as before, the upper (or lower) index i meaning its restriction to the interval $[s_i, s_{i+1}]$ and thus

$$
= |xy| + t\langle \eta - \xi, \dot\gamma(0) \rangle_{T_xM} + \frac{1}{2} t^2 \mathrm{I}_{\overline{\gamma}}(\overline{V}^\perp_{\xi,\eta}, \overline{V}^\perp_{\xi,\eta}) + o_{n,\gamma}(t^2)
$$

from which one obtains the assertion of the lemma as in the previous case. The second statement of the lemma follows from (2.9) and the fact that $\cosh(t) - 1 \le \sinh(t)$ for all nonnegative t. □

Remark 2.3. The proof of lemma 2.6 shows that the error term $o_\gamma(t^2)$ in (2.9) may be replaced by a uniform error estimate $o(t^2)$ as long as $x \neq y$ range over a compact $K \subset M \times M \setminus D(M)$.

Proof. By the smoothness assumption on (M,g) the injectivity radius of (M,g) is locally uniformly bounded away from zero and hence for x, $y \in M$ sufficiently close and t sufficiently small the geodesic variation γ_t of $\gamma_0 = \gamma_{xy}$ which connects $x_t = \exp_x(t\xi)$ with $y_t = \exp_y(t\eta)$ is uniquely given by $\gamma_t(s) = \exp_{x_t}(s \log_{x_t} y_t)$, which is smooth in t, $x, y \in M$ and $\xi \in T_xM$, $\eta \in T_yM$ for small t, for x, y sufficiently close and ξ, η bounded. Since the arclength functional $L : C^\infty([0,1], M) \to \mathbb{R}_+$ is also smooth we find that for any smooth local frame map $\phi \in \Gamma(M, O(M))$ the function

$$\mathcal{L} : [0,\epsilon) \times M \times M \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+ \quad \mathcal{L}(t,x,y,u,v) = L(\gamma_t)$$

with γ_t as above, where $\xi = \phi_x(u)$ and $\eta = \phi_y(v)$, is smooth in some neighbourhood of $\{0\} \times \{x = y\} \times \{(0,0)\}$. From this it follows that in fact the error term in (2.9) is uniform (for $\xi, \eta \in T_xM, \|\xi\| = \|\eta\| = 1$) as long as x and y range over a compact set sufficiently close to the diagonal in $M \times M$. If x and y range over a general compact set $K \subset M \times M$ we repeat the subpartitioning argument of the previous lemma along some geodesic γ_{xy}. It follows then from the locally uniform lower bound for the injectivity radius that the total number of necessary subdivisions of γ_{xy} is uniformly bounded for $x, y \in K$ which establishes the claim. □

In order to apply the previous lemma to the the sequence $(\hat{\Xi}^{l,(x,y)}_{.})_l$ we require in addition to condition $(*)$ on page 6 that $\Phi(.,.)$ satisfies

$$\Phi_2(u,v) \circ \Phi_1^{-1}(u,v) = /\!\!/_{\gamma_{uv}} \text{ on } (T_uM)^{\perp_{\gamma_{xy}}} \quad \forall\,(u,v) \in M \times M \setminus D(M) \qquad (**)$$

where γ_{uv} corresponds to some (w.r.t. u, $v \in M \times M \setminus D(M)$ symmetric) choice of connecting unit speed geodesics. A function $\Phi(.,.)$ satisfying $(*)$ and $(**)$ realizes the *coupling by reflection* method (cf. [Ken86, Cra91]) in our present context. In order to see that we actually may find at least one such map $\Phi(.,.)$ which is also measurable we may proceed similarly as in the proof of lemma 2.1: from a given measurable and symmetric choice $\gamma_{..} : M \times M \to C^1([0,1], M)$ and a continuous frame $\psi \in \Gamma(O(M))$ we obtain $\Phi_1(x,y)$ by an appropriate rotation of $\psi(x)$ such that $\Phi_1(x,y)e_1 = \dot\gamma_{xy}/\|\dot\gamma_{xy}\|$. $\Phi_2(x,y)$ is then obtained from $\Phi_1(x,y)$ by parallel transport and reflection w.r.t. the direction of γ_{xy}. Since these operations depend continuously (w.r.t. to the C^1-norm) on the curve γ_{xy}, $\Phi(.,.)$ inherits the measurability and continuity properties of the map $\gamma_{..} : M \times M \to C^1([0,1], M)$. For $\delta > 0$ let us introduce the functional $T_{D,\delta}$ on the path space of $M \times M$ by

$$T_{D,\delta} : C_{\mathbb{R}_+}(M \times M) \to \mathbb{R} \cup \{\infty\}, \quad T_{D,\delta}(\omega) = \inf\{s \geq 0 \,|\, d(\omega_s^1, \omega_s^2) \leq \delta\}.$$

with $\omega_s^i = \pi_i(\omega_s)$, $i = 1,2$ being the projections of the path ω onto the factors. Then the *coupling time* $T_D = T_{D,0}$ is the first hitting time of the diagonal $D(M) \subset M \times M$.

Theorem 2.3 (Coupling Probabilty Estimate). *Let* $Sec\, M \geq -k,\ , k > 0$ *and let* Φ *be chosen as above. Then for arbitrary* $x, y \in M$ *and any weak limit* $P^{\infty,(x,y)} = w\text{-}\lim_{l'\to\infty} (\hat{\bar{\Xi}}^{l',(x,y)}_{\cdot})_* P$ *on* $C_{\mathbb{R}_+}(M \times M)$ *the following estimate holds true:*

$$P^{\infty,(x,y)}(T_D = \infty) \leq \frac{d-1}{2}\sqrt{k}\, d(x,y).$$

Proof. For $x = y$ there is obviously nothing to prove. So let $(x,y) \notin D(M)$ and assume first that M is compact. For $\delta > 0$ let

$$T^l_{D,\delta} : \Omega \to \mathbb{R} \cup \{\infty\}, \quad T^l_{D,\delta}(\omega) = \inf\{t \geq 0 \,|\, d(\hat{\bar{\Xi}}^{l,(x,y)}_s(\omega)) \leq \delta\} = T_{D,\delta} \circ (\hat{\bar{\Xi}}^{l,(x,y)}_{\cdot})(\omega)$$

be the first hitting time of the set $D_\delta = \{(x,y) \in M \times M |\, d(x,y) \leq \delta\}$ for the process $(\hat{\bar{\Xi}}^{l,(x,y)}_s)_{s\geq 0}$, where $(\Omega, \mathcal{O}, P)$ is the initial probability space on which the random i.i.d. sequences $(\xi)_{i\in\mathbb{N}}$ (and $(\eta_i)_{i\in\mathbb{N}}$) are defined. By the choice of Φ for $(u,v) \in D(M)^c$, $\theta \in S^{d-1} \subset \mathbb{R}^d$ and

$$(u^\epsilon, v^\epsilon) = (\exp_u(\epsilon\Phi_1\theta), \exp_v(\epsilon\Phi_2\theta))$$

we obtain from lemma 2.6 and remark 2.3

$$d(u^\epsilon, v^\epsilon) \leq d(u,v) - 2\epsilon\lambda + \sqrt{k}\epsilon^2\chi + o(\epsilon^2)$$

where $\lambda = \mathrm{pr}_1\theta$ is the projection of θ onto the first coordinate axis and $\chi = \left\|\theta^\perp\right\|^2_{\mathbb{R}^d}$ is the squared length of the orthogonal part of θ. This estimate inserted into the inductive definition of $(\hat{\bar{\Xi}}^{l,(x,y)}_t)_{t\geq 0}$ yields in the case $\hat{\bar{\Xi}}^{\frac{1}{\sqrt{l}},(x,y)}_{\lfloor lt\rfloor} \in D(M)^c$

$$\begin{aligned} d(\hat{\bar{\Xi}}^{l,(x,y)}_t) = d(\hat{\bar{\Xi}}^{\frac{1}{\sqrt{l}},(x,y)}_{lt}) &\leq d(\hat{\bar{\Xi}}^{\frac{1}{\sqrt{l}},(x,y)}_{\lfloor lt\rfloor}) - 2\left(\frac{lt - \lfloor lt\rfloor}{\sqrt{l}}\right)\lambda_{\lfloor lt\rfloor+1} \\ &+ \sqrt{k}\left(\frac{lt - \lfloor lt\rfloor}{\sqrt{l}}\right)^2 \chi_{\lfloor lt\rfloor+1} + o\left[\left(\frac{lt - \lfloor lt\rfloor}{\sqrt{l}}\right)^2\right] \end{aligned}$$

with the random variables $\lambda_i = \mathrm{pr}_1\xi_i$ and $\chi_i = \left\|\xi_i^\perp\right\|^2_{\mathbb{R}^d}$, from which one deduces by iteration

$$\begin{aligned} &\leq d(x,y) - 2\frac{1}{\sqrt{l}}\sum_{i=0}^{\lfloor lt\rfloor}\lambda_{i+1} - 2\left(\frac{lt - \lfloor lt\rfloor}{\sqrt{l}}\right)\lambda_{\lfloor lt\rfloor+1} + \sqrt{k}\frac{1}{l}\sum_{i=0}^{\lfloor lt\rfloor}\chi_{i+1} \\ &\quad + \sqrt{k}\left(\frac{lt - \lfloor lt\rfloor}{\sqrt{l}}\right)^2\chi_{\lfloor lt\rfloor+1} + \lfloor lt\rfloor o\left(\frac{lt - \lfloor lt\rfloor}{\sqrt{l}}\right)^2 \\ &= d(x,y) - 2S^l_t + \sqrt{k}\frac{1}{l}\sum_{i=0}^{\lfloor lt\rfloor}\chi_{i+1} + \rho_t(l) =: r^l(t) \end{aligned} \tag{2.12}$$

at least on the set $\{T^l_{D,\delta} > t\}$, with

$$S^l_t := \frac{1}{\sqrt{l}}S_{lt},\ S_t := S_{\lfloor t\rfloor} + (t - \lfloor t\rfloor)S_{\lfloor t\rfloor+1},\ S_k := \sum_{i=0}^{k}\lambda_i$$

and

$$\rho_t(l) \to 0 \text{ for } l \to \infty.$$

Define furthermore the stopping times

$$T_\delta^l : \Omega \to \mathbb{R} \cup \{\infty\}, \quad T_N^l = \inf\{t \geq 0 \,|\, r_t^l \leq \delta\}$$

then the inequality above implies $\{T_{D,\delta}^l > n\} \subset \{T_\delta^l > n\}$ for all l, $n \in \mathbb{N}$ and hence

$$\mathbb{P}_{P,\Omega}(\{T_\delta^l > n\}) \geq \mathbb{P}_{P,\Omega}(\{T_{D,\delta}^l > n\}) = \int\limits_{\{T_{D,\delta} > n\}} (\hat{\bar{\Xi}}_{\cdot}^{l,(x,y)})_*(P)(d\omega) \tag{2.13}$$

where the second integral is taken on a subset of the path space $\Omega' = C(\mathbb{R}_+, M \times M)$ with respect to the image measure of P under $(\hat{\bar{\Xi}}_{\cdot}^{l,(x,y)})$. By assumption we have $P^{\infty,(x,y)} = \text{w-}\lim_{l' \to \infty} (\hat{\bar{\Xi}}_{\cdot}^{l',(x,y)})_* P$ and the lower semi-continuity of the function $T_{D,\delta}$ w.r.t. to the topology of locally uniform convergence on the path space implies that the set $\{T_{D,\delta} > n\} \subset C(\mathbb{R}_+, M \times M)$ is open. Thus from (2.13) it follows that

$$\mathbb{P}_{P^{\infty,(x,y)}}(\{T_{D,\delta} > n\}) = \int\limits_{\{T_{D,\delta} > n\}} (\tilde{\bar{\Xi}}_{\cdot}^{\infty,(x,y)})_*(P)(d\omega)$$

$$\leq \liminf_{l' \to \infty} \int\limits_{\{T_{D,\delta} > n\}} (\hat{\bar{\Xi}}_{\cdot}^{l',(x,y)})_*(P)(d\omega) \leq \liminf_{l' \to \infty} \mathbb{P}_{P,\Omega}(\{T_\delta^{l'} > n\}) = \mathbb{P}(\{T_\delta(r^\infty) > n\}).$$

The last equality is a consequence of Donsker's invariance principle applied to the sequence of processes $(r_\cdot^l)_{l \in \mathbb{N}}$: since each $(\lambda_i)_i$ and $(\chi_i)_i$ are independent sequences of i.i.d. random variables on $\{\Omega, P, \mathcal{A}\}$ with

$$\mathbb{E}(\lambda_i) = 0, \quad \mathbb{E}(\lambda_i^2) = \sqrt{\frac{1}{d}}, \quad \mathbb{E}(\chi_i) = \frac{d-1}{d}$$

one finds that $(r_\cdot^l)_l$ converges weakly to the process $r_\cdot^\infty$ with

$$r_t^\infty = d(x, y) + \frac{2}{\sqrt{d}} b_t + \sqrt{k} \frac{d-1}{d} t \tag{2.14}$$

such that in particular $\mathbb{P}_{P,\Omega}(\{T_\delta(r^\infty) = n\}) = 0$ and we can pass to the limit in the last term on the right hand side of (2.13). Letting n tend to infinity leads to

$$\mathbb{P}_{P^{\infty,(x,y)},\Omega'}(\{T_{D,\delta} = \infty\}) \leq \mathbb{P}_{P,\Omega}(\{T_\delta(r^\infty) = \infty\}),$$

where $\delta > 0$ was chosen arbitrarily from which we finally may conclude

$$\mathbb{P}_{P^{\infty,(x,y)},\Omega'}(\{T_D = \infty\}) \leq \mathbb{P}_{P,\Omega}(\{T_0(r^\infty) = \infty\}). \tag{2.15}$$

with T_0 being the first hitting time of the origin for the semi-martingale r^∞. Using a Girsanov transformation of $(\Omega, \mathcal{O}, P)$ ([KS88], sec. 3.5.C) the probability on the right hand side can be computed precisely to be

$$\mathbb{P}_{P,\Omega}(\{T_0(r^\infty) = \infty\}) = 1 - e^{-\frac{1}{2}\sqrt{k}(d-1)d(x,y)} \leq \frac{d-1}{2}\sqrt{k}\, d(x,y),$$

which is the claim in the compact case. For noncompact M we choose some open precompact $A \subset M$ such that (x, y) in K. We may stop the processes $(\hat{\Xi}_{\cdot}^{l',(x,y)})$ when they leave A and repeat the previous arguments for the the stopping time $T_{A,D,\delta} = T_{D,\delta} \wedge T_{A^c}$ with $T_{A^c} = \inf\{s \geq 0 \,|\, \omega_s \in A^c\}$ which gives instead of (2.13)

$$\mathbb{P}_{P,\Omega}(\{T_\delta^l \wedge T_{A^c} > n\}) \geq \int\limits_{\{T_{D,\delta} \wedge T_{A^c} > n\}} (\hat{\Xi}_{\cdot}^{l,(x,y)})_*(P)(d\omega)$$

From this we obtain (2.15) if we successively let tend $l \to \infty$, $A \to M \times M$, $\delta \to 0$ and $n \to \infty$. □

Remark 2.4. In the proof of theorem 2.3 we did not use that $P^{\infty,(x,y)}$ is (necessarily) a solution of the restricted coupling martingale problem. In fact, theorem 2.1 is not involved at all at this stage, except that it guarantees the existence of a weakly converging subsequence $(\hat{\Xi}_{\cdot}^{l',(x,y)})_* P$.

Remark 2.5. Also in the case of lower Ricci curvature bounds the same type of arguments should yield the extension of theorem 2.3. However, the difficulties arise from the fact that lower Ricci bounds lead to a uniform upper estimate of the expectation of the χ_i in (2.12) only. Since these random variables are also only asymptotically mutually independent, one has to find and apply an appropriate central limit theorem to the expression $\frac{1}{l}\sum_{i=0}^{\lfloor lt \rfloor} \chi_{i+1}$ in order to obtain the pathwise(!) upper bound for the distance process by the semimartingale (2.14).

For different (local and global) versions of the following result as well as for extensions to harmonic maps see the papers by W. Kendall. We state the global version only because it is easiest to formulate whereas its stochastic proof received most scepticism due to non-differentiability of the distance function on the cut locus.

Corollary 2.2 (Gradient estimate for harmonic functions). *If u is a harmonic, nonnegative and uniformly bounded function on M, then*

$$|u(x) - u(y)| \leq \|u\|_\infty (d-1)\sqrt{k}\, d(x,y) \tag{2.16}$$

Proof. From elliptic regularity theory we now that $u \in C^\infty(M)$. Let $x \neq y$ be given. Since $\Delta u = 0$ we find $L_c(u \otimes 1) = L_c(1 \otimes u) = 0$ and from theorem 2.1 it follows that both processes $((u \otimes 1)(\pi_s))_s$ and $((1 \otimes u)(\pi_s))_s$ are nonnegative continuous bounded martingales with respect to the probability measure $P^{\infty,(x,y)}$, where $\pi_{\cdot} = (\pi_1, \pi_2)_{\cdot}$ is the projection process on the path space $C_{\mathbb{R}_+}(M \times M)$. For any $s > 0$ we obtain by means of the optional stopping theorem

$$\begin{aligned} u(x) - u(y) &= (u \otimes 1)(\pi_0) - (1 \otimes u)(\pi_0) \\ &= \mathbb{E}_{P^{\infty,(x,y)}}[(u \otimes 1)(\pi_{s\wedge T_D}) - (1 \otimes u)(\pi_{s\wedge T_D})] \end{aligned}$$

which equals, since $(u \otimes 1)(\pi_{T_D}) = (1 \otimes u)(\pi_{T_D})$ on $\{T_D < \infty\}$

$$\begin{aligned} &= \mathbb{E}_{P^{\infty,(x,y)}}[\big((u \otimes 1)(\pi_{s\wedge T_D}) - (1 \otimes u)(\pi_{s\wedge T_D})\big)\mathbb{1}_{\{s<T_D\}}] \\ &\leq 2\,\|u\|_\infty\, \mathbb{P}_{P^{\infty,(x,y)}}(\{T_D \geq s\}) \end{aligned}$$

Passing to the limit for $s \to \infty$ and using theorem 2.3 proves the claim. □

2.4 Extension to Riemannian Polyhedra

As indicated in the introduction the stochastic proof of gradient estimates given above suggests to be extended to certain non-smooth spaces since it depends essentially only on two facts, namely the central limit theorem for the factor processes $(\pi_i(\Xi^k_{\cdot}))$ and the asymptotic quadruple estimate (2.10). Therefore one might think of situations where the same kind of argument applies.

Definition 2.1. *Let X be an n-dimensional topolgical manifold equipped with a complete metric d. We call (X, d) an n-dimensional Riemannian polyhedron with lower curvature bound $\kappa \in \mathbb{R}$ if $X = \bigcup_i P_i$ is obtained by gluing together (in a locally finite fashion) convex closed patches $P_i \subset M_i^n$ $(i \in I)$ of n-dimensional Riemannian manifolds with uniform lower sectional curvature bound κ along their boundaries, where*

i) the boundary $\partial P_i = \bigcup_j S_{ij} \subset M_i$ of each $P_i \subset M_i$ is the union of totally geodesic hypersurfaces S_{ij} in M_i

ii) each $S_{ij} \subset X$ is contained in the intersection of at most two $P_k \subset X$ and $S_{ij} \subset M_i$ where $S_{kl} \subset M_k$ are isometric whenever two adjacent patches $P_i \subset X$ and $P_k \subset X$ have a common face $S_{ij} \simeq S_{kl} \subset (X, d)$

iii) the sum of the dihedral angles for each face of codimension 2 is less or equal 2π.

Examples 2.1. The boundary ∂K of a convex Euclidean polyhedron $K \subset \mathbb{R}^n$ (with nonempty interior $\mathring{K}$) is a (n-1)-dimensional Riemannian polyhedron with lower curvature bound 0 in our sense. A simple example for the case $\kappa < 0$ in two dimensions is the surface of revolution obtained from a concave function $f : [a, b] \to \mathbb{R}_+,\ f \in C^1[a, b] \cap C^2([a, c) \cup (c, b])$ with $c \in (a, b)$ and

$$f'(c) = 0 \text{ and } f''/f = \begin{cases} -k_1^2 & \text{on } [a, c) \\ -k_2^2 & \text{on } (c, b]. \end{cases}$$

Constructions: The 2π-condition iii) above assures that (X, d) is an Alexandrov space with $\mathrm{Curv}(X) \geq \kappa$ (see section 3), and we can use the result in [Pet98] that a geodesic segment connecting two arbitrary metrically regular points does not hit a metrically singular point, i.e. a point whose tangent cone is not the full Euclidean space. Moreover, from condition ii) it follows in particular that metrically singular points can occur only inside the $(n-2)$-skeleton X^{n-2} of X. Thus there is a natural parallel translation along any geodesic segment γ_{xy} whenever x and y are regular and which is obtained piecewisely from the parallel translation on the Riemannian patches P_i and from the natural gluing of the tangent half-spaces for points $x \in X^{n-1} \setminus X^{n-2}$ lying on the intersection of two adjacent (n-1)-faces $S_{ij} \simeq S_{kl} \subset X$. Similarly we can define the exponential map $\exp_x$ for every regular point $x \in X$, i.e. for given $\xi \in T_xX$ we obtain a unique ("quasi-geodesic") curve $\mathbb{R}_+ \ni t \to \exp_x(t\xi)$ (and which can be represented as a union of geodesic segments on the patches P_i). With these constructions at our disposal we can verify a non-smooth version of the asymptotic quadrangle estimate of lemma 2.6:

Proposition 2.1. *Let (X, d) be a n-dimensional Riemannian polyhedron with lower curvature bound $\kappa \in \mathbb{R}$ and let $x, y \in X \setminus X^{n-2}$ be connected by some segment γ_{xy}. Then for $\xi \in T_xX$, $\|\xi\| = 1$ the estimate (2.10) holds, where the error term $o(t^2)$ can be chosen uniform if $x \neq y$ range over a compact subset of $X \setminus X^{n-2}$.*

Proof. Let us prove (2.10) for fixed $x, y \in X \setminus X^{n-2}$ and $\xi \in T_xX$ first, i.e. without addressing the problem of uniformity. Suppose furthermore that for some P_i we have $\gamma_{xy} \subset P_i$, i.e. γ_{xy} is entirely contained in the (closed) patch P_i, then we distinguish three cases:

i) If $\gamma_{xy} \subset \mathring{P}_i$ then due to lemma 2.6 there is nothing left to prove.

ii) $x \in \mathring{P}_i$ and $y \in P_i \cap P_j$ for some j. Since y is assumed to be regular the first order part of estimate (2.10) is obviously true and we may focus on the second order part which corresponds to orthogonal variations of the geodesic γ, i.e. we may assume that $\xi = \xi^\perp$ in (2.10). If γ_{xy} is orthogonal to the hypersurface $\partial P_i \cap \partial P_j$ at y or x and y are both in $P_i \cap P_j$ then again there is nothing to prove since we in this case we have to consider geodesic variations which take place completely on one of the patches $P_i \subset M_i$ or $P_j \subset M_j$ and we can apply lemma 2.6 on M_i or M_j respectively.

   Consequently we only have to treat the case that γ_{xy} is neither parallel nor orthogonal to $\partial P_i \cap \partial P_j$, i.e. $0 < \langle \frac{\dot\gamma_{xy}(|xy|)}{\|\dot\gamma_{xy}(|xy|)\|}, \nu \rangle_{T_yM_i} < 1$ where ν denotes the outward unit normal vector of ∂P_i.

Let $\eta = /\!\!/_{\gamma_{xy}} \xi$ be the parallel translate of a unit vector $\xi \in T_xM_i$ normal to $\dot\gamma_{xy}$. Then $\zeta = \eta - \frac{\langle \nu,\eta\rangle}{\langle \nu, \dot\gamma_{xy}(|xy|)\rangle}\dot\gamma_{xy}(|xy|) \in T_yM_i \cap T_yM_j$ is the unique vector in the intersection of the $\{\dot\gamma_{xy}(|xy|), \eta\}$-plane and the tangent hyperplane to ∂P_i in y which is determined by its w.r.t. $\dot\gamma_{xy}$ orthogonal projection η. For its length we obtain $\|\zeta\| = \sin^{-1}\alpha$ where α is the angle enclosed by $\dot\gamma_{xy}$ and η at y. Since $\partial P_i \subset M_i$ and $\partial P_j \subset M_j$ are totally geodesic the point $z = \exp_y(t\zeta)$ also lies on $\partial P_i \cap \partial P_j$ and the triangle inequality yields

$$d_X(\exp_x(t\xi), \exp_y(t /\!\!/_{\gamma_{xy}} \xi)) \le d_{M_i}(\exp_x(t\xi), z) + d_{M_j}(z, \exp_y(t\eta)) \tag{2.17}$$

where d_X, d_{M_i} and d_{M_j} denote the distance functions on X, M_i and M_j respectively. Now the estimate (2.9) of lemma 2.6 applied to ξ and ζ in M_i yields

$$d_{M_i}(\exp_x(t\xi), \exp_y(t\zeta)) \le d_{M_i}(x,y) - t\frac{\cos\alpha}{\sin\alpha} + t^2\sqrt{k}|\xi^\perp|^2 + o_\gamma(t^2)$$

since trivially $\langle \dot\gamma, \zeta\rangle = \frac{\cos\alpha}{\sin\alpha}$ and by construction $\zeta^\perp = \eta = /\!\!/_{\gamma_{xy}}\xi$. As for the distance $d_{M_j}(z, \exp_y(t\eta))$ remember that by the smoothness assumption of the curvature of M_j is locally uniformly bounded and from the Toponogov triangle comparison and the cosine formula on the model spaces $\mathbb{M}_{d,\kappa}$ we may infer with $\beta = \sphericalangle_{T_yM_j}(\zeta, \eta)$

$$d_{M_j}(z, \exp_y(t\eta) = t\sqrt{|\zeta|^2 + 1 - |\zeta|\cos\beta} + o(t^2) = t\frac{\cos\alpha}{\sin\alpha} + o(t^2)$$

because all vectors $\dot\gamma_{xy}$, η and ζ lie on a common hyperplane and as $\eta \perp \dot\gamma$ we have $\alpha = \pi/2 - \beta$. Inserting the the last two inequalities into (2.17) yields (2.10).

iii) $x \in P_i \cap P_k$ and $y \in P_i \cap P_j$. We may argue similarly as in ii) by subdividing the quadruple into two geodesic triangles on M_j and M_k and a remaining quadruple on M_i. - Alternatively, if $z \in \gamma_{xy} \cap \mathring{P}_i \neq \emptyset$ then one may subdivide $\gamma_{xy} = \gamma_{xz} * \gamma_{zy}$ and argue as in ii). If $\gamma_{xy} \cap \mathring{P}_i = \emptyset$, then again we have to deal with variations of γ_{xy} on a single Riemannian patch P_k only, where P_k depends on the direction $\xi \perp \dot\gamma_{xy}$.

The discussion above proves (2.10) when $\gamma_{xy} \subset P_i$ for some P_i. In the general case when γ_{xy} is not contained in a single patch we subdivide $\gamma_{xy} = \gamma_{i_1} * \gamma_{i_2} * \cdots * \gamma_{i_m}$ into pieces $\gamma_{i_k} \subset P_{i_k}$ lying entirely on one of the patches which we consider separately: let $\{x_1, \ldots, x_m\} = \gamma_{xy} \cap X^{n-1}$ be the set of (transversal) intersections of γ_{xy} and X^{n-1} and for each $k = 1, \ldots, m$, $t > 0$ let $z_t^k = \exp_{x_k}(tz_k)$ where the direction z^k (depending on the initial direction $\xi \in T_xX$) is chosen as in ii). As before the triangle inequality yields the simple upper bound $d(x_t, y_t) \le d(x_t, z_t^1) + d(z_t^1, z_t^2) + \cdots + d(z_t^{m-1}, z_t^m) + d(z_t^m, y_t)$ for the distance between $x_t = \exp_x(\xi_t)$ and $y_t = \exp_y(t /\!\!/_{\gamma_{xy}}\eta)$. On each patch P_{i_k} we may apply the previous discussion i) - iii) in order to derive asymptotic estimates for $d(x_t, z_t^1)$, $d(z_t^1, z_t^2)$, $\ldots$, $d(z_t^{m-1}, z_t^m)$ and $d(z_t^m, y_t)$, where it is important to note that for sufficiently small t the variations η_t^k of the pieces of γ_{i_k} which we construct on each patch P_{i_k} also lie entirely on P_{i_k}. (This follows from the fact that the segment γ_{xy} lies at a strictly positive distance away from X^{n-2} which comprises the set of points where more than just two patches intersect.) Hence we obtain an upper bound for the distance $d(x_t, y_t)$ in the global quadruple by a sum of distances $d(x_t, z_t^1)$, $d(z_t^1, z_t^2)$, $d(z_t^{m-1}, z_t^m)$ and $d(z_t^m, y_t)$ in geodesic quadrangles and triangles which are each entirely contained in a single patch. Analogously to the final step in ii) summing up the corresponding asymptotic upper estimates we recover (2.10) for the global quadruple due to the special choice of the directions $\{z^k | k \in 1, \dots, m\}$.

Finally, the uniformity assertion is obtained in a similar way by combining the arguments in lemma 2.3 on each patch P_i with the observation that for a given compact set $K \subset X \setminus X^{n-2}$ the collection of all segments $\{\gamma_{xy} |\, (x, y) \in K \times K\}$ also lies at a strictly positive distance away from X^{n-2}, which may be inferred from a simple compactness consideration. This implies that there is some $t_0 > 0$ such that for all $t \leq t_0$ and $x, y \in K$ all variations $\gamma_{xy,t}$ constructed in the previous paragraph determine a well-defined sequence of geodesic triangles and quadrangles located on the individual patches as above. Hence, by the smoothness of the patches (and the fact that only finitely many patches are involved for x, $y \in K \times K$) we may conclude precisely as in lemma 2.3 that the estimate (2.10) is in fact locally uniform in the sense stated above. □

As a second preparation for the probabilistic approach to a gradient estimate on (X, d) we need to state precisely what we understand by a Brownian motion in the present situation.

Definition 2.2. *The ('Dirichlet-')Laplacian Δ^X on (X, d) is defined as the generator of the Dirichlet form $(\mathcal{E}_c, D(\mathcal{E})_c)$ which is obtained as the $L^2(X, dm = \sum_i dm_{i\lfloor P_i})$-closure of the classical energy form $\mathcal{E}(f, f) = \sum_i \int_{P_i} |\nabla f|^2\, dm_i$ on the set of Lipschitz functions on (X, d) with compact support. A continuous Hunt process whose transition semigroup coincides with the semigroup associated to $(\mathcal{E}_c, D(\mathcal{E}_c))$ on $L^2(X, dm)$ is called a Brownian motion on (X, d).*

Lemma 2.7. *Let $\mathcal{A}$ denote the set $\bigcap_i C^\infty(\mathring{P}_i) \cap Lip(X) \cap C_c(X \setminus X^{n-2})$ of piecewise smooth Lipschitz functions with compact support in $X \setminus X^{n-2}$ satisfying the gluing condition for the the normal derivatives on adjacent $(n-1)$-dimensional faces*

$$\sum_{\partial P_i \cap \partial P_j \neq \emptyset} \frac{\partial}{\partial \nu_j} f = 0 \ \text{on}\ \partial P_i \cap \partial P_j \qquad (+)$$

and let $\Delta^X_\circ$ denote the Laplace operator defined on $\mathcal{A}$ by $\mathring{\Delta}^X f(x) = \sum_i \Delta^{P_i}(f_{|P_i})(x) 1\!\!1_{P_i}(x)$. Then $\Delta^X_\circ$ is essentially self adjoint on $L^2(X, dm)$.

Proof. We adopt the proof for smooth case (cf. [Dav89], thm. 5.2.3) with small alterations which are caused by the exceptional set X^{n-2}. Since $-\Delta^X_\circ$ is non-negative it suffices to prove that the range of $\mathrm{Id} - \Delta^X_\circ$ is dense (cf. [RS75], thm. X.26). So suppose $u \in$

$L^2(X, dm)$ is orthogonal to $(\mathrm{Id} - \Delta_\circ^X)\mathcal{A}$, i.e.

$$\langle \Delta_\circ^X \psi, u\rangle_{L^2(X,dm)} = \langle \psi, u\rangle_{L^2(X,dm)} \quad \forall\, \psi \in \mathcal{A}. \tag{2.18}$$

This means that $\Delta u = u$ weakly on each P_i and using elliptic regularity theory we find that is u is continuous on X and smooth on each P_i up to the boundary except in the corners $P_i \cap X^{n-2}$ which comprises the set where ∂P_i is not smooth. Equation (2.18) also implies that u satisfies the gluing condition (+) such that we can integrate by parts in order to obtain

$$\langle \nabla\psi, \nabla u\rangle_{L^2(X,dm)} = \sum_i \langle \nabla^{P_i}\psi, \nabla^{P_i} u\rangle_{L^2(P_i)} = -\langle \psi, u\rangle_{L^2(X,dm)} \tag{2.19}$$

which holds true for all $\psi \in \mathcal{A}$ and consequently also for all functions $\psi \in W_0^{1,2}(X \setminus X^{n-2})$ where we define $W_0^{1,2}(X \setminus X^{n-2})$ as the closure of $\mathcal{A}$ with respect to the norm $\|\psi\|_{1,2}^2 = \|\psi\|^2_{L^2(X,dm)} + \sum_i \left\|\nabla^{P_i}\psi_{|P_i}\right\|^2_{L^2(P_i)}$. But for $\eta \in \mathrm{Lip}(X) \cap C_c(X \setminus X^{n-2})$ the function ηu belongs to $W_0^{1,2}(X \setminus X^{n-2})$ (note that the condition (+) is void on $W_0^{1,2}(X \setminus X^{n-2})$) such that ψ in (2.19) may be replaced by ηu.

In a second step we would like to get rid of the condition that η vanishes on X^{n-2}. For this recall that $\dim_H X^{n-2} \le n-2$, hence for each compact $K \subset X$ we can construct a sequence of cut-off functions $(\rho_k)_k \in \mathrm{Lip}(X) \cap C_c(X \setminus X^{n-2})$ such that $\rho_k \to 1$ dm-a.e. on K and $\nabla\rho_k \to 0$ in $L^2(K, dm)$ (cf. the very general construction given in [KMS01] which can be applied in the present context). Consequently, for arbitrary $\eta \in \mathrm{Lip}(X) \cap C_c(X)$ we may construct a sequence $(\rho_k)_k$ of cut-off functions relative to $K = \mathrm{supp}(\eta)$ as above in order to obtain a (uniformly bounded) sequence $\eta_k = \rho_k\eta \in \mathrm{Lip}(X) \cap C_c(X \setminus X^{n-2})$ such that $\eta_k \to \eta$ $dm-$a.e. and $\nabla\eta_k \to \nabla\eta$ in $L^2(X, dm)$. Inserting $\psi = \eta_k^2 u$ in (2.19) and using Leibniz' rule for the gradient and Cauchy-Schwarz inequality we obtain

$$\begin{aligned}\int_X \eta_k^2 |\nabla u|^2 dm &= -\int_X \eta_k^2 u^2 dm - 2\int_X \eta_k \nabla\eta_k u \nabla u dm \\ &\le 2\, \|\eta_k|\nabla u|\|^{\frac{1}{2}}_{L^2(X)}\, \|u|\nabla\eta_k|\|^{\frac{1}{2}}_{L^2(X)},\end{aligned}$$

i.e. $\int_X \eta_k^2|\nabla u|^2 dm \le 4\int_X u^2|\nabla\eta_k|^2 dm$ which yields $\int_X \eta^2|\nabla u|^2 dm \le 4\int_X u^2|\nabla\eta|^2 dm < \infty$ by letting tend k to infinity, i.e. $\nabla u \in L^2_{\mathrm{loc}}(X, dm)$. Hence we see that if we set $\psi = \rho_k\eta$ in (2.19)

$$\langle \rho_k\nabla(\eta u), \nabla u\rangle_{L^2(X,dm)} + \langle \eta u\nabla\rho_k, \nabla u\rangle_{L^2(X,dm))} = \langle \rho_k\eta u, u\rangle_{L^2(X,dm)}$$

we may in fact pass to the limit for $k \to \infty$ in order to prove that

$$\langle \nabla(\eta u), \nabla u\rangle_{L^2(X,dm)} = \langle \eta u, u\rangle_{L^2(X,dm)} \quad \forall\eta \in \mathrm{Lip}(X) \cap C_c(X). \tag{2.20}$$

The rest of the argument is identical to the proof in [Dav89], which we recapitulate for the reader's comfort. Take a smooth function $\psi : \mathbb{R} \to [0,1]$ with $\psi_{|[0,1]} = 1$, $\psi_{|[2,\infty)} = 0$

and let $\phi_n : X \to [0,1]$ be given by

$$\phi_n(x) = \psi(d(x,y)/n)$$

where $d = d_X$ is the distance function on (X,d) and $y \in X$ is some fixed point. Then ϕ_n is a feasible test function for (2.20) and moreover, since $|d_X(x,y) - d_X(z,y)| \le d_X(x,z) \le d_{P_i}(z,y)$ for $z, y \in P_i$ and

$$|\nabla d(.,y)|(x) \le \limsup_{z \to x} \left| \frac{d_X(x,y) - d_X(z,y)}{d_{P_i}(x,z)} \right| \le 1 \quad \forall x \in \mathring{P}_i$$

the bound

$$\|\nabla \phi_n\|_{L^\infty} = \|\psi'(d(y,.)/n)\nabla d(y,.)\|_{L^\infty} \le \|\psi'\|_\infty /n$$

obviously holds true. Inserting ϕ_n^2 for η in (2.20) we get

$$0 \ge -\langle \phi_n^2 u, u\rangle = \langle \nabla(\phi_n^2 u), \nabla u\rangle = \langle \phi_n^2 \nabla u, \nabla u\rangle + 2\langle u\phi_n \nabla\phi_n, \nabla u\rangle$$

which yields

$$\int_X \phi_n^2 |\nabla u|^2 dm \le 4 \int_X u^2 |\nabla \phi_n|^2 dm \le \frac{4}{n} \|\psi\|_\infty \int_X u^2 dm$$

and thus finally $\nabla u = 0$ a.e. by sending $n \to \infty$, which contradicts (2.19) unless $u = 0$. □

The last preparation concerns the construction of the coupling process, where further singularities of the coupling map $\Phi(.,.)$ with $\Phi(x,y) : \mathbb{R}^n \times \mathbb{R}^n \to K_xX \times K_yX$ may be caused by the existence of non-Euclidean tangent cones K_x if $x \in X^{n-2}$. However, choosing beforehand a map $\Psi(,)$ on $X \times X^{n-2} \cup X^{n-2} \times X \cup \{(x,x) \,|\, x \in X\}$ with $\Psi(x,y) : \mathbb{R}^n \times \mathbb{R}^n \to K_xX \times K_yX$ (not necessarily isometric) and depending measurably on (x,y) we can find a globally defined measurable coupling map $\Phi(.,.)$ extending $\Psi(.)$ and satisfying $(*)$ and $(**)$ on $X \times X \setminus (X \times X^{n-2} \cup X^{n-2} \times X)$. This can be proved by slightly modifying the arguments of lemma 2.1.

Hence we have everything we need to define a sequence of coupled (quasi-)geodesic random walks on $X \times X$ from which we obtain as before the sequences $(\overset{\sim}{\overline{\Xi}}{}^{k,(x,y)}_{.})_k$ and $(\overset{\wedge}{\overline{\Xi}}{}^{k,(x,y)}_{.})_k$ by scaling.

Proposition 2.2. *For any $(x,y) \in X \times X$ the sequences $(\overset{\sim}{\overline{\Xi}}{}^{k,(x,y)}_{.})_k$ and $(\overset{\wedge}{\overline{\Xi}}{}^{k,(x,y)}_{.})_k$ are tight on $D_{\mathbb{R}_+}(X \times X)$ and $C_0(\mathbb{R}_+, X \times X)$ respectively. For any subsequence k' the sequence of measures $(\overset{\wedge}{\overline{\Xi}}{}^{k',(x,y)}_{.})_* P$ on $D_{\mathbb{R}_+}(X \times X)$ is weakly convergent if and only if $(\overset{\sim}{\overline{\Xi}}{}^{k',(x,y)}_{.})_* P$ is, in which case the limits coincide. Under any weak limit $P^\infty_{(x,y)} = w\text{-}\lim_{k' \to \infty} (\overset{\wedge}{\overline{\Xi}}{}^{k',(x,y)}_{.})_* P$ the marginal processes $\pi^1_.$ and $\pi^2_.$ are Brownian motions on (X,d) starting in x and y respectively.*

Proof. The tightness assertion and the coincidence of the limits of any jointly converging subsequences $(\overset{\sim}{\Xi}_{\cdot}^{k',(x,y)})_{k'}$ and $(\overset{\wedge}{\Xi}_{\cdot}^{k',(x,y)})_{k'}$ is proved precisely in the same manner as in the smooth case. It just remains to identify the limit of the marginal processes. Due to lemma 2.7 above we know that the generator of $(\mathcal{E}_c, D(\mathcal{E}_c))$ is given as the closure of the Laplacian $\Delta_\circ^X = \sum_i \mathbb{1}_{P_i}\Delta^{M_i}$ on the set $\mathcal{A}$ of piecewise smooth Lipschitz functions with compact support on X, vanishing on X^{n-2} and satisfying (+), i.e. $\mathcal{A}$ is a core for the generator of $(\mathcal{E}_c, D(\mathcal{E}_c))$. Now for $u \in \mathcal{A}$ it is obvious that the sequence $A_k u$ converges locally uniformly and hence also in $L^2(X, dm)$ to $\Delta_\circ^X u$, where the operators A_k are defined by scaled tangential mean values as in (2.1). Since the $(A_k)_k$ are the generators for the marginals of $(\overset{\sim}{\Xi}_{\cdot}^{k,(x,y)})_k$ we may infer from Kurtz' semigroup approximation theorem (cf. [EK86], thm. 1.6.1) that also the associated semigroups $(P_t^k)_k$ converge strongly in $L^2(X, dm)$ to the semigroup P_t which is generated by the closure of $\Delta_\circ^X$, i.e. the semigroup associated to $(\mathcal{E}_c, D(\mathcal{E}_c))$. □

Lemma 2.8. *Let $u \in L^\infty(X, dm) \cap \mathcal{D}(\mathcal{E}_c)$ weakly harmonic on (X, d), i.e. $\mathcal{E}(u, \xi) = 0$ for all $\xi \in \mathcal{D}(\mathcal{E}_c)$, and let $P_{(x,y)}^\infty = w\text{-}\lim_{k'\to\infty}(\overset{\wedge}{\Xi}_{\cdot}^{k',(x,y)})_* P$ be a weak limit of a subsequence $(\overset{\wedge}{\Xi}_{\cdot}^{k',(x,y)})_* P$. Then under $P_{(x,y)}^\infty$ the processes $t \to u(\omega_t^1) = (u \otimes 1)(\omega_t)$ and $t \to u(\omega_t^2) = (1 \otimes u)(\omega_t)$ are martingales with respect to to the canonical filtration $(\overline{\mathcal{F}}_t = \sigma\{\pi_s^i \,|\, s \le t,\, i = 1, 2\})_{t\ge 0}$ on $C_{\mathbb{R}_+}(X \times X)$.*

Proof. Note that due to elliptic regularity theory we have that $u \in C^0(X) \cap C^\infty(\overset{\circ}{X}{}^{n-1})$ and that u satisfies the gluing condition (+). Hence we find that $A_k(u) \to 0$ locally uniformly on $X \setminus X^{n-2}$, where A_k is approximate Laplacian operator (2.1). We would like to use this property when we pass to the limit for $k' \to \infty$. It remains to justify this limit. Let us call for short $\nu = P_{(x,y)}^\infty = \text{w-}\lim_{k'\to\infty}(\overset{\wedge}{\Xi}_{\cdot}^{k',(x,y)})_* P$ for a suitable subsequence k' and $\hat{\nu}^k = (\overset{\wedge}{\Xi}_{\cdot}^{k,(x,y)})_* P$.

For $\rho > 0$ we may find some open neighbourhood $C_\rho \subset X$ of X^{n-2} satisfying

i) $\overline{B_{\rho/2}(X^{n-2})} \subset C_\rho \subset \overline{C_\rho} \subset B_\rho(X^{n-2})$

ii) ∂C_ρ intersects X^{n-1} transversally and

iii) $\partial C_\rho \cap X \setminus X^{n-1}$ is smooth.

Let $T_\rho^i = \inf\{t \ge 0 | \omega_t^1 \in \overline{C_\rho}\}$ for $i = 1, 2$ the hitting time for the marginals of $\overline{C_\rho}$ and let $D_i = \{\omega \in C_{\mathbb{R}_+}(X \times X) \,|\, T_\rho^i$ is not continuous in $\omega\}$, then $\nu(D_i) = 0$ for $i = 1, 2$. This is seen as follows: since the hitting time of a closed set $C \subset X$ is lower semi-continuous on $C_{\mathbb{R}_+}(X)$ for each $\omega \in D_i$ we necessarily have $T_\rho^i(\omega) < \infty$ and it exists a sequence $\omega^\epsilon \to \omega^i \in C_{\mathbb{R}_+}(X)$ such that $T_\rho^i(\omega^i) + \delta < \liminf_\epsilon T_\rho^i(\omega^\epsilon)$ for some $\delta > 0$. Note that by condition ii) on C_ρ the set $X^{n-1} \cap \partial C_\rho$ has (Hausdorff-)dimension $\le n - 2$

and hence is polar for Brownian motion on (X, d) and that by proposition 2.2 under ν the marginal processes are Brownian motions on X. Thus $T^i_\rho(\omega) < \infty$ implies (ν-almost surely) $\omega^i_{T^i_\rho(\omega)} \in \partial C_\rho \cap X \setminus X^{n-1}$. But then $T^i_\rho(\omega^i) + \epsilon < \liminf_\epsilon T^i_\rho(\omega^\epsilon)$ implies the existence of some $\epsilon_0 > 0$ such that $\omega^i_{T^i_\rho(\omega)+\epsilon'} \notin C_\rho$ for all $\epsilon' \le \epsilon_0$. Using the strong Markov property of the marginal processes under ν and the regularity of $\partial C_\rho \cap X \setminus X^{n-1}$ we finally deduce that the set of such paths has indeed vanishing ν-measure.

On account of $\hat{\nu}^{k'} \Rightarrow \nu$ and the ν-almost sure continuity of functional $\Sigma_\rho : D_{\mathbb{R}_+}(X \times X) \to D_{\mathbb{R}_+}(X \times X)$, $(\Sigma_\rho \omega)(t) = \omega_{t \wedge T^1_\rho(\omega) \wedge T^2_\rho(\omega)}$ we find (by thm. 5.1. of [Bil68]) that $(\Sigma_\rho)_* \hat{\nu}^{k'} \Rightarrow \nu_\rho := (\Sigma_\rho)_* \nu$ for $k' \to \infty$. Set $\overline{T}_\rho = T^1_\rho \wedge T^2_\rho$, then the Markov property of $\hat{\overline{\Xi}}^{k,(x,y)}_{.}$ and the optional sampling theorem yield that for all $t \ge s_l \ge \ldots s_1 \ge 0$ and $v, g_1, \ldots g_l \in C_b(X \times X)$

$$\left\langle v(\omega_t) - v(\omega_0) - \int_0^t \overline{A}_k v(\omega_s) ds, g_1(\omega_{s_1}) \ldots g_l(\omega_{s_l}) \right\rangle_{\hat{\nu}^k_\rho}$$

$$= \left\langle v(\omega_{t \wedge \overline{T}_\rho}) - v(\omega_0) - \int_0^{t \wedge \overline{T}_\rho} \overline{A}_k v(\omega_s) ds, g_1(\omega_{s_1 \wedge \overline{T}_\rho}) \ldots g_l(\omega_{s_l \wedge \overline{T}_\rho}) \right\rangle_{\hat{\nu}^k}$$

$$= \left\langle v(\omega_{s_l \wedge \overline{T}_\rho}) - v(\omega_0) - \int_0^{s_l \wedge \overline{T}_\rho} \overline{A}_k v(\omega_s) ds, g_1(\omega_{s_1 \wedge \overline{T}_\rho}) \ldots g_l(\omega_{s_l \wedge \overline{T}_\rho}) \right\rangle_{\hat{\nu}^k}$$

$$= \left\langle v(\omega_{s_l}) - v(\omega_0) - \int_0^{s_l} \overline{A}_k v(\omega_s) ds, g_1(\omega_{s_1}) \ldots g_l(\omega_{s_l}) \right\rangle_{\hat{\nu}^k_\rho}, \tag{2.21}$$

where $\overline{A}_k$ is the generator of $\hat{\overline{\Xi}}^{k,(x,y)}_{.}$. If we put $v = u \otimes 1$ for u as above and use the fact that

$$\overline{A}_k(u \otimes 1) = (A_k u) \otimes 1 \to 0 \text{ uniformly on } X \setminus B_\rho(X^{n-2})$$

we see that for each $t \ge s_l$ the sequence of functionals $(T^k_t)_k$ defined by

$$T^k_t : D_{\mathbb{R}_+}((X \setminus B_\rho(X^{n-2})) \times (X \setminus B_\rho(X^{n-2})) \to \mathbb{R}$$

$$T^k_t(\omega) = (u \otimes 1)(\omega_t) - (u \otimes 1)(\omega_0) - \int_0^t \overline{A}_k(u \otimes 1)(\omega_s) ds g_1(\omega_{s_1}) \ldots g_l(\omega_{s_l})$$

converges uniformly on compacts $K \subset D_{\mathbb{R}_+}((X \setminus B_{\rho/2}(X^{n-2})) \times (X \setminus B_{\rho/2}(X^{n-2}))$ to the functional

$$T_t(\omega) = (u(\omega^1_t) - u(\omega^1_0)) g_1(\omega_{s_1}) \ldots g_l(\omega_{s_l}).$$

defined on $D_{\mathbb{R}_+}(X \times X) \supset D_{\mathbb{R}_+}((X \setminus B_{\rho/2}(X^{n-2})) \times (X \setminus B_{\rho/2}(X^{n-2}))$. Hence, due to $\hat{\nu}^{k'}_\rho \Rightarrow \nu_\rho$ we may pass to the limit in (2.21) for $k' \to \infty$ giving

$$\left\langle (u(\omega^1_t) - u(\omega^1_0)) g_1(\omega_{s_1}) \ldots g_l(\omega_{s_l}) \right\rangle_{\nu_\rho} = \left\langle (u(\omega^1_{s_l}) - u(\omega^1_0)) g_1(\omega_{s_1}) \ldots g_l(\omega_{s_l}) \right\rangle_{\nu_\rho}. \tag{2.22}$$

Moreover, by definition of ν_ρ (2.22) is equivalent to

$$\begin{aligned}&\Big\langle (u(\omega^1_{t\wedge \overline{T}_\rho}) - u(\omega^1_0))g_1(\omega_{s_1\wedge \overline{T}_\rho})\dots g_l(\omega_{s_l\wedge \overline{T}_\rho})\Big\rangle_\nu \\ &= \Big\langle (u(\omega^1_{s_l\wedge \overline{T}_\rho}) - u(\omega^1_0))g_1(\omega_{s_1\wedge \overline{T}_\rho})\dots g_l(\omega_{s_l\wedge \overline{T}_\rho})\Big\rangle_\nu . \end{aligned} \tag{2.23}$$

Finally, using again that under ν the marginal processes are Brownian motions on X and X^{n-2} is polar we deduce $\overline{T}_\rho \geq T^1_{B_\rho(X^{n-2})} \wedge T^2_{B_\rho(X^{n-2})} \to \infty$ for $\rho \to 0$ ν-almost surely, such that taking the limit for $\rho \to 0$ in (2.23) yields

$$\big\langle (u(\omega^1_t) - u(\omega^1_0))g_1(\omega_{s_1})\dots g_l(\omega_{s_l})\big\rangle_\nu = \big\langle (u(\omega^1_{s_l}) - u(\omega^1_0))g_1(\omega_{s_1})\dots g_l(\omega_{s_l})\big\rangle_\nu ,$$

which amounts to the statement that the process $t \to u(\omega^1_t)$ is a $((\overline{\mathcal{F}}_t),\nu)$-martingale. □

Proposition 2.3. *Let (X,d) be an n-dimensional Riemannian polyhedron with lower sectional curvature bound κ and let for arbitrary x, $y \in X$ the measure $P^{\infty,(x,y)} = w\text{-}\lim_{l'\to\infty}(\hat{\overline{\Xi}}^{l',(x,y)}_{\cdot})_* P$ on $C_{\mathbb{R}_+}(X\times X)$ be a weak limit of some suitably chosen subsequence k'. Then the coupling probability estimate holds true as in the smooth case, i.e.*

$$P^{\infty,(x,y)}(T_D = \infty) \leq \frac{n-1}{2}\sqrt{k}\, d(x,y).$$

Proof. We give only a sketch, because there is essentially nothing new in the arguments. Using proposition 2.1 we may proceed as in the proof of theorem 2.3 if we restrict of the discussion onto the set of paths stopped by the stopping time $\overline{T}\rho$ for $\rho > 0$. In analogy to the proof of lemma 2.8 the final step is to send $\rho \to 0$ which then yields the claim. □

Hence with optional sampling from lemma 2.8 and proposition 2.3 we may conclude

Theorem 2.4. *Let (X,d) be a d-dimensional generalized locally finite Riemannian polyhedron with lower sectional curvature bound κ, then any bounded and weakly harmonic function $u \in \mathcal{D}(\mathcal{E}_c)$ on (X,d) satisfies the gradient estimate (2.16).*

Remark 2.6. It should be noted that the estimate (2.16) is different from Yau's original estimate (4.19) because it yields an upper bound for the supremum of the gradient by the supremum of the function itself, whereas (4.19) is a bound on the supremum of the logarithmic derivative of a nonnegative harmonic function. In two dimensions our analytic proof of the stronger estimate (4.19) given in section four covers the case of generalized two-dimensional Riemannian polyhedra.

3 Alexandrov Spaces

Basic Concepts

In this section we give a short overview of the most fundamental concepts of Alexandrov spaces which generalize Riemannian manifolds with lower sectional curvature bound. The general reference is the paper [BGP92], certain parts of which are thoroughly explained in [Shi93] and [BBI01].

Most definitions of curvature bounds for metric spaces (X, d) rely on the comparison of the behaviour of the distance function d on n-tuples of points $(x_i) \in X$ with the distance function $\overline{d}$ on appropriately chosen n-tuples $(\overline{x}_i) \in \mathbb{M}_{n,k}$ from the simply connected Riemannian manifold of constant curvature $k \in \mathbb{R}$, cf. the concept of $\mathcal{K}$-curvature classes in [Gro99]. In [BGP92] four essentially equivalent definitions for geodesic metric spaces with lower curvature bounds are given, among which the four points property is probably the most general, because it does not even require completeness nor the existence of geodesics in (X, d).

Definition 3.1 (Metric space with Curvature $\geq k$). *An arbitrary metric space (X, d) is said to have curvature bounded from below, i.e. Curv$(X) \geq k$, $k \in \mathbb{R}$ iff every point $x \in X$ admits a neighborhood U_x (with $U_x \subset B_{\pi/\sqrt{k}}(x)$ if $k > 0$) such that for any quadruple of points $(a; p, q, r)$ taken from U_x the four-points property holds:*

$$\overline{\sphericalangle} paq + \overline{\sphericalangle} qar + \overline{\sphericalangle} rap \leq 2\pi. \tag{3.1}$$

Here $\overline{\sphericalangle} stv$ denotes the angle at $\overline{t}$ of a geodesic triangle $\Delta(\overline{s}, \overline{t}, \overline{v}) \subset \mathbb{M}_{2,k}$ with sidelengths given by $\overline{d}(\overline{s}, \overline{t}) = d(s, t)$, $\overline{d}(\overline{t}, \overline{v}) = d(t, v)$ and $\overline{d}(\overline{v}, \overline{s}) = d(v, s)$.

By an inner (or path metric) space we mean a metric space with the property that any two points can be joined by a rectifiable curve with arclength equal to the distance of the given points, i.e. for all x, $y \in X$ there is a (necessarily) continuous curve $\gamma_{xy} : [0, d(x, y)] \to X$ such that $L(\gamma_{xy}) = d(x, y)$, where the length of a curve $c : [a, b] \to X$ is defined by

$$L(c) = \sup \sum_{a=t_0 \leq t_1 \leq \cdots \leq t_K = b} d(c(t_{i+1}), c(t_i))$$

with the supremum being taken over all finite partitions $a = t_0 \leq t_1 \leq \cdots \leq t_K = b$. For locally compact and complete metric spaces this is equivalent to the existence of a midpoint, i.e. for all x, $y \in X$ there is a $z \in X$ such that

$$d(x, z) = d(y, z) = \frac{1}{2} d(x, y).$$

A curve joining two points x and y as above is called a geodesic segment, any curve being locally a geodesic segment is called geodesic (arc). We may assume without restriction that any geodesic is parametrized according to arclength. Finally a geodesic triangle in an inner metric space is defined by its three vertices and three geodesic segments joining each pair of them. In the sequel we are going to deal with locally compact inner metric spaces only.

Definition 3.2 (Alexandrov Space). *A locally compact and complete inner metric space (X, d) with Curv$(X) \geq k$ is called Alexandrov space with curvature bounded below by k.*

Proposition 3.1 (Alexandrov Convexity, local version). *For a locally compact inner metric space the condition Curv$(X) \geq k$ is equivalent to the Alexandrov convexity of geodesic hinges: for each $x \in X$ there is an $\epsilon > 0$ such that for any pair of geodesics γ, η with $\gamma(0) = \eta(0) = x$ the function*

$$\theta(s,t) = \overline{\sphericalangle}(\gamma(s)x\eta(t)) \text{ is non-increasing for } s,\ t \leq \epsilon.$$

This formulation is very important because it permits the introduction of the following crucial concepts.

Definition 3.3 (Angles and Tangent Cones). *Let γ, η be two geodesics emanating from $x \in X$ as above then*

$$\sphericalangle(\gamma,\eta) := \lim_{s,t\to 0} \theta(s,t) =: d_{\sphericalangle}(\gamma,\eta)$$

defines the angle (or angular distance) between η and γ. The space of directions $(\Sigma_x, d_{\sphericalangle})$ is the (closure of the) set of equivalence classes of all geodesics emanating from x with respect to the angular distance $d_{\sphericalangle}$

$$\Sigma_x = \left(\{\gamma_{xy} | y \in X\}/_{\sim_{d_{\sphericalangle}}}, d_{\sphericalangle}\right)^{\sim}.$$

The tangent cone (K_x, d_x) is the topological cone $(\Sigma_x \times \mathbb{R}_+)/(\Sigma_x \times \{0\})$ over Σ_x equipped with the metric induced on $\mathbb{R}_+ \times \Sigma_x$ by the Euclidean cosine law

$$d_x^2[(\alpha,s),(\beta,t)] = s^2 + t^2 - 2st\cos d_{\sphericalangle}(\alpha,\beta).$$

Remark 3.1. In fact, in the definition above it would be better to talk of the angular distance between γ and η only, because $\sphericalangle(\gamma,\eta)$ as defined above yields the shorter of the two angular rotations from γ to η and from η to γ respectively, i.e. it does not distinguish between 'inner' or 'outer' angle.

Generally, for a geodesic triangle $\Delta(p,q,r) \subset X$, which is defined by its vertices and arbitrary but fixed geodesic segments pq, qr and $rp \subset X$, let $\overline{\Delta}(p,q,r) = \Delta(\overline{p},\overline{q},\overline{r})$ denote the (up to congruence) uniquely defined geodesic triangle in $\mathbb{M}_{2,k}$ such that $\overline{d}(\overline{p},\overline{q}) = d(p,q)$, $\overline{d}(\overline{q},\overline{r}) = d(q,r)$ and $\overline{d}(\overline{r},\overline{p}) = d(r,p)$, which will be called comparison triangle for

$\Delta(p,q,r)$. If $k \leq 0$ one can always find such a triangle, in the case $k > 0$ one has to impose for its circumference $d(p,q) + d(r,q) + d(q,p) \leq 2\pi/\sqrt{k}$. Having introduced the notion of angles between geodesics one can now state two more equivalent properties (cf. [BGP92]):

Proposition 3.2. *For a locally compact inner metric space the following properties are equivalent:*

i) Curv$(X) \geq k$ (in the sense of definition 3.1)

ii) for every geodesic triangle $\Delta(p,q,r) \subset U_x$ the angles are bounded above by the corresponding angles of the comparison triangle $\overline{\Delta}(p,q,r) \subset \mathbb{M}_{2,k}$

iii) for every geodesic triangle $\Delta(p,q,r) \subset U_x$ and a corresponding comparison triangle $\overline{\Delta}(p,q,r) \subset \mathbb{M}_{2,k}$ the inequality

$$d(p, m_{qr}) \geq \overline{d}(\overline{p}, \overline{m}_{\overline{qr}})$$

holds, where $m_{qr} \in pq$ and $\overline{m}_{\overline{qr}} \in \overline{qr}$ are the midpoints of the geodesics rq and $\overline{rq}$ in X and $\mathbb{M}_{2,k}$ respectively.

The last statement means that every geodesic triangle $\Delta \subset U_x \subset X$ in such a neighborhood of any given point in $x \in X$ is *more convex* than the corresponding comparison triangle $\overline{\Delta} \subset \mathbb{M}_{2,k}$, where the *convexity* of a triangle is measured by the distances from its vertices to interior points on the opposing side. Finally, all previous essentially equivalent local conditions can be turned into global ones ([BGP92], thm. 3.2.):

Theorem 3.1 (Globalization Theorem). *If (X,d) is a complete space with Curv$(X) \geq k$, then inequality (3.1) remains true for any geodesic triangle in (X,d).*

Remark 3.2. In the case $k > 0$ the assumption about the circumference of triangles is superfluous since it will be automatically satisfied. Also the diameter of X is bounded from above by $\pi/\sqrt{k}$.

Examples 3.1. ([BGP92, Shi93])

i) Riemannian manifolds with sectional curvature bounded from below and quotients M/Γ of such Riemannian manifolds by groups Γ acting isometrically (not necessarily free or discrete)

ii) simplicial n-dimensional Riemannian complexes (obtained from gluing together Riemannian simplexes of constant curvature k) which satisfy the 2π-gluing-condition along the faces of codimension 2

iii) boundaries of convex subsets in Riemannian manifolds with lower sectional curvature bound, and as a special case surfaces of revolution $\subset \mathbb{R}^3$ obtained from graphs of convex functions

iv) spherical suspensions or cones of Alexandrov spaces, for instance $C(\mathbb{R}P^n)$, the Euclidean cone over the real projective space of dimension n

v) spaces obtained by gluing two Alexandrov spaces along their boundaries if the boundaries are intrinsically isometric ([Pet97])

vi) Hausdorff-limits of Riemannian manifolds with uniform sectional curvature bound. - For this recall the definition of the Hausdorff-distance of two subsets $A,\ B \subset (X,d)$ of a metric space

$$d_H^X(A,B) = \inf\{\epsilon > 0 \,|\, A \subset U_\epsilon(B),\ B \subset U_\epsilon(A)\}.$$

Now the Hausdorff-distance[§] of two metric spaces (A,d_A), (B,d_B) is defined by

$$d_H(A,B) = \inf d_H^X(f(A),g(B))$$

where the infimum is taken over all metric spaces (X,d) and all isometric embeddings $f : A \to X$ and $g : B \to X$. Then d_H induces a complete metric between compact metric spaces (cf. [Gro99]), i.e. the space $\mathcal{X}_c$ of (equivalence classes of) compact metric spaces together with the function d_H forms a complete (and contractible) not locally compact metric space. By Gromov's compactness theorem for each choice of $n \in \mathbb{N}$, $\kappa \in \mathbb{R}$, $D \in \mathbb{R}_+$ the set $\mathcal{M}(n,\kappa,D)$ of n-dimensional Riemannian manifolds with lower Ricci curvature bound $(n-1)\kappa$ and diameter less than D is precompact in $(\mathcal{X}_c, d_H)$. Moreover, if one assumes also the sectional curvature to be bounded below by $k \in \mathbb{R}$ then any Hausdorff limit of a converging sequence will be an Alexandrov space with curvature bounded below by k (cf. [BBI01]).

Dimension and Regularity

In contrast to singular spaces with upper curvature bounds Alexandrov spaces exhibit a great amount of regularity and even carry a natural weak Riemannian structure. Firstly, each Alexandrov space of finite rough dimension[¶] has in fact integer Hausdorff dimension, which then coincides with its topological dimension. The proof of this fact is based upon

[§] The Lipschitz distance $d_L(A,B)$ between two metric spaces (A,d_A) and (B,d_B) is defined by $d_L(A,B) = \inf\{|\log \mathrm{dil} f| + |\log \mathrm{dil} f^{-1}| | f : A \to B \text{ bi-Lip homeom.}\}$. For compact spaces the convergence with respect to the Hausdorff-distance is essentially equivalent to the Lipschitz convergence of corresponding ϵ-nets for arbitrary $\epsilon > 0$, cf. [Gro99, Pet86].

[¶] The rough dimension of a bounded subset $Z \subset (X,d)$ is defined as $\dim_r(Z) = \inf\{\alpha > 0 \,|\, \lim_{\epsilon\to 0} \beta_Z(\epsilon)\epsilon^\alpha = 0\}$, where $\beta_Z(\epsilon)$ is maximal number of points $\{a_i\} \in Z$ with $d(x_i,x_j) \geq \epsilon$ if $i \neq j$.

the local constancy of the rough dimension in X and the existence of a dense open set of manifold points p, which are characterized by existence of a maximal system of almost orthogonal almost geodesic curves passing through p (*strainer* at p), which in turn can be used for the construction of bi-Lipschitz homeomorphisms of a certain neighborhood $U_p \subset X$ onto an open subset $\mathbb{R}^n$ (see the appendix B for a short sketch of the ideas). In particular each finite dimensional Alexandrov space is locally compact and supports a natural nontrivial measure m_X, which is its n-dimensional Hausdorff measure when $\dim(X) = n \in \mathbb{N}$. Throughout the rest of this work we are going to deal with finite dimensional Alexandrov spaces only. For each $\delta > 0$ the system of *natural coordinate functions*

$$\phi_{\{a_i\}} : X \supset V_p \to \mathbb{R}^n \quad q \to (d(a_1, q), \dots, d(a_n, q))$$

which are induced from the (n, δ)-strainers[‖] $\{(a_{-i}, a_{+i}) \,|\, i = 1, \dots, n\}$ at $p \in X_\delta = X \setminus S_\delta$ yields a (locally bi-Lipschitz) topological atlas for X_δ for sufficiently small δ ([BGP92]).

Definition 3.4. *Let $S_\delta = \{p \in X \,|\, \nexists\, (n, \delta)$-strainer at $p\}$. Then the set $S_X = \bigcup_{\delta>0} S_\delta = \lim_{\delta\to 0} S_\delta$ is called the (metrically) singular set.*

S_X is equivalently characterized by

$$S_X = \{x \in X \,|\, \Sigma_p \not\simeq S^{n-1}\},$$

which follows from the compactness of Σ_p and a suitable sphere theorem (cf. thm. 8.9 in [Shi93]) and it can be shown that $\dim_H(S_X) \le n - 1$ ([BGP92, OS94]), but except being the countable union of closed sets no further topological information on S_X is obtained. Conversely, $X \setminus S_X = \bigcap_{0<\delta\ll 1} X_\delta$ with $X_\delta = \{x \in X \mid \mathrm{rad}(\Sigma_x) > \pi - \delta\}$[**] is the intersection of Lipschitz manifolds which are open in X. However, $X \setminus S_X$ need not be a manifold in general but it still carries a so called weak C^1-Riemannian structure[††], which is revealed by a careful analysis of the regularity properties of the natural coordinate functions.

Theorem 3.2 (Weak C^1-Riemannian Structure ([OS94])). *For properly chosen $U_{\phi_{\{a_i\}}}$ the system of maps $\big\{\phi_{\{a_i\}} : X \supset U_{\phi_{\{a_i\}}} \to \mathbb{R}^n \,|\, \{(a_{-i}, a_{+i})\}$ is a (n, δ)-strainer at p, $p \in X \setminus S_X, \delta \le 1/2n\big\}$ gives rise to a weak C^1-Riemannian structure on $X \setminus S_X$, whose induced intrinsic distance coincides with the initial distance d on $X \setminus S_X$.*

In section 5.1 it will be shown that the metrically singular points do not comprise all points in which the local behaviour of X deviates considerably from that of a smooth Riemannian manifold.

[‖] cf. appendix B for the precise definition of a (n, δ)-strainer at $p \in X$

[**] The radius $\mathrm{rad}(A)$ of a subset A in a metric space (X, d) is defined by $\mathrm{rad}(A) := \inf_{\xi\in A} \sup_{\eta\in A} d(\xi, \eta)$.

[††] cf. appendix B for a definition of a weak Riemannian structure

4 Gradient Estimates on Alexandrov Surfaces

In this section we give an analytic proof of gradient estimates for harmonic functions on two-dimensional Alexandrov spaces with Lipschitz continuous metric tensor (g_{ij}). Since the paper [CH98] contains a proof of gradient estimates for Riemannian manifolds with Lipschitz metric tensor satisfying a distributional lower curvature bound it remains for us to derive this property for Alexandrov surfaces out of the geometric curvature bounds given in terms of geodesic triangles[‡‡]. By our result we can enlarge significantly the set of examples of non-smooth spaces given in [CH98] which admit Yau's gradient estimate.

4.1 An Integral Gauss-Bonnet-Formula

Thus the main task is to overcome the technical difficulties arising from the low regularity of the metric tensor, the main step being an extension of the Gauss-Bonnet theorem to a non-smooth situation. Here we rely on the existence and semi-boundedness of the excess measure e_X on an arbitrary Alexandrov space as introduced and investigated in the paper [Mac98], see also section 6.2 for a short exposition. Let $\Box_h(x)$ denote the open Euclidean square in $\mathbb{R}^2$ of sidelength $h > 0$ with lower left corner x and assume that $\Box_h(x)$ is contained in the image $\Omega = \psi(U)$, $U \subset X$ of a natural local chart ψ of (X, d) and let $e_x(\Box_h(x))$ denote the excess measure of the set $\psi^{-1}(\Box_h(x)) \subset X$. Then we obtain the following

Proposition 4.1 (Gauss-Bonnet). *Let (X, d) be a two-dimensional Alexandrov space with Curv$(X) \geq \kappa \in \mathbb{R}$ and discrete metrically singular set S_X such that the system of natural coordinates with the associated metric tensors (g_{ij}) yield a $C^{1,1}$ Riemannian structure for $X \setminus S_X$, i.e. the chart transition maps are $C^{1,1}_{loc}$ and the metric tensors (g_{ij}) are locally uniformly Lipschitz continuous. Let $\Omega = \phi_{\{a_i\}}(U) \subset \mathbb{R}^2$ be the image of $U \subset X \setminus S_X$ under $\phi_{\{a_i\}}$ and let $g_{\{a_i\}} = (g_{ij}) : \Omega \to \mathbb{R}^{2\times 2}$, $g_{ij} \in Lip_{loc}(\Omega)$, be the Riemannian tensor on U with respect to ϕ. Then for all $h > 0$ and almost all $x \in \Omega$ with dist$(x, \partial\Omega) > h$ the set $\Box_h(x)$ is e_X-measurable and for any $\xi \in C^0_c(\Omega)$ and sufficiently small $h > 0$ the following identity applies:*

$$\begin{aligned} \int_\Omega e_X(\Box_h(x))\xi(x)dx &= \int_\Omega \xi(x)\Big(\sum_{i=1}^4 \sphericalangle_i^{\text{int}}(\Box_h(x)) - 2\pi\Big)dx \\ &- \int_\Omega\int_0^h \Big[\xi(x-te_2)\tfrac{\sqrt{\det g}}{g_{22}}(x)\Gamma^1_{22}(x) - \xi(x-he_2-te_1)\tfrac{\sqrt{\det g}}{g_{11}}(x)\Gamma^2_{11}(x)\Big]dt\,dx \qquad (4.1) \\ &+ \int_\Omega\int_0^h \Big[\xi(x-(h-t)e_2-he_1)\tfrac{\sqrt{\det g}}{g_{22}}(x)\Gamma^1_{22}(x) - \xi(x-(h-t)e_1)\tfrac{\sqrt{\det g}}{g_{11}}(x)\Gamma^2_{11}(x)\Big]dt\,dx\,. \end{aligned}$$

[‡‡]Compare the works by I. Nikolaev (cf. survey in [BNR93]) who proved that metric space spaces with two-sided curvature bounds can be approximated in Gromov-Hausdorff distance by smooth Riemannian manifolds. Here we have only lower curvature bounds but impose a priori a certain regularity on (X, d).

Remark 4.1. In the case $(g_{ij}) \in \mathrm{BV}_{\mathrm{loc}}(\Omega) \cap C^0(\Omega)$ - at least on a formal level - it would be perfectly possible to replace the measures $\Gamma^i_{jj}(x)dx$ in the formula above by the finite Radon measures $\Gamma^i_{jj}(dx)$, but unfortunately we were not able to prove (4.1) under these weaker assumptions. Because of this reason we state our main theorem of this section under the assumption that $(g_{ij}) \in \mathrm{Lip}_{\mathrm{loc}}(\Omega)$.

For the proof of proposition 4.1 some preparation is needed. At first we take a look at the smooth situation. Suppose (M, g) is a smooth (two-dimensional) Riemannian manifold and let $\Omega \subset \mathbb{R}^2$ be contained in the image of some chart $\psi : M \supset U \to \Omega \subset \mathbb{R}^2$ with associated metric tensor g_{ij}. For every $x \in \Omega$ and $v \in \mathbb{R}^2$, $s > 0$ we study the behaviour of the geodesic $\gamma^v_{x,s} = \gamma_{x,x+sv}$ on the manifold M connecting the points that correspond to the Euclidean points x and $x + sv \in \Omega$ for small s. Throughout this section we assume that every geodesic between fixed endpoints $p, q \in M$ is parameterized on the unit interval and we identify γ with its image under ψ as long as it is entirely contained in U. Then we obtain the following

Lemma 4.1. *In the smooth case for s tending to zero the asypmtotic formula*

$$\dot{\gamma}^v_{x,s}(0) = sv - \frac{s^2}{2}\Gamma_x(v, v) + o(s^2) \tag{4.2}$$

holds true locally uniformly with respect to $x \in \Omega$, $v \in \mathbb{R}^2$.

Proof. The assertion is a consequence of the geodesic equation. By definition of the geodesic $\gamma^v_{x,s}$

$$\begin{aligned}
\dot{\gamma}^v_{x,s}(\rho) - sv &= \dot{\gamma}^v_{x,s}(\rho) - (\gamma^v_{x,s}(1) - \gamma^v_{x,s}(0)) = \int_0^1 [\dot{\gamma}^v_{x,s}(\rho) - \dot{\gamma}^v_{x,s}(\tau)]d\tau \\
&= \int_0^1 \int_\tau^\rho \ddot{\gamma}^v_{x,s}(\sigma) d\sigma d\tau = \int_0^1 \int_\tau^\rho \Gamma_{\gamma^v_{x,s}(\sigma)}(\dot{\gamma}^v_{x,s}(\sigma), \dot{\gamma}^v_{x,s}(\sigma)) d\sigma d\tau \qquad (4.3) \\
&\leq Cs^2 \int_0^1 \left| \int_\tau^\rho \left\| \Gamma_{\gamma^v_{x,s}(\sigma)} \right\| d\sigma \right| d\tau \leq Cs^2 \sup_{\sigma \in [0,1]} \left\| \Gamma_{\gamma^v_{x,s}(\sigma)} \right\| \int_0^1 |\rho - \tau| d\tau = O(s^2) \qquad (4.4)
\end{aligned}$$

where we have put $\Gamma_z : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$, $\Gamma_z(\xi, \eta) = \Gamma^k_{ij}(z)\xi_i\eta_j e_k$, $z \in \Omega$ and $O(s)$ is a "big O" Landau-function. In (4.3) we used the fact that

$$\left\| \dot{\gamma}^v_{x,s} \right\|_{g_{\gamma^v_{x,s}}} = \mathrm{const} = d(x, s + sv) \leq C_g s$$

locally uniformly w.r.t. x and v. Hence $\dot{\gamma}^v_{x,s}(\rho) = sv + O(s^2)$ for all $\rho \in [0, 1]$. Plugging this back into the second integral integral of (4.3) for the case $\rho = 0$ gives

$$\dot{\gamma}^v_{x,s}(0) - sv = -\frac{s^2}{2}\Gamma_x(v, v) + s^2 \int_0^1 \int_0^\tau \left[\Gamma_x - \Gamma_{\gamma^v_{x,s}(\sigma)} \right] (v, v) d\sigma d\tau + o(s^2) \tag{4.5}$$

From the continuity of $\Gamma(.)$ it follows that $\int\limits_0^1 \int_0^\tau \left[\Gamma_x - \Gamma_{\gamma^v_{x,s}(\sigma)}\right](v,v)d\sigma d\tau = \vartheta(s)$ and since all the error estimates above hold true locally uniformly with respect to x and v the claim of the lemma is established. □

Our next goal is the derivation of (4.2) in the given non-smooth situation. Here the main device is the approximation of (g_{ij}) by smooth metrics $(g^\epsilon_{ij} = \phi_\epsilon * g_{ij})$, where ϕ_ϵ is a sequence of smooth symmetric Dirac functions.

Lemma 4.2. *The sequence $(g^\epsilon_{ij}) \in \mathbb{R}^{2\times 2}_{sym}$ has the following properties*

$$
\begin{array}{cc}
(g^\epsilon_{ij}) \stackrel{\epsilon\to 0}{\longrightarrow} (g_{ij}) & \textit{locally uniformly on } \Omega \textit{ and} \\
g^\epsilon_{ij,k} \stackrel{\epsilon\to 0}{\longrightarrow} D_{x_k} g_{ij} & \textit{strongly in } L^p_{loc}(\Omega) \textit{ for all } p \in (1,\infty) \\
\left\|g^\epsilon_{ij,k}\right\|_{C^0(K)} \leq [g_{ij}]_{Lip(K)} & \textit{for all } K \subset \Omega \textit{ compact.}
\end{array}
$$

□

With (g^ϵ_{ij}) given as above and $v \in \mathbb{R}^2$ define furthermore

$$
\begin{aligned}
\overline{\gamma}^{v,\epsilon}_{x,s}(\delta) &= \frac{1}{m^\epsilon(B_{s^3}(x+sv))} \int\limits_{B_1(0)\subset\mathbb{R}^2} \gamma^\epsilon_{x,x+s(v+s^2\theta)}(\delta)\sqrt{\det g^\epsilon_{ij}(x+s(v+s^2\theta))}d\theta \\
\overline{\gamma}^{v}_{x,s}(\delta) &= \frac{1}{m(B_{s^3}(x+sv))} \int\limits_{B_1(0)\subset\mathbb{R}^2} \gamma_{x,x+s(v+s^2\theta)}(\delta)\sqrt{\det g_{ij}(x+s(v+s^2\theta))}d\theta
\end{aligned}
\tag{4.6}
$$

where $\gamma^\epsilon_{x,x+s(v+s^2\theta)} : [0,1] \to \Omega$ and $\gamma_{x,x+s(v+s^2\theta)} : [0,1] \to \Omega$ denote geodesic segments with respect to (g^ϵ_{ij}) and (g_{ij}) respectively connecting the point x with $x+s(v+s^2\theta)$, where $\theta \in B_1(0)$ and $m^\epsilon(B_{s^3}(x+sv)) = \int_{B_{s^3}(x+sv)}(\det g^\epsilon_{ij}(y))^{\frac{1}{2}}dy$, $m(B_{s^3}(x+sv) = \int_{B_{s^3}(x+sv))}(\det g_{ij}(y))^{\frac{1}{2}}dy$. This complicated choice of $\overline{\gamma}^v_{x,s}(\delta)$ and $\overline{\gamma}^{v,\epsilon}_{x,s}(\delta)$ is motivated by the following property.

Lemma 4.3. *For all $x \in \Omega$ and $\delta \in [0,1]$ $\lim_{\epsilon\to 0}\overline{\gamma}^{v,\epsilon}_{x,s}(\delta) = \overline{\gamma}^v_{x,s}(\delta)$.*

Proof. Since $g^\epsilon_{ij} \to g_{ij}$ uniformly, we obtain that for each pair of points x, y in Ω sufficiently close the sequence of geodesics γ^ϵ_{xy} is uniformly bounded in $W^{1,\infty}(\mathbb{R}_+, U) = \{\gamma : \mathbb{R}_+ \to U \mid dx\text{-ess sup}\sqrt{\dot\gamma_i g_{ij}(\gamma)\dot\gamma_j} < \infty\}$ which implies in particular the precompactness of this sequence w.r.t. (locally) uniform convergence. Since the arclength functional $L_{(g_{ij})}$ is lower semicontinuous w.r.t. to weak convergence in $W^{1,2}([0,1],U) = \{\gamma : \mathbb{R}_+ \to U \mid \|\sqrt{\dot\gamma_i g_{ij}(\gamma)\dot\gamma_j}\|_{L^2(\mathbb{R}_+,dx)} < \infty\}$ any limit function of γ^ϵ_{xy} will be geodesic segment from x to y in X. From the fact that the cut locus $C_x = \{y \in X \mid \gamma_{xy} \text{ not unique}\}$ of every point $x \in X$ has zero Hausdorff measure in X ([OS94], prop. 3.1) we may conclude the assertion of the lemma. □

Remark 4.2. The estimate (4.4) prevails for the (g_{ij})-geodesic segments $\gamma^v_{x,s}$ in the sense that

$$\forall K \subset \Omega\ \exists C \in \mathbb{R}_+ : \quad m\text{-}\operatorname*{ess\,sup}_{\theta\in B_{s^3}(x+sv)} \|\dot\gamma_{x,\theta}(0) - sv\| \le Cs^2 \quad \forall x \in K, \tag{4.7}$$

which may be seen as follows: using the analogue of (4.4) for the (g^ϵ_{ij})-geodesics $\gamma^\epsilon_{x,\theta}$, $\theta \in B_{s^3}(x+sv)$, we obtain the corresponding estimate (4.7$^\epsilon$) for the approximating metrics (g^ϵ_{ij}), where the constant C on the right hand side can be chosen uniform with respect to ϵ. (This is a simple consequence of the uniform bound on $\|\Gamma^\epsilon(.,.)\|$). Hence we may pass to the limit in (4.7$^\epsilon$) for $\epsilon \to 0$ provided that $\dot\gamma^\epsilon_{x,\theta}(0) \to \dot\gamma_{x,\theta}(0)$. However, by the geodesic equation and by the uniform boundedness of $\|\Gamma^\epsilon(.,.)\|$ it is evident that the Lipschitz norm of $\dot\gamma^\epsilon_{x,\theta}(.)$ is uniformly bounded, which implies relative compactness of the family $(\gamma^\epsilon_{x,\theta}(.))_{\epsilon\ge 0}$ in $C^1([0,1],\mathbb{R}^2)$. Together with the m-almost sure uniqueness of the $(g_{ij})-$geodesic $\gamma_{x,\theta}$ with respect to the endpoint θ we may conclude $\dot\gamma^\epsilon_{x,\theta}(0) \to \dot\gamma_{x,\theta}(0)$ for $m-$a.e. θ such that we can in fact pass to the limit for m-almost all $\theta \in B_{s^3}(x+sv)$ which yields the assertion. - From the same compactness and uniqueness arguments we may conclude also that the function $\Omega \ni x \to \dot{\overline{\gamma}}^v_{x,s}(0)$ with

$$\dot{\overline{\gamma}}^v_{x,s}(0) := \lim_{\epsilon\to 0} \dot{\overline{\gamma}}^{\epsilon,v}_{x,s}(0) = \fint_{B_{s^3}(x+sv)} \dot\gamma_{x,\theta}(0) m(d\theta) = \lim_{\delta\to 0}\frac{1}{\delta}\left[\overline{\gamma}^v_{x,s}(\delta) - x\right] = \dot{\overline{\gamma}}^v_{x,s}(0)$$

is well defined and locally uniformly bounded , i.e. $\dot{\overline{\gamma}}^v_{.,s}(0) \in L^\infty_{\text{loc}}(\Omega)$.

Now we can formulate the key lemma for the proof of proposition 4.2.

Lemma 4.4. *Under the conditions of proposition 4.1 formula (4.2) remains true for* $\dot{\overline{\gamma}}^v_{x,s}(0)$ *in integrated sense, i.e. for all* $\xi \in C^0_c(\Omega)$

$$\lim_{s\to 0}\frac{1}{s^2}\int_\Omega \xi(x)\left[\dot{\overline{\gamma}}^v_{x,s}(0) - sv\right]dx = -\frac{1}{2}\int_\Omega \xi(x)\Gamma_x(v,v)dx \tag{4.8}$$

Proof. For smooth (g_{ij}) from lemma 4.1 we obtain (4.8) because the error term in (4.2) is estimated by $o(s^2)$ locally uniformly w.r.t. x and v and thus

$$\begin{aligned}
\dot{\overline{\gamma}}^v_{x,s}(0) &= \frac{1}{m(B_{s^2}(x+sv))}\int_{B_1(0)} \dot\gamma_{x,x+s(v+s^2\theta)}(0)\sqrt{\det g_{ij}(x+s(v+s^2\theta))}d\theta \\
&= \frac{1}{m(B_{s^2}(x+sv))}\int_{B_1(0)} \left\{s(v+s^2\theta) - \frac{s^2}{2}\Gamma_x(v+s^2\theta, v+s^2\theta) + o(s^2)\right\} \\
&\qquad\qquad \times \sqrt{\det g_{ij}(x+s(v+s^2\theta))}d\theta \\
&= sv - \frac{s^2}{2}\Gamma_x(v,v) + o(s^2)
\end{aligned}$$

For general $(g_{ij}) \in \text{Lip}_{\text{loc}}(\Omega)$ we use the approximation by smooth (g^ϵ_{ij}) furnished by lemma 4.2. Hence for all $\epsilon > 0$ we obtain equation (4.8) with the quantities $\dot{\overline{\gamma}}^v_{x,s}$ and g_{ij}

replaced by $\dot{\overline{\gamma}}^{v,\epsilon}_{x,s}$ and ${g_{ij}}^\epsilon$, which we denote (4.8^ϵ). In order to pass to the limit in (4.8^ϵ) for ϵ tending to zero it is sufficient to show that

i) for s fixed $\int\limits_\Omega \xi(x)\dot{\overline{\gamma}}^{v}_{x,s}(0)dx = \lim\limits_{\epsilon\to 0}\int\limits_\Omega \xi(x)\dot{\overline{\gamma}}^{v,\epsilon}_{x,s}(0)dx$ and

ii) the convergence in (4.8^ϵ) for s tending to zero is uniform with respect to ϵ.

i) From lemma 4.3 we know that

$$\int\limits_\Omega \xi(x)\left[\overline{\gamma}^{v,\epsilon}_{x,s}(\delta)-x\right]dx \overset{\epsilon\to 0}{\longrightarrow} \int\limits_\Omega \xi(x)\left[\overline{\gamma}^{v}_{x,s}(\delta)-x\right]dx \quad \forall\,\xi\in C^0_c(\Omega), s\geq 0, \delta\geq 0.$$

Hence, if we can show that the limit

$$\lim_{\delta\to 0}\frac{1}{\delta}\int\limits_\Omega \xi(x)\left[\overline{\gamma}^{v,\epsilon}_{x,s}(\delta)-x\right]dx = \int\limits_\Omega \xi(x)\dot{\overline{\gamma}}^{v,\epsilon}_{x,s}(0)dx \tag{4.9}$$

is uniform w.r.t. ϵ then i) is proved. In fact, for any compact $K\subset\Omega$ and $x\in K$

$$\begin{aligned}
&\Big|\frac{1}{\delta}[\overline{\gamma}^{\epsilon}_{x,s}(\delta)-x]-\dot{\overline{\gamma}}^{\epsilon}_{x,s}(0)\Big| \\
&\quad=\left|\frac{1}{\delta}\int\limits_0^\delta\int\limits_0^\sigma \frac{1}{m^\epsilon(B_{s^3}(x+sv))}\int\limits_{B_{s^3}(x+sv)} \ddot{\gamma}^{\epsilon}_{x,\theta}(\tau)m^\epsilon(d\theta)d\tau d\sigma\right| \\
&\quad\leq \frac{1}{\delta}\int\limits_0^\delta\int\limits_0^\sigma \frac{1}{m^\epsilon(B_{s^3}(x+sv))}\int\limits_{B_{s^3}(x+sv)} \left\|\Gamma^\epsilon_{\gamma^\epsilon_{x,\theta}(\tau)}\right\|_{\mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R}^2} m^\epsilon(d\theta)d\tau d\sigma \\
&\quad\leq \frac{1}{2}\delta\sup_{x\in B_{2s}(K)}\|\Gamma^\epsilon_x\| \leq C_{K,s}\delta
\end{aligned} \tag{4.10}$$

by the uniform boundedness of $\sup\limits_{x\in K}\|\Gamma^\epsilon_x\|_{\mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R}^2}$ for all $K\subset\Omega$ compact. Obviously this implies (4.9).

ii) By definition of $\overline{\gamma}^{\epsilon,v}_{x,s}$

$$\begin{aligned}
\dot{\overline{\gamma}}^{\epsilon,v}_{x,s}(0)-sv = &\fint\limits_{B_{s^3}(x+sv)}\left(\dot{\gamma}^\epsilon_{x,\theta}(0)-[\gamma^\epsilon_{x,\theta}(1)-\gamma^\epsilon_{x,\theta}(0)]\right)m^\epsilon(d\theta) \\
&+\fint\limits_{B_{s^3}(x+sv)}(sv+x-\theta)m^\epsilon(d\theta).
\end{aligned}$$

The uniform boundedness of $\sup\limits_{x\in K}\|\Gamma^\epsilon_x\|_{\mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R}^2}$ for all $K\subset\Omega$ compact implies via (4.3) for $\gamma^{\epsilon,v}_{x,s}$ that $\dot{\gamma}^{\epsilon,v}_{x,s}(\rho)=sv+O(s^2)\,\forall\,\rho\in[0,1]$, $x\in K$, $v\in K'$ uniformly w.r.t. $\epsilon>0$, such that we can repeat the steps leading to (4.5) in a similar manner which yields

$$\dot{\gamma}^{\epsilon,v}_{x,s}(0)-sv = -\frac{s^2}{2}\Gamma^\epsilon_x(v,v)+s^2\int\limits_0^1\int\limits_0^\tau\fint\limits_{B_{s^3}(x+sv)}\left[\frac{1}{2}\Gamma^\epsilon_x-\Gamma^\epsilon_{\gamma^\epsilon_{x,\theta}(\sigma)}\right](v,v)m^\epsilon(d\theta)d\sigma d\tau+o(s^2).$$

If we multiply this equation with a compactly supported test function ξ and integrate with respect to x we obtain

$$\left|\frac{1}{s^2}\int_\Omega \xi(x)[\dot\gamma^{\epsilon,v}_{x,s}(0)-sv]dx+\frac{1}{2}\int_\Omega \xi(x)\Gamma^\epsilon_x(v,v)dx\right|\le\frac{1}{2}\|\xi\|_{L^2}\left\|\Gamma^\epsilon_\cdot(v,v)-\Gamma^{\epsilon,s}_\cdot(v,v)\right\|_{L^2}+\vartheta(s)$$

with

$$\Gamma^{\epsilon,s}_x(v,v)=2\int_0^1\int_0^\tau \fint_{B_{s^3}(x+sv)}\Gamma^\epsilon_{\gamma^\epsilon_{x,\theta}(\sigma)}(v,v)m^\epsilon(d\theta)d\sigma d\tau$$

which converges by a general property of L^p-functions to $\Gamma_\cdot(v,v)$ in L^p for all $p\in(1,\infty)$ if ϵ and δ tend to zero. This proves ii) and hence the lemma. □

The subsequent considerations require a clarification of the meaning of angles between geodesics emanating from the same point: for each $x\in\Omega\subset X$ we have a natural orientation on the tangent cone K_x as induced from the chart $\phi_{\{a_i\}}$ and the orientation of $\mathbb{R}^2$. Hence we may talk about rotations of K_x in positive or negative direction and consequently if K_x is non-singular for given $\theta\in\Sigma_x$ any $\eta\in\Sigma_x$ is uniquely determined by $\cos\sphericalangle(\theta,\eta)$ and the orientation of the basis $\{\eta,\theta\}$ in K_x. Moreover, the distinction between $\sphericalangle(\theta,\eta)$ and $\sphericalangle(\eta,\theta)$ makes sense and which are related by the formula $\sphericalangle(\theta,\eta)+\sphericalangle(\eta,\theta)=L(\Sigma_x)$, where $L(\Sigma_x)$ denotes the length of the space of directions over x. If K_x is regular, equivalently if $L(\Sigma_x)=2\pi$, and γ and η are geodesic segments parameterized by arclength emanating from x then we set

$$\sphericalangle(\gamma,\eta)=\operatorname{sign}(\gamma,\eta)d_\sphericalangle(\gamma,\eta)\bmod 2\pi,$$

where

$$\operatorname{sign}(\gamma,\eta)=\begin{cases}+1 & \text{if } \{\dot\gamma,\dot\eta\} \text{ pos. orientated in } K_x\\ -1 & \text{if } \{\dot\gamma,\dot\eta\} \text{ neg. orientated in } K_x.\end{cases}$$

If $x\in X$ is contained in an open orientable domain $O\subset X$, i.e. if the chart transitions $\phi_{\{a_i\}}\circ\phi^{-1}_{\{b_i\}}$ preserve the orientation of K_y for $y\in O$ then this definition of the angle between two geodesic segments is obviously independent of the specific choice of $\phi_{\{a_i\}}$. Moreover, if the images of γ and η under $\phi_{\{a_i\}}$ are C^1-curves in Ω, the properties of the weak Riemannian structure imply

$$d_{\sphericalangle_{g_x}}(\gamma,\eta)=\arccos(\langle\dot\gamma,\dot\eta\rangle_{g_x})$$

and analogously we define the angle $\sphericalangle_{g_x}$ between two arbitrary C^1-curves γ, η in $\Omega\subset\mathbb{R}^2$ starting from $x\in\Omega$ measured w.r.t. (g_{ij}) by the same formula as above

$$\sphericalangle_{g_x}(\gamma,\eta)=\operatorname{sign}(\dot\gamma,\dot\eta)d_{\sphericalangle_{g_x}}(\gamma,\eta).\tag{4.11}$$

Corollary 4.1. *Under the conditions of proposition 4.1 for all* $\xi \in C_c^0(\Omega)$

$$\lim_{s\to 0}\frac{1}{s}\int\limits_{\Omega}\fint_{B_{s^3}(x+sv)}\xi(x)\big[\sphericalangle_{g_x}(\dot\gamma_{x\theta},v_x^\perp)-\pi/2\big]m(d\theta)dx=\frac{1}{2}\int\limits_{\Omega}\xi(x)\frac{1}{\|v\|_{g_x}}\langle\Gamma_x(v,v),v_x^\perp\rangle_{g_x}dx$$
$$=\lim_{s\to 0}\frac{1}{s}\int\limits_{\Omega}\fint_{B_{s^3}(x-sv)}\xi(x)\big[\sphericalangle_{g_x}(v_x^\perp,\dot\gamma_{\theta x})-\pi/2\big]m(d\theta)dx \tag{4.12}$$

where the vectors $v_x^\perp$ *and* $v_\theta^\perp$ *are are obtained from* v *by a positive rotation of* v *about the angle* $\pi/2$ *with respect to the angular distance* $d_{\sphericalangle_{g_x}}$ *and* $d_{\sphericalangle_{g_\theta}}$ *respectively.*

Proof. We show the first identity in (4.12), the proof of the second is similar. By (4.7) we find

$$\left|\sin\frac{\sphericalangle_{g_x}(\dot\gamma_{x\theta},v)}{2}\right|=\frac{1}{2}\left\|\frac{\dot\gamma_{x\theta}}{\|\dot\gamma_{x\theta}\|_{g_x}}-\frac{v}{\|v\|_{g_x}}\right\|_{g_x}\le\frac{\|\dot\gamma_{x\theta}-sv\|\,\|sv\|\,|\,\|sv\|-\|\dot\gamma_{x\theta}\|\,|}{\|sv\|\,\|\dot\gamma_{x\theta}\|}=O(s) \tag{4.13}$$

for m-a.e. $\theta\in B_{s^3}(x+sv)$ which implies via the Taylor-expansion of arcsin about zero, i.e. $\sphericalangle_{g_x}(\dot\gamma_{x\theta},v)=\sin\sphericalangle_{g_x}(\dot\gamma_{x\theta},v)+o(s^2)$, that $\sphericalangle_{g_x}(\dot\gamma_{x\theta},v)=O(s)$. From this and from the triangle inequality for angular distances we see that $|\sphericalangle_{g_x}(\dot\gamma_{x\theta},v_x^\perp)-\pi/2|=O(s)$ and thus

$$\sphericalangle_{g_x}(\dot\gamma_{x\theta},v_x^\perp)-\pi/2=\sin\big(\sphericalangle_{g_x}(\dot\gamma_{x\theta},v_x^\perp)-\pi/2\big)+o(s^2)=-\cos\sphericalangle_{g_x}(\dot\gamma_{x\theta},v_x^\perp)+o(s^2)$$

for m-a.e. $\theta\in B_{s^3}(x+sv)$ locally uniformly w.r.t. $x\in\Omega$. Hence we may replace $\sphericalangle_{g_x}(\dot\gamma_{x\theta},v_x^\perp)-\pi/2$ by

$$-\cos\sphericalangle_{g_x}(\dot\gamma_{x\theta},v_x^\perp)=-\langle\frac{\dot\gamma_{x\theta}}{\|\dot\gamma_{x\theta}\|_{g_x}},v_x^\perp\rangle_{g_x}=-\langle\frac{\dot\gamma_{x\theta}-sv}{\|\dot\gamma_{x\theta}\|_{g_x}},v_x^\perp\rangle_{g_x}.$$

in the first integral in (4.12), which we rewrite

$$-\frac{1}{s}\int\limits_{\Omega}\fint_{B_{s^3}(x+sv)}\xi(x)\langle\frac{\dot\gamma_{x\theta}-sv}{\|sv\|_{g_x}},v_x^\perp\rangle_{g_x}m(d\theta)$$
$$-\frac{1}{s}\int\limits_{\Omega}\fint_{B_{s^3}(x+sv)}\xi(x)\langle\frac{(\dot\gamma_{x\theta}-sv)(\|sv\|_{g_x}-\|\dot\gamma_{x\theta}\|_{g_x})}{\|\dot\gamma_{x\theta}\|_{g_x}\|sv\|_{g_x}},v_x^\perp\rangle_{g_x}m(d\theta). \tag{4.14}$$

Furthermore, from (4.10) and the fact that $\dot\gamma_{x\theta}^\epsilon\to\dot\gamma_{x\theta}$ for m-a.e. $\theta\in B_{s^3}(x+sv)$ it follows that (4.4) remains true for m-a.e. $\theta\in B_{s^3}(x+sv)$ locally uniformly w.r.t. x and hence

$$\left\|\frac{(\dot\gamma_{x\theta}-sv)(\|sv\|_{g_x}-\|\dot\gamma_{x\theta}\|_{g_x})}{\|\dot\gamma_{x\theta}\|_{g_x}\|sv\|_{g_x}}\right\|\le\frac{O(s^2)O(s^2)}{\|v\|\,s(1+O(s^2))}\ \text{for } m\text{-a.e. } \theta\in B_{s^3}(x+sv)$$

which implies that the second integral in (4.14) tends to zero for $s\to 0$. Since all the functions depending on x except $\dot\gamma_{x\theta}$ in the first integral of (4.14) are continuous and are cut-of by ξ equality (4.8) proves the claim. □

Remark 4.3. The proofs of lemma 4.4 and corollary 4.1 show that if we replace $\xi(x)$ in (4.12) by $\xi(x+\lambda w)$ with $\lambda \in \mathbb{R}$ and $w \in \mathbb{R}^2$ then the convergence is locally uniform with respect to λ.

Proof of proposition 4.1. For $k \in \mathbb{N}$ and $x \in \Omega$ divide each side of the boundary $\partial(\Box_h(x)) = \bigcup_{i=1}^4 \partial^i(\Box_h(x))$ into $2k$ subintervals of length $s = s_k = \frac{h}{2k}$. Let $(v_h^i(x))^{i=1,\dots,4}$, $v_h^1(x) = x$ denote the four vertices of the square $\Box_h(x)$ (numbered in clockwise direction) and $(\sigma_j^1)_{j=1,\dots,2k-1}^{i=1,\dots,4}$ the interior partitioning points on the sides of $\partial(\Box_h(x))$. Further set $\sigma_0^i = v^i$ and $\sigma_{2k}^i = v^{i+1}$ where the upper index i is always taken modulo 4. Then for each odd interior partitioning point σ_{2l+1}^i $(l = 0, \dots, k-1)$ we choose some θ_{2l+1}^i in the Euclidean ball $B_{s^3}(\sigma_{2l+1}^i)$ and (w.r.t. (g_{ij})) geodesic segments $\gamma_{\sigma_{2l}^i \theta_{2l+1}^i}$ and $\gamma_{\theta_{2l+1}^i \sigma_{2(l+1)}^i}$ connecting θ_{2l+1}^i with its neighboring even partitioning points σ_{2l}^i and $\sigma_{2(l+1)}^i$ respectively. The concatenation of these segments defines the boundary of a polygonal domain $\tilde{\Box}_h(x)[k, (\theta_i^j)]$ depending on the choice of the (θ_j^i). By the local equivalence of the Euclidean and the (w.r.t. (g_{ij})) Riemannian distances on Ω the topological type of $\tilde{\Box}_h(x)[k, (\theta_i^j)]$ in $\mathbb{R}^2$ and in the Alexandrov space (X, d) are the same so that if the geodesic segments $\gamma_{\sigma_{2l}^i \theta_{2l+1}^i}$ and $\gamma_{\theta_{2l+1}^i \sigma_{2(l+1)}^i}$ are close to the corresponding Euclidean segments one may conclude that the domain $\tilde{\Box}_h(x)[k, (\theta_i^j)]$ is simply connected in (X, d) and orientable, i.e. for the Euler-Poincaré characteristic of $\tilde{\Box}_h(x)[k, (\theta_i^j)]$ in X we obtain $\chi_X(\tilde{\Box}_h(x)[k, (\theta_i^j)]) = 1$. Consequently, the combinatorial Gauss-Bonnet theorem (X, d) (cf. [Mac98]) applied to the domain $\tilde{\Box}_h(x)[k, (\theta_i^j)]$ yields

$$
\begin{aligned}
e_X(\tilde{\Box}_h(x)[k, (\theta_i^j)]) = 2\pi - \sum_{\substack{i=1,\dots,4 \\ l=0,\dots,k-1}} \left(\pi - \sphericalangle^{\text{int}}(\gamma_{\sigma_{2l}^i \theta_{2l+1}^i}, \gamma_{\theta_{2l+1}^i \sigma_{2(l+1)}^i})\right) \\
- \sum_{\substack{i=1,\dots,4 \\ l=1,\dots,k-1}} \left(\pi - \sphericalangle^{\text{int}}(\gamma_{\theta_{2l-1}^i \sigma_{2l}^i}, \gamma_{\sigma_{2l}^i \theta_{2l+1}^i})\right) - \sum_{i=1,\dots,4} \left(\pi - \sphericalangle^{\text{int}}(\gamma_{\theta_{2k-1}^{i-1} v^i}, \gamma_{v^i \theta_1^i})\right).
\end{aligned}
\tag{4.15}
$$

The idea is now, of course, to multiply this equation by a test function $\xi(x)$, integrate with respect to x and (θ_j^i) and let tend k tend to infinity. Taking the mean in (4.15) over all (θ_j^i) ranging over the Euclidean ball $B_{s^3}(\sigma_{2l+1}^i)$ with respect to the Riemannian measure m induced from (g_{ij}) and integrating this with respect to dx against a test function gives the identity

$$
\int_\Omega \xi(x) \fint_{B_{s^3}(\sigma_1^1(x))} \cdots \fint_{B_{s^3}(\sigma_{2k-1}^4(x))} e_X(\tilde{\Box}_h(x)[k, (\theta_i^j)(x)]) m(d\theta_1^1) \cdots m(d\theta_{2k-1}^4) dx = I_k(h) \tag{4.16}
$$

with

$$I_k(h) = \int_\Omega \xi(x) \left[\sum_{i=1}^4 \fint_{B_{s^3}(\sigma_{2k-1}^{i-1})} \fint_{B_{s^3}(\sigma_1^i)} \sphericalangle^{\mathrm{int}}(\gamma_{\theta_1 v^i}, \gamma_{v^i\theta_2}) m(d\theta_1) m(d\theta_2) - 2\pi \right] dx$$
$$- \int_\Omega \xi(x) \sum_{i=1}^4 \sum_{l=0}^{k-1} \fint_{B_{s^3}(\sigma_{2l+1}^i)} \left(\pi - \sphericalangle^{\mathrm{int}}(\gamma_{\sigma_{2l}^i\theta}, \gamma_{\theta\sigma_{2l+2}^i})\right) m(d\theta) dx$$
$$- \int_\Omega \xi(x) \sum_{i=1}^4 \sum_{l=0}^{k-1} \fint_{B_{s^3}(\sigma_{2l-1}^i)} \fint_{B_{s^3}(\sigma_{2l+1}^i)} \left(\pi - \sphericalangle^{\mathrm{int}}(\gamma_{\theta_1\sigma_{2l}^i}, \gamma_{\sigma_l^i\theta_2})\right) m(d\theta_1) m(d\theta_2) dx.$$

Let us identify the limit $\lim_{k\to\infty} I_k(h)$ first. We rewrite $I_k(h)$ as

$$\begin{aligned} &\int_\Omega \xi(x) \left[\sum_{i=1}^4 \fint_{B_{s^3}(\sigma_{2k-1}^{i-1})} \fint_{B_{s^3}(\sigma_1^i)} \sphericalangle^{\mathrm{int}}(\gamma_{\theta_1 v^i}, \gamma_{v^i\theta_2}) m(d\theta_1) m(d\theta_2) - 2\pi \right] dx \\ &\quad - \sum_{i=1}^4 \int_0^h \int_\Omega \xi(x) f_k^i(x,t) dx dt \end{aligned} \tag{4.17}$$

with $f_k^i : \Omega \times [0,h] \to \mathbb{R}$ defined for $i = 1, \dots, 4$ by

$$f_k^i(x,t) = \begin{cases} \frac{1}{s_k} \fint_{B_{s^3}(\sigma_{2l+1}^i)} \left(\pi - \sphericalangle^{\mathrm{int}}(\gamma_{\sigma_{2l}^i\theta}, \gamma_{\theta\sigma_{2l+2}^i})\right) m(d\theta) & \\ \qquad\qquad \text{if } t \in (\frac{h}{2k}(2l+1) - \frac{s_k}{2}, \frac{h}{2k}(2l+1) + \frac{s_k}{2}], & l = 0, \dots, k-1 \\ \frac{1}{s_k} \fint_{B_{s^3}(\sigma_{2l-1}^i)} \fint_{B_{s^3}(\sigma_{2l+1}^i)} \left(\pi - \sphericalangle^{\mathrm{int}}(\gamma_{\theta_1\sigma_{2l}^i}, \gamma_{\sigma_{2l}^i\theta_2})\right) m(d\theta_1) m(d\theta_2) & \\ \qquad\qquad \text{if } t \in (\frac{h}{2k}(2l) - \frac{s_k}{2}, \frac{h}{2k}(2l) + \frac{s_k}{2}], & l = 1, \dots, k-1 \\ 0 \qquad\qquad \text{if } t \in [0, \frac{s_k}{2}] \cup (h - \frac{s_k}{2}, h]. & \end{cases}$$

Estimate (4.7) and a discussion similar to the proof of corollary 4.1 show that the first integral in (4.17) converges to $\int_\Omega \xi(x) \big(\sum_{i=1}^4 \sphericalangle_i^{\mathrm{int}}(\square_h(x)) - 2\pi\big) m(dx)$. (This could also be deduced directly from (4.12) and the decomposition of angles in Euclidean tangent cones, compare (4.18) below). So it remains to prove the convergence of $\int_0^h \int_\Omega \xi(x) f_k^i(x,t) m(dx) dt$. Let us consider f_k^1 first. Note that relative to the orientation on $\mathbb{R}^2$ for the interior angles above we have

$$\sphericalangle^{\mathrm{int}}(\gamma_{\sigma_{2l}^i\theta}, \gamma_{\theta\sigma_{2l+2}^i}) = \sphericalangle_{g_\theta}(\dot\gamma_{\sigma_{2l}^i\theta}, \dot\gamma_{\theta\sigma_{2l+2}^i}) \text{ and } \sphericalangle^{\mathrm{int}}(\gamma_{\theta_1\sigma_{2l}^i}, \gamma_{\sigma_{2l}^i\theta_2}) = \sphericalangle_{g_{\sigma_{2l}^i}}(\dot\gamma_{\theta_1\sigma_{2l}^i}, \dot\gamma_{\sigma_{2l}^i\theta_2}),$$

whenever the differentials of those curves exist. Furthermore, since the tangent cones over all $x \in \Omega$ are Euclidean we may write

$$\begin{aligned} \pi - \sphericalangle^{\mathrm{int}}(\gamma_{\theta_1\sigma_{2l}^i}, \gamma_{\sigma_{2l}^i\theta_2}) &= \pi - \sphericalangle_{g_{\sigma_{2l}^i}}(\dot\gamma_{\theta_1\sigma_{2l}^i}, \dot\gamma_{\sigma_{2l}^i\theta_2}) \\ &= \sphericalangle_{g_{\sigma_{2l}^i}}(\dot\gamma_{\sigma_{2l}^i\theta_2}, \dot\gamma_{\theta_1\sigma_{2l}^i}) - \pi \\ &= [\sphericalangle_{g_{\sigma_{2l}^i}}(\dot\gamma_{\sigma_{2l}^i\theta_2}, e_{2\,\sigma_{2l}^i}^{\perp}) - \pi/2] + [\sphericalangle_{g_{\sigma_{2l}^i}}(e_{2\,\sigma_{2l}^i}^{\perp}, \dot\gamma_{\theta_1\sigma_{2l}^i}) - \pi/2] \end{aligned} \tag{4.18}$$

which is always true since we calculate the oriented angles modulo 2π. Together with the same kind of reasoning applied to $\sphericalangle_{g_{\sigma^i_{2l}}}(\dot\gamma_{\theta_1\sigma^i_{2l}}, \dot\gamma_{\sigma^i_{2l}\theta_2})$ we obtain the decomposition $f^1_k = f^{1,+}_k + f^{1,-}_k$. Now for any dyad $r \in h\cdot\{\frac{m}{2^l}|m, l \in \mathbb{N},\, m \le 2^l\}$ we apply the shift $\tau : \Omega \to \Omega$, $\tau(x) = x - r\cdot e_2$ and then the identity (4.12) of corollary 4.1 to the integrals $\int_\Omega \xi(x) f^{1,+}_k(x, r)dx$ and $\int_\Omega \xi(x) f^{1,-}_k(x, r)dx$ respectively which gives rise to

$$\lim_{k\to\infty} \int_\Omega \xi(x) f^1_k(x, r)dx = \int_\Omega \xi(x - re_2)\frac{1}{\|e_2\|_{g_x}} \langle \Gamma_x(e_2, e_2), e_{2x}^{\perp}\rangle_{g_x} dx$$

and since this limit is locally uniform w.r.t. r (see remark 4.3) we can replace r by any $s \in [0, h]$. Invoking (4.4) once more we see that there is a constant $C = C_\xi$ such that $\sup_k |\int_\Omega \xi(x) f^{i,\pm}_k(x, s)dx| < C$ for all $s \in [0, h]$ and hence we may conclude

$$\lim_{k\to\infty} \int_0^h \int_\Omega \xi(x) f^1_k(x, s)dxds = \int_0^h \int_\Omega \xi(x - se_2)\frac{1}{\|e_2\|_{g_x}} \langle \Gamma_x(e_2, e_2), e_{2x}^{\perp}\rangle_{g_x} dxds$$

by means of Lebesgue's dominated convergence theorem applied to the functions $s \to \int_\Omega \xi(x) f^{1,\pm}_k(x, s)dx$ on $[0, h]$. Finally, $e_{2x}^{\perp} = -(g_{22}(x)e_1 - g_{12}(x)e_2)/\sqrt{g_{22}(x) \det g(x)}$ and hence

$$\frac{1}{\|e_2\|_{g_x}} \langle \Gamma_x(e_2, e_2), e_{2x}^{\perp}\rangle_{g_x} = -\frac{\Gamma^1_{22}(x)\langle e_1, g_{22}(x)e_1 - g_{12}(x)e_2\rangle_{g_x}}{g_{22}(x)\sqrt{\det g(x)}} = -\frac{\sqrt{\det g(x)}}{g_{22}(x)}\Gamma^1_{22}(x).$$

In the other cases $f^i_k(.,.)$, $i = 2, 3, 4$ we argue in the same manner, noting that instead of $e_2^{\perp}$ in (4.18) we take $e_1^{\perp} = -(g_{12}e_1 - g_{11}e_2)/\sqrt{g_{11} \det g}$ for $i = 2$ and $-e^{\perp}_{i-1}$ for $i = 3, 4$. Hence we have shown that the for k tending to $I_k(h)$ converges to the right hand side of (4.1), where $h > 0$ was chosen arbitrarily.

It remains to identify the limit of the left hand side in (4.16), where we assume first that $\xi \ge 0$. From the local equivalence of the Riemannian and Euclidean metric it is clear that for sufficiently large k there is an $\epsilon > 0$ with $\epsilon \to 0$ for $k \to \infty$ such that for all $x \in \mathrm{supp}(\xi)$ with $h_\epsilon = h - 2\epsilon$ and $x_\epsilon = x + \sqrt{2}\epsilon(e_1 + e_2)$

$$\overline{\Box_{h_\epsilon}(x_\epsilon)} \subset \tilde\Box_h(x)[k, (\theta^j_i)(x)]$$

for all possible choices of $(\theta^i_j) \in B_{s^3}(\sigma^i_{2l+1})$. Since we may assume w.l.o.g. that e_X is nonnegative (otherwise we argue with the measure $e^+_X(dx) = e_X(dx) + |\kappa| m_X(dx)$ (c.f. [Mac98] or section 6.2)) we may infer from this that

$$e_X(\overline{\Box_{h_\epsilon}(x_\epsilon)}) \le e_X(\tilde\Box_h(x)[k, (\theta^j_i)(x)])$$

from which we deduce by passing to the limit for $k \to \infty$ in (4.16) that for all $\epsilon > 0$

$$\int_\Omega \xi(x) e_X(\overline{\Box_{h_\epsilon}(x_\epsilon)})dx \le \lim_{k\to\infty} I_k(h) =: I(h).$$

Hence, using the fact that $\overline{\square_{h_\epsilon}(x_\epsilon)} \nearrow \square_h(x)$ for each x and $\epsilon \to 0$ as well as the monotone convergence theorem we obtain

$$\int_\Omega \xi(x) e_{X*}(\square_h(x))dx \leq I(h),$$

where e_{X*} denotes the inner measure associated to e_X. (An analogous discussion for the enlarged squares $\square_{h_\epsilon}(x_\epsilon)$ with $h_\epsilon = h + 2\epsilon$ and $x_\epsilon = x - \sqrt{2}\epsilon(e_1 + e_2)$ would yield $\int_\Omega \xi(x) e_X(\square_h(x))dx \geq I(h)$, where $e_X = e_X^*$ by definition of e_X.) But now we are done since

$$\int_\Omega e_X(\square_h(x))\xi(x)dx \leq \int_\Omega e_{X*}(\square_{h+\epsilon}(x))\xi(x)dx \leq I(h+\epsilon)$$

for all $\epsilon > 0$ and the continuity of (g_{ij}) implies that the function $h \to I(h)$ is continuous. Hence

$$\int_\Omega e_X(\square_h(x))\xi(x)dx \leq I(h) = \lim_{\epsilon \to 0} I(h-\epsilon) \leq \int_\Omega e_{X*}(\square_h(x))\xi(x)dx,$$

which implies

$$e_{X*}(\square_h(x)) = e_X(\square_h(x)) \text{ for } dx\text{-a.e.} x \in \Omega \setminus B_h(\partial\Omega)$$

since $e_{X*} \leq e_X$, i.e. $\square_h(x)$ is e_X-measurable for dx-a.e. $x \in \Omega \setminus B_h(\partial\Omega)$ as well as

$$\int_\Omega \xi(x) e_X(\square_h(x))dx = \lim I(h) \quad \forall\, 0 \leq \xi \in C_c^0(\Omega).$$

which is (4.1) for non-negative ξ. For general $\xi \in C_c^0(\Omega)$ the claim follows from the decomposition $\xi = \xi_+ - \xi_-$. $\square$

4.2 Distributional Gaussian Curvature Bounds

Proposition 4.2 (Weak lower curvature bounds). *Let (X, d) be a two-dimensional Alexandrov space with Curv$(X) \geq \kappa \in \mathbb{R}$ and discrete metrically singular set S_X such that the system of natural coordinates with the associated metric tensors (g_{ij}) yield a $C^{1,1}$ Riemannian structure for $X \setminus S_X$, i.e. the chart transition maps are $C^{1,1}_{loc}$ and the metric tensors (g_{ij}) are locally uniformly Lipschitz continuous. Then the Gaussian curvature k is bounded from below on U by κ in the distributional sense, i.e. let $\Omega = \phi_{\{a_i\}}(U) \subset \mathbb{R}^2$ be the image of $U \subset X \setminus S_X$ under $\phi_{\{a_i\}}$ and let $g_{\{a_i\}} = (g_{ij}) : \Omega \to \mathbb{R}^{2\times 2}$, $g_{ij} \in Lip_{loc}(\Omega)$, be the Riemannian tensor on U with respect to ϕ, then for all $0 \leq \xi \in C_c^\infty(\Omega)$ the inequality*

$$\int_\Omega \left[\xi_{,12} \arccos\left(\frac{g_{12}}{\sqrt{g_{11}g_{22}}}\right) - \xi_{,1}\frac{\sqrt{\det g}}{g_{22}}\Gamma^1_{22} - \xi_{,2}\frac{\sqrt{\det g}}{g_{11}}\Gamma^2_{11}(x)\right]dx \geq \kappa \int_\Omega \xi\sqrt{\det g}\,dx$$

holds true, where the functions $\Gamma^k_{ij}(x) = \frac{1}{2}g^{kl}(x)\{D_{x_i}g_{jl}(x) + D_{x_j}g_{il}(x) - D_{x_l}g_{ij}(x)\}$ are obtained from the weak differentials $D_{x_l}g_{ij}$ of g_{ij}.

Proof. Remember that $\lim_{h\to 0} m_X(\Box_h(x))/h^2 = \sqrt{\det g_{ij}(x)}$ locally uniformly on Ω. Thus if we multiply $\xi(x)$ in the integral on the left hand side of our Gauss-Bonnet formula (4.1) by $1 = m_X(\Box_h(x))/m_X(\Box_h(x))$, then the result in [Mac98] that

$$\lim_{h\to 0} \frac{e(\Box_h(x))}{m(\Box_h(x))} \geq \kappa \quad m\text{-a.e.}$$

together with Fatou's lemma implies

$$\liminf_{h\to 0} \frac{1}{h^2} \int_\Omega \xi(x) e(\Box_h(x)) dx \geq \kappa \int_\Omega \xi(x) \sqrt{\det g_{ij}(x)} dx.$$

Thus it remains to compute the limit of the right hand side of (4.1) divided by h^2 for h tending to zero. From the regularity of ξ it follows easily that

$$\lim_{h\to 0} \frac{1}{h^2} \int_\Omega \int_0^h \left[\xi(x - te_2) - \xi(x - (h-t)e_2 - he_1)\right] \frac{\sqrt{\det g}}{g_{22}}(x) \Gamma^1_{22}(x) dx dt$$
$$= \int_\Omega \xi_{,1}(x) \frac{\sqrt{\det g}}{g_{22}}(x) \Gamma^1_{22}(x) dx$$

where $\xi_{,1}$ denotes the differential of ξ with respect to the i-th direction in $\mathbb{R}^2$ and analogously

$$\lim_{h\to 0} \frac{1}{h^2} \int_\Omega \int_0^h \left[\xi(x - (h-t)e_1) - \xi(x - he_2 - te_1)\right] \frac{\sqrt{\det g}}{g_{11}}(x) \Gamma^2_{11}(x) dx dt$$
$$= \int_\Omega \xi_{,2}(x) \frac{\sqrt{\det g}}{g_{11}}(x) \Gamma^2_{11}(x) dx.$$

As for the last remaining integral we may use the fact that $K_x \simeq \mathbb{R}^2$ for each $x \in \Omega$ once more in order to write

$$\begin{aligned}\sum_{i=1}^4 \sphericalangle_i^{\text{int}}(\Box_h(x)) - 2\pi =& \sphericalangle_x(e_1, e_2) + \sphericalangle_{x+he_2}(-e_2, e_1) \\ &+ \sphericalangle_{x+h(e_1+e_2)}(-e_1, -e_2) + \sphericalangle_{x+he_1}(e_2, -e_1) - 2\pi \\ =& \sphericalangle_x(e_1, e_2) + \pi - \sphericalangle_{x+he_2}(e_1, e_2) + \sphericalangle_{x+h(e_1+e_2)}(e_1, e_2) \\ &+ \pi - \sphericalangle_{x+he_1}(e_1, e_2) - 2\pi \\ =& \sphericalangle_{g_x}(e_1, e_2) - \sphericalangle_{g_{x+he_2}}(e_1, e_2) + \sphericalangle_{g_{x+h(e_1+e_2)}}(e_1, e_2) - \sphericalangle_{g_{x+he_1}}(e_1, e_2)\end{aligned}$$

which equals, since trivially $\text{sign}(e_1, e_2) = 1$, by (4.11)

$$\begin{aligned}=& \arccos\left(\frac{g_{12}}{\sqrt{g_{11}g_{22}}}\right)(x) - \arccos\left(\frac{g_{12}}{\sqrt{g_{11}g_{22}}}\right)(x + he_2) \\ &+ \arccos\left(\frac{g_{12}}{\sqrt{g_{11}g_{22}}}\right)(x + h(e_1 + e_2)) - \arccos\left(\frac{g_{12}}{\sqrt{g_{11}g_{22}}}\right)(x + he_1).\end{aligned}$$

From this it is now easy to see that in fact

$$\lim_{h\to 0}\frac{1}{h^2}\int\limits_{\Omega}\xi(x)\big(\sum_{i=1}^{4}\sphericalangle_i^{\text{int}}(\Box_h(x))-2\pi\big)dx=\int\limits_{\Omega}\xi_{,12}\arccos(\frac{g_{12}}{\sqrt{g_{11}g_{22}}})dx$$

which completes the proof. □

Remark 4.4. For general Alexandrov spaces the assumption that $g_{ij}\in \mathrm{Lip}_{\mathrm{loc}}$ seems to be too strong. In fact, the weaker assumption $g_{ij}\in \mathrm{BV}_{\mathrm{loc}}(\Omega)\cap C^0(\Omega)$ would be more natural, which follows from the results in [OS94] and [Per98], but for the reasons indicated in remark 4.1 we have to confine ourselves to the situation stated above.

In the following corollary the Laplacian and harmonic functions on Alexandrov spaces are understood in terms of the weak Riemannian structure, see section 5.2 or the papers [KMS01, KS98]. On account of the distributional Gaussian curvature bounds we now may engage the paper [CH98] which readily yields

Corollary 4.2 (Gradient Estimate). *Let (X,d) an Alexandrov space with Curv $(X)\geq -k^2$ as in proposition 4.2 then any nonnegative harmonic function u on X satisfies Yau's gradient estimate*

$$\left\|\frac{\nabla u}{u}\right\|_{L^\infty(X,m)}\leq\sqrt{k}. \tag{4.19}$$

Remark 4.5. Using cut-off functions as in lemma 3.3 in [KMS01] it is obviously possible to extend the previous statement a littler further to Alexandrov surfaces with a sufficiently small exceptional set outside which the assumptions of propositions 4.2 apply. For instance, this is the case if the singular set is discrete.

Examples 4.1.

i) Riemannian surfaces with lower sectional curvature bounds,

ii) surfaces of revolution in $\mathbb{R}^2$ induced by Lipschitz continuous functions $f:[a,b]\to\mathbb{R}_+$ satisfying $f''\leq k^2 f$ in a weak sense, compare [CH98] and

iii) as a special case of ii) - or iv) - one may take the double $S_h\cup_{\partial S_h}S_h$ of a spherical cap $S_h=\{(x,y,z)\in S^2\subset\mathbb{R}^3\,|\,z\geq h\}$ with its induced inner metric.

iv) two-dimensional Riemannian simplicial complexes with a discrete set of 0-simplexes,

v) Alexandrov surfaces obtained by more general surgery operations, i.e. by cutting and gluing two-dimensional Riemannian pieces of constant curvature in such a way that metrically singular points occur only in a discrete fashion, for instance spaces of the kind X_f or X_k of section 5.1 ex. iii) and iv) respectively,

5 The Heat Kernel on Alexandrov Spaces

5.1 Volume Regularity

On a smooth Riemannian manifold (M^d, g) one obtains from the expansion of the volume density function $\det d\exp_x$ for the pull back of the Riemannian volume form on the tangent space at x

$$\frac{d\exp_x^* \mathrm{vol}_M}{d\mathcal{L}^n}(z) = \det d\exp_x(z) = 1 - \frac{1}{6}\,\mathrm{Ricc}(z,z) + o(|z|^2). \tag{5.1}$$

This formula implies in particular that both tangential and intrinsic mean value operators can be used for the approximation of the Laplace-Beltrami operator on M:

Lemma 5.1. *Let (M, g) be a smooth (i.e. C^3) Riemannian manifold and $f : M \mapsto \mathbb{R}$ some C^3-function. Then for all $x \in M$*

$$\begin{aligned}\lim_{r\to 0}\frac{1}{r^2}\left[f(x) - \fint_{B_r(x)} f(z)vol_M(dz)\right] &= \lim_{r\to 0}\frac{1}{r^2}\left[f(x) - \fint_{\mathbb{B}_r(0_x)\subset T_xM} f(\exp_x z)dz\right]\\ &= -\frac{1}{2(n+2)}\Delta^M f(x).\end{aligned}$$

Proof.

$$\begin{aligned}\fint_{B_r(x)} f(z)\mathrm{vol}(dz) &= \frac{1}{\mathrm{vol}(B_r(x))}\int_{\mathbb{B}_r(0_x)} f(\exp_x(z))\det d_z\exp_x dz\\ = \frac{1}{\mathrm{vol}(B_r(x))}&\int_{\mathbb{B}_r(0_x)}\left(f(x) + df_x(z) + \frac{1}{2}\mathrm{Hess}_x f(z,z) - \frac{1}{6}f(x)\,\mathrm{Ricc}(z,z) + o(|z|^2)\right)dz\end{aligned}$$

and hence, since $\int_{\mathbb{B}_r(0)} A(z,z) = (n+2)^{-1}r^2|\mathbb{B}_r|\mathrm{tr}(A)$

$$= \frac{|\mathbb{B}_r|}{\mathrm{vol}(B_r(x))}\left((1 - \frac{s(x)r^2}{6(n+2)})f(x) + \frac{r^2}{2(n+2)}\Delta f(x) + o(r^2)\right).$$

Thus

$$\begin{aligned}\frac{1}{r^2}\left[f(x) - \fint_{B_r(x)} f(z)\mathrm{vol}_M(dz)\right] &= \frac{-|\mathbb{B}_r|}{\mathrm{vol}(B_r(x))}\frac{1}{2(n+2)}\Delta f(x) + \theta(r)\\ &+ \frac{|\mathbb{B}_r|}{\mathrm{vol}(B_r(x))}\frac{1}{r^2}\left(\frac{\mathrm{vol}(B_r(x))}{|\mathbb{B}_r|} - \left(1 - \frac{s(x)r^2}{6(n+2)}\right)\right)f(x).\end{aligned}$$

By virtue of the expansion of $(r \mapsto \frac{\mathrm{vol}(B_r(x))}{|\mathbb{B}_r|})$ about zero (which follows from integrating (5.1) over the ball $\mathbb{B}_r(0_x)$) the second term on the right hand side converges to zero and the first equality follows. The second is now also obvious. □

Remark 5.1.
The same assertion is true for the analogous spherical mean value operators. Moreover, by the Kato-Trotter formula from the preceding assertion one deduces for the family of rescaled mean value operators

$$M_t f(x) := \fint_{B_{\sqrt{t/(n+2)}}(x)} f(y)m(dy) \qquad \lim_{k\to\infty} M_{t/k}^k f(x) = e^{t\frac{1}{2}\Delta} f(x)$$

on a smooth n-dimensional Riemannian manifold M and for sufficiently smooth f, compare [Blu84]. - Similar statements can be made for the operators

$$K_t f(x) := (2\pi t)^{-n/2} \int_M \exp(-d(x,y)^2/2t) f(y) \mathrm{vol}(dy)$$
$$Q_t f(x) := (2\pi t)^{-n/2} \int_M \exp(-\frac{d(x,y)^2}{2t} + t\frac{s(x)+s(y)}{12}) f(y) \mathrm{vol}(dy)$$

for which one obtains ([Sun84], [AD99] resp.)

$$K_{t/k}^k f(x) \to \exp(t\frac{1}{2}\Delta - \frac{1}{6}s) f(x)$$
$$Q_{t/k}^k f(x) \to \exp(t\frac{1}{2}\Delta) f(x)$$

if $k \to \infty$ for all $x \in M$, where $s(x)$ is the scalar curvature in x.

The proof of lemma 5.1 relies on the integrated version of (5.1) only, i.e. on the asymptotic behaviour of the volume density functions $q_r : X \mapsto \mathbb{R}$

$$q_r(x) = \frac{m(B_r(x))}{b_{n,k}(r)}$$

for r tending to zero. Here $b_{n,k}(r)$ denotes the volume of an r-ball in the model space $\mathbb{M}_{n,k}$. This is the motivation for the following

Definition 5.1 (Volume Regularity). *For $\alpha > 0$ a point x in an n-dimensional Alexandrov space with curvature bounded below by k is said to be α-volume regular iff*

$$1 - q_r(x) = o(r^\alpha) \text{ for } r \to 0.$$

(X, d) is called locally volume (L^p, α)-regular with exceptional set $\mathbb{S}_X^{p,\alpha}$, iff $\mathbb{S}_X^{p,\alpha} \subset X$ is closed and

$$r^{-\alpha}(1 - q_r(.)) \to 0 \text{ in } L^p_{loc}(X \setminus \mathbb{S}_X^{p,\alpha}) \text{ for } r \to 0.$$

Remarks 5.2.
a) The existence of the limit of $m(B_r(p))/b_{n,k}(r)$ for r tending to zero follows from the Bishop-Gromov volume comparison on Alexandrov spaces (cf. [Yam96]).

b) In [She93] one finds the statement that if $\lim\limits_{r\to 0} \frac{m(B_r(x))}{b_{n,k}(r)} > \frac{1}{2}$ then p is (topologically) a manifold point. This can be seen as well from the results in [BGP92] and (5.2).
c) A volume α-regular point is also metrically regular. This is seen as follows: since the tangent cone K_p is the pointed Gromov-Hausdorff-limit of the rescaled metric r-balls centered at p, i.e. for all $R > 0$

$$(B^r_R, d^r) = (B^R(p), \frac{1}{r} d^X) \xrightarrow{GH} (\mathbb{B}_R(0_p), d_p) \text{ for } r \to 0,$$

by theorem 10.8 in [BGP92] one obtains that the associated Hausdorff measures converge weakly, too. Thus

$$m_{K_p}(\mathbb{B}_1(0_p)) = \lim_{r\to 0} m^r(B^r_1) = \lim_{r\to 0} \frac{m(B_r(p))}{r^n}. \tag{5.2}$$

The condition of volume α-regularity implies that the limit in (5.2) equals one, and hence $\Sigma_p = \partial\mathbb{B}_1(0_p)$ has the measure ω_n and therefore must be isometric to S^{n-1}. - In general the converse assertion is true only for $\alpha = 0$, see example iii) below.

Examples 5.1.

i) A smooth Riemannian manifold is trivially volume (L^∞, α)-regular with empty exceptional set $\mathbb{S}^{\infty,\alpha}_X$ for $\alpha < 2$.

ii) Locally finite simplicial Riemannian complexes obtained from Riemannian simplices of constant curvature are locally (L^∞, ∞)-volume-regular with $\mathbb{S}^{\infty,\infty}_X = S_X$.

iii) Take a planar circular cone $C_f := B_1(0)\setminus\Sigma_f$ in the two-dimensional Euclidean plane, i.e. the unit disc minus some sector $\Sigma_f := \{(x_1, x_2) \in \mathbb{R}^2 | \, x_1 > 0, |x_2| < f(x_1)\}$. If f is chosen to be convex, due to the gluing theorem [Pet97], the metric space X_f, which is obtained by gluing C_f (endowed with the induced Euclidean distance) along the graph of $\pm f$, is an Alexandrov space of nonnegative curvature. For linear f, X_f is a flat cone which is a special case of ii). Assume f is not linear and differentiable.
Note that if $f'(0) = 0$ then X_f is metrically regular. Let $p = (s, \pm f(s)) \in \text{graph}(f) \subset X_f$ then for sufficiently small r the ball $B_r(p) \subset X_f$ is given by union of the intersection of the Euclidean r-Balls $B^+_r(p)$ (and $B^-_r(p)$) centered at $(s, f(s))$ (and $(s, -f(s))$) with C_f respectively and it is sufficient to compute $|B^+_r(p) \cap \Sigma_f|/\pi r^2$, where $|.|$ denotes the two-dimensional Lebesgue measure. If we shift p to the origin and rotate the picture $B^+_r(p)\cap\Sigma_f$ is the intersection of the upper half r-ball around 0 with the epigraph of a convex function $\tilde f$ such that $\tilde f(0) = \tilde f'(0) = 0$. If $x_\pm = x_\pm(r)$ denote the x-coordinates of the intersection points of $\partial B_r(0)$ with the graph of $\tilde f$ in the first and second quadrant respectively, i.e. $r = x_\pm\sqrt{1 + (\tilde f(x\pm)/x_\pm)^2}$,

$x_+ > 0$, $x_- < 0$

then the estimates

$$\frac{1}{2} - \frac{\int_{x_-}^{x_+} \tilde{f}(s)ds}{\pi r^2} \geq \frac{|B_r^+(p) \cap \Sigma_f|}{\pi r^2}$$
$$\geq \frac{1}{2} - \frac{\int_{x_-}^{x_+} \tilde{f}(s)ds + \tilde{f}(x_-)(r - x_-) + \tilde{f}(x_+)(r - x_+)}{\pi r^2}$$

hold trivially and since

$$\frac{(r - x_\pm)\tilde{f}(x_\pm)}{r^{2+\alpha}} \approx \frac{\tilde{f}(x_\pm)}{|x_\pm|^{1+\alpha}} \text{ and } \frac{\int_0^{x_\pm} \tilde{f}(s)ds}{|x_\pm|^{2+\alpha}} \approx \frac{f(x_\pm)}{|x_\pm|^{1+\alpha}}$$

for small $r \approx |x_\pm|$ one obtains for sufficiently small r

$$C_1 \frac{\tilde{f}(x_+)}{x_+^{1+\alpha}} + \frac{\tilde{f}(x_-)}{|x_-|^{1+\alpha}} \leq \frac{1}{r^\alpha}(1 - q_r(p)) \leq C_2 \frac{\tilde{f}(x_+)}{x_+^{1+\alpha}} + \frac{\tilde{f}(x_-)}{|x_-|^{1+\alpha}}. \tag{5.3}$$

For instance, if the function $\tilde{f}$ has growth $x^{1+\gamma}$ around 0 then p is an α-volume regular point iff $\gamma > \alpha$. The function $\tilde{f}$ is given as the image of f under a affine transformation of f depending on the point p. Suppose that the estimate $c_1 r^{1+\gamma} \leq \tilde{f}(r) \leq c_2 r^{1+\gamma}$ for $r \in (-\epsilon, \epsilon)$ holds true locally uniformly with respect to p. Since the upper estimate in (5.3) obviously applies to points in a r-neighborhood of graph(f), one obtains that X_f is locally (L^p, α)-volume regular with empty exceptional set $\mathbb{S}_X^{p,\alpha}$, if $\alpha < (1 + \gamma p)/p$, in particular for $\mathbb{S}_X^{1,1} = \emptyset$ for any $\gamma > 0$.

iv) We sketch without proofs an idea how to construct a two-dimensional Alexandrov space X with $\mathbb{S}_X^{1,1} \neq \emptyset$ based on the previous example. Suppose that the lower estimate in (5.3) holds (with a different constant C_1) in the $r/2$-neighborhood of graph$(\pm f)$. Then the idea is to iterate the cut-off-and-glue procedure of iii) with cusps similar to Σ_f such that the volume of the r-neighbourhoods of the resulting non-volume regular points in the limit space decays like r^β with $\beta < 1$ for r tending to zero.

To start with take the Euclidean rectangle $[0, 1] \times [0, L] \subset \mathbb{R}^2$ with $L > 0$ and a self similar cantor set $K = [0, 1] \setminus \bigcup_{k \in \mathbb{N}} D_k \hookrightarrow [0, 1] \times \{0\} \subset [0, 1] \times [0, L]$, where each D_k

consists of a finite number of mutually disjoint open intervals, $D_{k+1} \subset [0,1]\setminus\bigcup_{i\le k} D_i$ and $\bigcup_k D_k$ is dense in $[0,1]$. Then for each interval (a,b) of D_k, $k\in\mathbb{N}$ we fit a copy $\Lambda^{(a,b)}$ of Σ_f into (a,b) by rotating Σ_f about $-\pi/2$ to point in positive $y-$direction and shifting it along the (x,y)-plane into the 'gap' (a,b), i.e. such that $\partial\Lambda^{(a,b)}\cap\{y=0\}=\{(a,0),(b,0)\}$. Choosing L sufficiently large we find $\Lambda^{(a,b)}\cap\{y=L\}=\emptyset$. Let

$$C_{K,f}=([0,1]\times[0,L])\setminus\bigcup_{(a,b)\in D_k,\,k\in\mathbb{N}}\Lambda^{(a,b)}$$

be the resulting Euclidean domain after removing all cusps $\{\Lambda^{(a,b))}\}_{(a,b)\in D_k,\,k\in\mathbb{N}}$ from $[0,1]\times[0,L]$ and denote the fractal part of its boundary by

$$F_{K,f}=\bigcup_{(a,b)\in D_k,\,k\in\mathbb{N}}\partial\Lambda^{(a,b)}\cap([0,1]\times[0,L])\,.$$

Then it follows from the construction of $C_{K,f}$ that for $z\in K\hookrightarrow[0,1]\times\{0\}\subset C_{K,f}$ and any neighbourhood $U_z\subset\mathbb{R}^2$ of z we have a lower bound for the two-dimensional measure of the r-neighbourhood of $F_{K,f}$ in U_z given by

$$|B_r\left(F_{K,f}\right)\cap C_{K,f}\cap U_z|\ge C_3r^{\beta}\qquad\forall\,r\le r_0 \tag{5.4}$$

with some $\beta<1$ that depends on the choice of f and K.

Using Petrunin's gluing theorem we may show that $C_{K,f}$ gives rise to an Alexandrov surface $X_{K,f}$ with nonnegative curvature and empty boundary if we glue $C_{K,f}$ along the adjacent pairs of branches of $\partial\Lambda^{(a,b)}\cap([0,1]\times[0,L])$ for $(a,b)\in D_k$, $k\in\mathbb{N}$ as well as along the two vertical boundary segments $\{0\}\times[0,L]$ and $\{1\}\times[0,L]$ and finally attach a disk to the sphere $[0,1]\times\{L\}$. - Alternatively the space $X_{K,f}$ can be characterized as the Gromov-Hausdorff-limit of $X^l_{K,f}$ for $l\to\infty$, where $X^l_{K,f}$ is obtained similarly to $X_{K,f}$ from

$$C^l_{K,f}=([0,1]\times[0,L])\setminus\bigcup_{(a,b)\in D_k,\,k\le l}\Lambda^{(a,b)}$$

by gluing together $\{0\}\times[0,L]$ and $\{1\}\times[0,L]$ and the adjacent branches of $\partial\Lambda^{(a,b)}\cap([0,1]\times[0,L])$ for $(a,b)\in D_k$, $k\le l$, by attaching a flat disk to sphere $[0,1]\times\{L\}$ and an appropriate convex flat polygonal domain to the lower part of the boundary $\partial C^l_{K,f}\cap\{y=0\}$. Since $\bigcup_k D_k$ is dense in $[0,1]$ the set $\partial C^l_{K,f}\cap\{y=0\}\hookrightarrow X^l_{K,f}$ converges to a single point in $X_{K,f}$ which we denote by z. Then by (5.4) for any neighborhood $U_z\subset X_{K,f}$ one finds that $m_{X_{K,f}}(B_r(\Lambda)\cap U_z)\ge C\cdot r^{\beta}$ where Λ denotes the image of $\bigcup_{(a,b)\in D_k,\,k\in\mathbb{N}}\partial\Lambda^{(a,b)}\hookrightarrow X_{K,f}$ and which yields together with the left hand side of (5.3)

$$\int_{U_z}\frac{1}{r}(1-q_r(x))m_{X_{K,f}}(dx)\ge C(U_z,f)\frac{r^{\beta}\cdot r^{1+\gamma}}{r^2}\nearrow\infty\text{ for }r\to0$$

if $\beta+\gamma<1$ and hence $z\in\mathbb{S}^{1,1}_{X_{K,f}}\neq\emptyset$.

5.2 Dirichlet Forms and Laplacians on Metric Spaces

Recently the development of differential calculus and potential theory on general metric spaces have gained much attention. In the framework of Alexandrov spaces a direct and very natural construction of Sobolev spaces, Laplacians and associated diffusion processes is possible. We present the earlier results in [Stu96, Stu98, KS01] concerning the definition and certain properties of a canonical intrinsic Dirichlet form. The idea is simple and based on the approximation of the gradient of a function on smooth spaces by difference quotients. Accordingly one defines for an open subset G in X, $r > 0$ and a measurable function $u : X \mapsto \mathbb{R}$ the family of approximating Dirichlet (or energy-) forms by

$$\mathcal{E}^r_G(u) = \frac{C}{2} \int\limits_G \int\limits_{B^*_r(x)} \left(\frac{u(x) - u(y)}{d(x,y)} \right)^2 m_r(dy) m_r(dx) \tag{5.5}$$

with $B^*_r(x)$ being the punctured geodesic ball of radius r around x and the measure $m_r(dx) = dx/\sqrt{m(B_r(x))}$. The constant C plays the role of a dimension, but can be chosen arbitrarily. The generator of this form

$$A_r u(x) = \rho_r(x) \left[u(x) - \int\limits_X u(y) \sigma_r(x,dy) \right]$$

with the Markov transition kernel

$$\sigma_r(x,dy) = \frac{1}{s_r(x)} \frac{1}{d^2(x,y)} \frac{1}{\sqrt{m(B_r(y))}} \mathbb{1}_{B^*_r(x)}(y) m(dy)$$

$$s_r(x) = \int\limits_{B^*_r(x)} \frac{1}{d^2(x,y)} \frac{1}{\sqrt{m(B_r(y))}} m(dy), \quad \rho_r(x) = \frac{C s_r(x)}{\sqrt{m(B_r(x))}}$$

is a non-local bounded symmetric operator on $L^2(G,m)$ giving rise to the L^2-semigroup $P^r_t = e^{-tA_r}$ whose associated continuous time Markov jump process $(\Xi^r_t)_{t\geq 0}$ is characterized by the one-step transition function $\sigma_r(x,dy)$ and the exponential jump rate $\rho_r(x)$ at a given point $x \in X$. In [Stu98] it is shown that if some metric measure space (X,d,m) possesses the so called measure contraction property, which is equivalent to lower Ricci curvature bounds in the smooth Riemannian case, the forms of type (5.5) converge to a limiting form $\mathcal{E}$ for r tending to zero in the Γ-sense, which in particular preserves the Dirichlet form properties for the limiting functional $\mathcal{E}_G$ and hence we obtain an m-symmetric Hunt process Ξ generated by $\mathcal{E}_G$.

Moreover, since the convergence of $\mathcal{E}^r_G$ is in fact a little better, namely in the stronger sense of Mosco (cf. [Mos94]). the semigroups P^r_t converge strongly to P_t, which means that the processes Ξ^r convergence to Ξ in the sense of finite dimensional distributions. Thus the limiting form $\mathcal{E}$ generates a process which is the scaling limit of a family of intrinsic jump process on (X, d, m) whose data are just the metric d and the measure m.

Kuwae and Shioya [KS01] modified this result for forms of the type

$$\mathcal{E}_G^{b,r}(u) = \frac{n}{2}\frac{1}{b_{n,k}(r)} \int\limits_G \int\limits_{B_r^*(x)} \left(\frac{u(x)-u(y)}{d(x,y)}\right)^2 m(dy)m(dx)$$

defined on Alexandrov spaces with lower curvature bound, for which they introduced the notion of generalized measure contraction property (see [KS01] for details). Also, they proved that on Alexandrov spaces in the limit both approximating forms $\mathcal{E}_G^{b,r}$ and $\mathcal{E}_G^r$ (for $C = n$) give the same result (ibid. Corollary 5.1). Summing up these results one obtains the following assertions:

Theorem 5.1 (Existence and uniqueness of the canonical Dirichlet form). *On an n-dimensional Alexandrov space with lower curvature bound k and some $G \subset X$ with $m(G) < \infty$ both sequences $\mathcal{E}_G^r$ and $\mathcal{E}_G^{b,r}$ have Γ-limits on $L^2(G)$ which coincide with their common pointwise limit $\mathcal{E}_G$ on Lip(X). The $L^2(G)$-closure $(\mathcal{E}_G, D(\mathcal{E}_G))$ of $\mathcal{E}_G$ on Lip(X) is a strongly local and regular Dirichlet form on $L^2(G)$.*

The generator Δ^G of this form will be called Laplacian for it coincides (up to a trivial time scaling) with the Neumann Laplace-Beltrami operator if X is smooth (see lemma 5.1). Moreover, since the measure m is doubling and a local weak Poincaré inequality applies, by Moser iteration one can show the existence and Hölder continuity of the heat kernel q_t^G for the corresponding semigroup [Stu96, Stu98]. In particular the semigroup has the Feller property, which will be used later on. Also, the analogous results hold true if one considers the Dirichlet Laplacian instead which is defined as the generator of the closure of $(\mathcal{E}, \mathrm{Lip}_c(G))$. For the sake of completeness we should mention also the approach in [KMS01] for the definition of the Laplacian on Alexandrov spaces, where the energy form $\mathcal{E}$ is defined just as in the smooth case by employing the weak Riemannian structure of (X, d). - In any case the m-symmetric Hunt process on X which is associated with $\mathcal{E}$ is called canonical because of its close and unique relation with the metric d (see also section 5.7).

For the comparison with the Riemannian case we introduce a third sequence of operators and corresponding forms which will be useful later and which are given by

$$A^{E,r}u(x) = n\tfrac{m(B_r(x))}{b_{n,k}(r)}\tfrac{1}{r^2}\left(u(x) - \fint\limits_{B_r(x)} u(y)m(dy)\right)$$

$$\mathcal{E}_G^{E,r}(u) = \tfrac{n}{2b_{n,k}(r)} \int\limits_G \int\limits_{B_r(x)} \tfrac{|u(x)-u(y)|^2}{r^2} m(dx)m(dy).$$

Lemma 5.2. *On Lip(G) the limit of* $\mathcal{E}_G^{E,r}$ *for* $r \to 0$ *exists and coincides with* $\frac{n}{n+2}\mathcal{E}_G$.

Proof. We compare $\mathcal{E}_G^{E,r}$ with $\mathcal{E}_G^{b,r}$. For this purpose define for $\alpha \in (0,1)$

$$\mathcal{E}_{G,B_\alpha}^{E,r}(u) = \frac{n}{2b_{n,k}(r)} \int_G \int_{B_{\alpha r}(x)} \frac{|u(x)-u(y)|^2}{r^2} m(dx)m(dy)$$

$$\mathcal{E}_{G,A_\alpha}^{E,r}(u) = \frac{n}{2b_{n,k}(r)} \int_G \int_{B_r(x)\setminus B_{\alpha r}(x)} \frac{|u(x)-u(y)|^2}{r^2} m(dx)m(dy)$$

and $\mathcal{E}_{G,B_\alpha}^{b,r}$ and $\mathcal{E}_{G,A_\alpha}^{b,r}$ analogously. Then we obtain

$$\begin{aligned}
\mathcal{E}_G^{b,r}(u) &= \mathcal{E}_{G,B_\alpha}^{b,r}(u) + \mathcal{E}_{G,A_\alpha}^{b,r}(u) \\
&\leq \mathcal{E}_{G,B_\alpha}^{b,r}(u) + \frac{1}{\alpha^2}\mathcal{E}_{G,A_\alpha}^{E,r}(u) \\
&= \mathcal{E}_{G,B_\alpha}^{b,r}(u) - \frac{1}{\alpha^2}\mathcal{E}_{G,B_\alpha}^{E,r}(u) + \frac{1}{\alpha^2}\mathcal{E}_G^{E,r}(u) \\
&= \frac{b_{n,k}(r\alpha)}{b_{n,k}(r)}\mathcal{E}_G^{b,\alpha r}(u) - \frac{b_{n,k}(r\alpha)}{b_{n,k}(r)}\mathcal{E}_G^{E,\alpha r}(u) + \frac{1}{\alpha^2}\mathcal{E}_G^{E,r}(u).
\end{aligned}$$

Due to the convergence of $\mathcal{E}_G^{b,r}(u)$ and $\lim_{r\to 0}\frac{b_{n,k}(r\alpha)}{b_{n,k}(r)} = \alpha^n$ in the limit this yields

$$\begin{aligned}
\mathcal{E}_G(u) &\leq \frac{1}{\alpha^2(1-\alpha^n)} \liminf_{r\to 0} \left(\mathcal{E}_G^{E,r}(u) - \alpha^2 \frac{b_{n,k}(r\alpha)}{b_{n,k}(r)} \mathcal{E}_G^{E,\alpha r}(u) \right) \\
&\leq \frac{1}{\alpha^2(1-\alpha^n)} \left(\liminf_{r\to 0} \mathcal{E}_G^{E,r}(u) - \alpha^{n+2} \liminf_{r\to 0} \mathcal{E}_G^{E,\alpha r}(u) \right) \\
&= \frac{1-\alpha^{n+2}}{\alpha^2(1-\alpha^n)} \liminf_{r\to 0} \mathcal{E}_G^{E,r}(u).
\end{aligned}$$

As $\lim_{\alpha\to 1}\frac{1-\alpha^{n+2}}{\alpha^2(1-\alpha^n)} = \frac{n+2}{n}$ we see $\mathcal{E}_G(u) \leq \frac{n+2}{n}\liminf_{r\to 0}\mathcal{E}_G^{E,r}(u)$ if we send α to 1. The reverse inequality can be proved in a similar way. One writes

$$\begin{aligned}
\mathcal{E}_G^{E,r}(u) &= \mathcal{E}_{G,B_\alpha}^{E,r}(u) + \mathcal{E}_{G,A_\alpha}^{E,r}(u) \\
&\leq \frac{\alpha^2 b_{n,k}(r\alpha)}{b_{n,k}(r)} \mathcal{E}_G^{E,\alpha r}(u) + \mathcal{E}_G^{b,r}(u) - \frac{b_{n,k}(r\alpha)}{b_{n,k}(r)}\mathcal{E}_G^{b,\alpha r}(u).
\end{aligned}$$

Hence

$$\begin{aligned}
(1-\alpha^n)\mathcal{E}_G(u) &= \lim_{r\to 0}\left(\mathcal{E}_G^{b,r}(u) - \frac{b_{n,k}(r\alpha)}{b_{n,k}(r)}\mathcal{E}_G^{b,\alpha r}(u) \right) \\
&\geq \limsup_{r\to 0}\left(\mathcal{E}_G^{E,r}(u) - \frac{\alpha^2 b_{n,k}(r\alpha)}{b_{n,k}(r)}\mathcal{E}_G^{E,\alpha r}(u) \right) \\
&\geq (1-\alpha^{n+2})\limsup_{r\to 0}\mathcal{E}_G^{E,r}(u).
\end{aligned}$$

Consequently, upon dividing this inequality by $(1-\alpha^n)$ and letting tend α to 1 one obtains $\mathcal{E}_G(u) \geq \frac{n+2}{n}\limsup_{r\to 0}\mathcal{E}_G^{E,r}(u)$ and the claim follows. □

5.3 Laplacian Comparison

The main result of this section is the comparison theorem for Δ^X which is a generalization of the well known Laplacian comparison theorem for Riemannian manifolds with lower Ricci curvature bound. Since a precise characterization of $\mathrm{Dom}(\Delta^X)$ is not available, the inequality for Δ^X is stated in the weak form, which involves the Dirichlet form $\mathcal{E}$ which is generated by Δ^X:

Theorem 5.2 (Laplacian Comparison). *Let X be an n-dimensional Alexandrov space with curvature bounded below by $k \in \mathbb{R}$ which is locally volume $(L^1, 1)$-regular with exceptional set $\mathbb{S}_X^{1,1}$ of rough dimension $\leq n-2$. Then for any $f \in C^3(\mathbb{R})$ with $f' \leq 0$, $0 \leq \zeta \in D(\mathcal{E})$ and $p \in X$*

$$\mathcal{E}(f \circ d_p, \zeta) \leq \langle -S_k^{1-n}(S_k^{n-1} f')' \circ d_p, \zeta \rangle_{L^2(X,m)} \tag{5.6}$$

where $d_p(x) := d(p, x)$ and

$$S_k(t) = \begin{cases} 1/\sqrt{k} \sin(\sqrt{k}t) & \text{if } k > 0 \\ t & \text{if } k = 0 \\ 1/\sqrt{(-k)} \sinh(\sqrt{-k}t) & \text{if } k < 0. \end{cases}$$

Proof. The proof is based upon the choice of a suitable metric on each tangent cone K_x, $x \in X$ and a modification $\Lambda_x : X \mapsto K_x$ of the inverse of the exponential map which allows to apply the Alexandrov convexity of (X, d) - For each $x \in X$ we equip the tangent cone K_x over x with the hyperbolic, spherical or flat metric d_k^x defined by the corresponding cosine law, i.e.

$$\begin{aligned} \cosh(\sqrt{-k} d_k^x[(\alpha, s), (\beta, t)]) &= \cosh(\sqrt{-k}s) \cosh(\sqrt{-k}t) \\ &\quad - \sinh(\sqrt{-k}s) \sinh(\sqrt{-k}t) \cos d_{\sphericalangle}(\alpha, \beta) \\ \cos(\sqrt{k} d_k^x[(\alpha, s), (\beta, t)]) &= \cos(\sqrt{k}s) \cos(\sqrt{k}t) \\ &\quad - \sin(\sqrt{k}s) \sin(\sqrt{k}t) \cos d_{\sphericalangle}(\alpha, \beta) \\ (d_k^x[(\alpha, s), (\beta, t)])^2 &= s^2 + t^2 - 2st \cos d_{\sphericalangle}(\alpha, \beta) \end{aligned}$$

depending on whether $k < 0$, $k > 0$ or $k = 0$ respectively in order to obtain a new curved tangent cone $K_{x,k} = \left((\Sigma_x \times \mathbb{R}_+)/ \sim_{d_k^x}, d_k^x\right)^{\sim}$, which will be denoted $(\mathbb{M}_x, d^x)$. Let $\mathbb{M}_{n,k}(X) = \dot{\bigcup}_{x \in X} \mathbb{M}_x$ denote the corresponding curved tangent cone bundle and note that $(\mathbb{M}_x, d^x) \backsimeq \mathbb{M}_{n,k}$ for each $x \in X \setminus S_X$. Then the map $\Lambda_x : X \mapsto \mathbb{M}_x$ is chosen as the canonical radially isometric extension of (some choice of) the map

$$\lambda_x : X \mapsto \Sigma_x \quad \lambda_x(z) = \gamma'_{xz}(0)$$

which is the projection of z onto one of the directions by which it is seen from x. It is possible to choose λ_x in such a way that the map $\Lambda : X \times X \mapsto \mathbb{M}_{n,k}(X), (x, z) \mapsto \Lambda_x(z) \in$

$\mathbb{M}_{n,k}(x)$ is measurable, where $\mathbb{M}_{n,k}(X)$ is endowed with the product sigma algebra. (Also there is a natural measure on $\mathbb{M}_{n,k}(X)$ as a product of m and the n-dimensional Hausdorff measure in each fiber.) Now $\mathrm{Curv}(X) \geq k$ implies that for $x \in X \setminus S_X$ the map Λ_x is expanding. In fact, for $y,\, z \in M$ let $\overline{y},\, \overline{z}$ be the image points of $y,\, z$ under Λ_x and denote $0_x = \Lambda_x(x)$, then by construction of Λ_x

$$d^x(0_x, \overline{y}) = d(x,y), \quad d^x(0_x, \overline{z}) = d(x,z)$$
$$\sphericalangle(\gamma_{xy}, \gamma_{xz}) = d_{\sphericalangle}(\gamma'_{xy}, \gamma'_{xz}) = \sphericalangle(\gamma_{0_x\overline{y}}, \gamma_{0_x\overline{z}})$$

where $\gamma_{0_x\overline{y}}$ and $\gamma_{0_x\overline{z}}$ denote the uniquely defined geodesics in $\mathbb{M}_x$ joining 0_x with $\overline{y}$ and $\overline{z}$ respectively. Since $\mathbb{M}_x \backsimeq \mathbb{M}_{n,k}$ one obtains

$$d^x(\Lambda_x(y), \Lambda_x(z)) \geq d(y,z) \tag{5.7}$$

because the contrary would mean a contradiction to the Alexandrov convexity for geodesic hinges (proposition 3.1) in the global version.

Let now be f a function as required and $p \in X$. Then $f \circ d_p$ is Lipschitz and thus in $D(\mathcal{E})$. Let us first assume that the nonnegative test function $\zeta \in D(\mathcal{E})$ has compact support. On account of the lemma 5.2 and the polarization identity we know that

$$\langle A^r(f \circ d_p), \zeta \rangle_{L^2(X,m)} = \mathcal{E}^{E,r}(f \circ d_p, \zeta) \to \frac{n}{n+2}\mathcal{E}(f \circ d_p, \zeta) \text{ for } r \to 0.$$

If p_x denotes the image point of p under Λ_x the monotonicity property of f together with (5.7) yields

$$A^r(f \circ d_p)(x) \leq q_r(x)\frac{n}{r^2}\left(f(d^x(p_x, 0_x)) - \fint_{B_r(x)} f(d^x(p_x, \Lambda_x(y)))m(dy) \right)$$

with $q_r(x) = \frac{m(B_r(x))}{b_{n,k}(r)}$. Also by (5.7) the image measure of m under Λ_x on $\mathbb{M}_{n,k}(x)$ is absolutely continuous with respect to the volume measure vol_x on $\mathbb{M}_x$

$$\frac{d\left((\Lambda_x)_* m\right)}{d\mathrm{vol}_x} =: \rho_x \leq 1 \quad \mathrm{vol}_x\text{-a.e.}$$

Thus by the general integral transformation formula

$$\begin{aligned} A^r(f \circ d_p)(x) &\leq \frac{q_r(x)n}{r^2}\left(f(d^x(p_x, 0_x)) - \frac{1}{q_r(x)} \fint_{\mathbb{B}_r(0_x)} f \circ d^x_{p_x}(z)\rho_x(z)\mathrm{vol}_x(dz) \right) \\ &= q_r(x)\frac{n}{r^2}\left(f(d^x(p_x, 0_x)) - \fint_{\mathbb{B}_r(0_x)} f \circ d^x_{p_x}(z)\mathrm{vol}_x(dz) \right) \\ &\quad + q_r(x)\frac{n}{r^2} \fint_{\mathbb{B}_r(0_x)} f \circ d^x_{p_x}(z)\left[1 - \frac{\rho_x(z)}{q_r(x)}\right]\mathrm{vol}_x(dz). \end{aligned} \tag{5.8}$$

As $\mathbb{S}_X^{1,1}$ has rough dimension $\leq n-2$ from lemma 3.3 in [KMS01] it follows that we can find a sequence of cut-off functions in $D(\mathcal{E})$ vanishing on some neighborhood of $\big(\{p\}\cup\mathbb{S}_X^{1,1}\big)\cap \mathrm{supp}(\zeta)$ and converging to the constant function 1 in the Dirichlet space $(D(\mathcal{E}),||.||_1^{\mathcal{E}})$. So we may assume that ζ is zero on some neighborhood of $\{p\}\cup\mathbb{S}_X^{1,1}$.

The volume regularity of X implies that $q_r(x)\to 1$ for (some subsequence if necessary) $r\to 0$ m-a.e. on $\mathrm{supp}(\xi)$ and thus by lemma 5.1 and the special structure of the Laplace-Beltrami operator $\Delta^{\mathbb{M}_{n,k}}$ acting on radial functions one obtains for the first term on the right hand side of (5.8) and $x\in\mathrm{supp}(\xi)$

$$\lim_{r\to 0}\frac{q_r(x)n}{r^2}\left(f(d^x(p_x,0_x))-\fint_{\mathbb{B}_r(0_x)} f\circ d^x_{p_x}(z)\rho_x(z)\mathrm{vol}_x(dz)\right)$$
$$=-\frac{n}{n+2}\Delta^{K_x}(f\circ d^x_{p_x})(0_x)=\frac{n}{n+2}S_k^{1-n}(S_k^{n-1}f')'\circ d_{p_x}(0_x)$$
$$=\frac{n}{n+2}S_k^{1-n}(S_k^{n-1}f')'(d(p,x)).$$

For the second term on the right hand side of (5.8) note first that we may assume without loss of generality that $f(0)=0$. Then by the regularity of f

$$\frac{n}{r^2}\fint_{\mathbb{B}_r(0_x)} f\circ d^x_{p_x}(z)\left[1-\frac{\rho(z)}{q_r(x)}\right]\mathrm{vol}(dz)$$
$$=\frac{n}{r^2}\frac{f'(d_p(x))}{d_p(x)}\fint_{\mathbb{B}_r(0_x)}\langle p_x,z\rangle\left[1-\frac{\rho(z)}{q_r(x)}\right]\mathrm{vol}(dz)$$
$$+\frac{1}{2}\frac{n}{r^2}\fint_{\mathbb{B}_r(0_x)}\mathrm{Hess}_{0_x}[f\circ d^x_{p_x}](z,z)+o_f(|z|^2)\left[1-\frac{\rho(z)}{q_r(x)}\right]\mathrm{vol}(dz).$$

Using the relation $1-\frac{\rho(z)}{q_r(x)}=(1-q_r^{-1}(x))(1-\frac{1-\rho(z)}{1-q_r(x)})$ and rescaling the integrals we see that the second term in the right hand side of (5.8) equals

$$n\frac{q_r(x)-1}{r}\frac{f'(d_p(x))}{d_p(x)}\fint_{\mathbb{B}_1(0_x)}\langle p_x,z\rangle\frac{1-\rho(rz)}{1-q_r(x)}\mathrm{vol}(dz)+$$
$$\frac{n}{2}(1-q_r(x))\fint_{\mathbb{B}_1(0_x)}\mathrm{Hess}_{0_x}[f\circ d^x_{p_x}](z,z)\left[1-\frac{1-\rho(zr)}{1-q_r(x)}\right]\mathrm{vol}(dz)+\vartheta_f(r)$$
$$\leq nf'(d_p(x))\frac{q_r(x)-1}{r}+n(1-q_r(x))C(f'',\mathrm{supp}(\zeta))+\vartheta_f(r)$$

with $\vartheta_f(r)\to 0$ for $r\to 0$. Thus we get the desired inequality for compactly supported ζ from the local $(L^1,1)$-volume regularity if we multiply (5.8) by ζ, integrate over X and let r tend to zero.

For general $\zeta\in C^0(X)\cap D(\mathcal{E})$ take a sequence of smooth nonnegative functions with compact support $\eta_k:\mathbb{R}_+\mapsto[0,1]$ such that $\eta(t)=1$ for $t\in[0,k]$ and set $\zeta_k=\zeta\cdot\eta_k\circ d_p$

in order to obtain (5.6) for ζ_k. The Dirichlet form $\mathcal{E}$ is strongly local and hence the corresponding energy measure $\mu_{\langle .,. \rangle}$, which is a measure valued symmetric bilinear form on $D(\mathcal{E})$ defined by

$$\int_X \phi(x)\mu_{\langle u,u\rangle}(dx) = 2\mathcal{E}(u,\phi u) - \mathcal{E}(u^2,\phi) \quad \forall \phi \in C_0(X)$$

has the derivation property

$$d\mu_{\langle u\cdot v,w\rangle} = u d\mu_{\langle v,w\rangle} + v d\mu_{\langle u,w\rangle}$$

for all u, v and $w \in \mathcal{D}(\mathcal{E})$ (cf. [FOT94], section 3.3.2). Also from the construction of $\mathcal{E}$ it is obvious that $\mu_{\langle u,v\rangle} \ll m$ for u, $v \in D(\mathcal{E})$ and thus $\mathcal{E}$ admits the Carré du Champ operator Γ which is defined via the corresponding density, i.e.

$$\Gamma(u,v) := \frac{d\mu_{\langle u,v\rangle}}{dm} \in L^1(X,m) \quad \forall\, u,\, v \in D(\mathcal{E})$$

and which yields the representation $\mathcal{E}(u,v) = \int_X \Gamma(u,v)dm$ for u, $v \in D(\mathcal{E})$. Hence a twofold application of Lebesgue's theorem and $\lim_{k\to\infty}\Gamma(f\circ d_p, \eta_k \circ d_p) = 0$ m-a.e. yield

$$\begin{aligned}
\mathcal{E}(f\circ d_p,\zeta) &= \int_X \Gamma(f\circ d_p,\zeta)dm \\
&= \lim_{k\to\infty}\int_X \eta_k\circ d_p\Gamma(f\circ d_p,\zeta)dm + \lim_{k\to\infty}\int_X \zeta\Gamma(f\circ d_p,\eta_k\circ d_p)dm \\
&= \lim_{k\to\infty}\int_X \Gamma(f\circ d_p,\eta_k\circ d_p\cdot\zeta)dm = \lim_{k\to\infty}\mathcal{E}(f\circ d_p,\zeta_k) \\
&\leq -\lim_{k\to\infty}\langle S_k^{1-n}(S_k^{n-1}f')'\circ d_p,\zeta_k\rangle_{L^2(X,m)} \\
&= -\langle S_k^{1-n}(S_k^{n-1}f')'\circ d_p,\zeta\rangle_m .
\end{aligned}$$

□

The weak inequality (5.6) becomes a pointwise bound if $f\circ d_p \in D(\Delta^X)$, in which case the classical result is completely recovered:

Corollary 5.1. *If $f\circ d_p \in D(\Delta^X)$ for some $p\in X$ and non-increasing $f\in C^3(\mathbb{R})$ then*

$$\Delta^X(f\circ d_p)(x) \geq S_k^{1-n}(S_k^{n-1}f')'\circ d_p(x) \quad \textit{for } m\textit{-a. a. } x \textit{ in } X.$$

Remark 5.3. Probably one might think about a different approach to proving a comparison theorem for the Laplacian on Alexandrov spaces via some sort of second variation formula for arclength or generalized Jacobi fields, compare [Ots98] for some work in this direction. However, such an approach seems to require much more work on geodesics on Alexandrov spaces and is probably less extendable to more general situations than the proof given above.

Remark 5.4. Under the weaker assumption that $1 - q_r(x) \leq O(r)$ locally uniformly on $X \setminus \mathbb{S}_X$ one obtains an additional drift term in the Laplacian comparison principle which then takes the form

$$\mathcal{E}(f \circ d_p, \zeta) \leq (-S_k^{1-n}(S_k^{n-1} f')' \circ d_p, \zeta)_{L^2(X,m)} \\ - (n+2) \sup_{\nu} \int_{\mathbb{M}_{n,k}(X)} \frac{f'(d_p(x))}{d_p(x)} \langle p_x, z \rangle \zeta(x) \nu(dx, dz) \tag{5.9}$$

where supremum with respect to ν is taken over all weak accumulation points of the weakly precompact sequence of measures on $\mathbb{M}_{n,k}(X)$

$$\nu_r(dz, dx) = \frac{1 - q_r(x)}{r} \frac{1 - \rho_x(rz)}{1 - q_r(x)} 1\!\!1_{\mathbb{B}_1(0_x)}(z) \mathrm{vol}_x(dz) m(dx).$$

If $z \mapsto \rho_x(z)$ is differentiable in 0_x we find

$$\nu_r(dx, dz) \to \nu(dx, dz) = d\rho_x(z) 1\!\!1_{\mathbb{B}_1(0_x)}(z) \mathrm{vol}_x(dz) m(dx)$$

and the drift part in (5.9) becomes

$$\int_X \frac{f'(d_p(x))}{d_p(x)} d\rho_x(p_x) \zeta(x) m(dx).$$

However, the drift term in (5.9) can be interpreted as a measure for the local approximation of X by its tangent spaces.

5.4 Heat Kernel Comparison

The following paragraphs contain a number of corollaries to the weak Laplacian comparison, the first concerning the associated Dirichlet heat kernel. For this recall that the *Dirichlet Laplacian* is the generator of $(\mathcal{E}_\Omega, D_c(\mathcal{E}_\Omega))$ obtained from taking the closure of the set of compactly in Ω supported Lipschitz functions with respect to the Dirichlet norm $\|.\|_1^2 = \|.\|_{L^2(\Omega)}^2 + \mathcal{E}_X(.)$. As usual the fundamental solution q_t^G of the corresponding heat equation is called Dirchlet heat kernel.

Theorem 5.3. *Let q_t^G be the Dirichlet heat kernel on some domain $G \subset X$ and let $x \ni B_r(x) \subset X$. Then for all $y \in B_r(x)$ and $\overline{x}, \overline{y} \in \mathbb{M}_{n,k}$ with $d(x, y) = \overline{d}(\overline{x}, \overline{y})$*

$$q_t^G(x, y) \geq q_t^{k,r}(\overline{x}, \overline{y}) \tag{5.10}$$

where $\overline{d}$ denotes the distance on $\mathbb{M}_{n,k}$ and $q_t^{k,r}$ is the Dirichlet heat kernel of $B_r(\overline{x}) \subset \mathbb{M}_{n,k}$.

Example 5.2. Consider the heat kernel $q_t^{C_f}$ on C_f (where C_f is defined as in example iii) of section 5.1) satisfying the boundary conditions

$$f = 0 \text{ on } S^1 \cap \partial C_f, \quad \frac{\partial u}{\partial \nu}(x, f(x)) = -\frac{\partial u}{\partial \nu}(x, -f(x)) \text{ on } \partial \Sigma_f$$

where $\frac{\partial u}{\partial \nu}$ is the exterior normal derivative of u. The conditions on $\partial\Sigma_f$ are chosen in such a way that solving this boundary value problem is consistent with gluing the two half-sectors together along the graph of $\pm f$. By the heat kernel comparison theorem we now get a lower bound for the flat heat kernel $q_t^{C_f}$ on C_f of the form $q_t^{C_f}(x,y) \geq q_t(x,y)$, where q_t is the Euclidean Dirichlet heat kernel on B_1.

The proof of theorem 5.3 is an application of theorem 5.2 to the heat kernel on the model space together with following simple version of a parabolic maximum principle for Δ^X:

Lemma 5.3. *Let $f : \Omega \times (0,T) \to \mathbb{R}$ with $f \in L^2([0,T], D_c(\mathcal{E}_\Omega)) \cap C([0,T]), L^2(\Omega))$. If $f_0 := f(0,.) \leq 0$ m-a. e. and $Lf \geq 0$ with $L = \Delta^\Omega - \partial_t$ in the following weak sense*

$$\int_\sigma^\tau \mathcal{E}(f(t,.),\xi)dt \leq -\left\langle f(.),\xi\right\rangle_m\big|_\sigma^\tau \tag{5.11}$$

for all $\sigma,\tau \in (0,T)$ and $0 \leq \xi \in D_c(\mathcal{E}_\Omega)$, then $f(t,x) \leq 0$ for m-a.e. $x \in \Omega$ and $t \in [0,T]$.

Proof. For $\epsilon > 0$ consider $f_\epsilon(t,x) = \frac{1}{\epsilon}\int_t^{t+\epsilon} f(s,x)ds$. Then f_ϵ is a subsolution to the heat equation in the following sense: for all nonnegative $\xi \in L^2([0,T], D_c(\mathcal{E}_\Omega))$ and $\sigma,\tau \in (0,T-\epsilon)$:

$$\begin{aligned}
\int_\sigma^\tau \mathcal{E}(f_\epsilon(t,.),\xi(t))dt &= \int_\sigma^\tau \frac{1}{\epsilon}\int_t^{t+\epsilon} \mathcal{E}(f(s,.),\xi(t))dsdt \\
&\leq -\int_\sigma^\tau \frac{1}{\epsilon}\langle f(t+\epsilon,.) - f(t,.),\xi(t,.)\rangle_m dt \\
&= -\int_\sigma^\tau \langle \partial_t f_\epsilon(t,.),\xi(t,.)\rangle_m dt
\end{aligned}$$

Now we take ξ to be $j(f_\epsilon)$ where $j \in C^\infty(\mathbb{R})$ with $j' \geq 0$, $\|j'\|_\infty < \infty$ and $j = 0$ on $(-\infty,\delta]$, $j > 0$ on $(\delta,+\infty)$ for some $\delta > 0$. From the definition of $D_c(\mathcal{E}_\Omega)$ it follows that this choice of ξ is admissible. Inserting this into the last inequality yields

$$0 \leq \int_\sigma^\tau\int_\Omega j'(f_\epsilon(t,x))\mu_{\langle f_\epsilon(t)\rangle}(dx)dt \leq \int_\sigma^\tau \underbrace{\langle \partial_t f_\epsilon(t), j(f_\epsilon)\rangle_m}_{=\partial_t\langle J(f_\epsilon(t))\rangle_m} dt = \langle J(f_\epsilon(\sigma))\rangle_m - \langle J(f_\epsilon(\tau))\rangle_m$$

with $J(t) = \int_0^t j(s)ds$. Using $f_\epsilon(\sigma) \to f_\epsilon(0)$ and $f_\epsilon(0) \to f_0$ in $L^2(\Omega)$ as well as $J(f_0) = 0$ m-a.s., by sending first $\sigma \to 0$ and then $\epsilon \to 0$ we find that $\langle J(f(\sigma))\rangle_m \leq 0$. From the definition of j and J resp. this implies that $f(t,x) \leq \delta$ for m a.e. $x \in \Omega$. Sending $\delta \to 0$ yields the claim. □

Proof of theorem 5.3. Let us first assume that $\overline{B_r(x)} \subset X$. If $q_t^{k,r}(.,.)$ denotes the Dirichlet heat kernel on $B_r(\overline{x}) \subset \mathbb{M}_{n,k}$ then there is the uniquely defined real valued function $(t,s) \mapsto h_t(s) = h_t^{k,r}(s)$ satisfying the differential equation $\partial_t h_t(s) = -S_k^{1-n}\partial_s(S_k^{n-1}\partial_s h_t(s))$ for $(t,s) \in \mathbb{R}_+ \times (0,r)$ and such that $q_t^{k,r}(\overline{x},\overline{y}) = h_t(\overline{d}(\overline{x},\overline{y}))$ (see [Cha84]). Furthermore, since $s \mapsto h_t(s)$ is non-increasing (ibid., lemma 2.3), the continuation of h_t (still denoted by $h_t(s)$)

$$\mathbb{R}_+ \times \mathbb{R}_+ \ni (t,s) \to h_t^{k,r}(s) = \begin{cases} h_t(s) & \text{for } s \in [0,r] \\ 0 & \text{for } s > r \end{cases}$$

is locally Lipschitz in both variables, non-increasing in s and satisfies

$$\partial_t h_t(s) \geq -S_k^{1-n}\partial_s(S_k^{n-1}\partial_s h_t(s)) \tag{5.12}$$

in the distributional sense. In order to prove (5.12) it is sufficient to note that for the Laplacian $\Delta^{\mathbb{M}_{n,k}}$ of the function $\overline{q}_t^{k,r,\overline{x}} : \mathbb{R}_+ \times \mathbb{M}_{n,k} \to \mathbb{R}_+$, $(t,\overline{y}) \to h_t(d(\overline{x},\overline{y}))$ one finds

$$-\Delta^{\mathbb{M}_{n,k}}\overline{q}_t^{k,r,\overline{x}} = -\Delta_y^{\mathbb{M}_{n,k}} q_t^{k,r}(\overline{x},.)1\!\!1_{B_r(\overline{x})} + \frac{\partial q_t^{k,r}(\overline{x},.)}{\partial \nu_{|\partial B_r(\overline{x})}} d\mathcal{H}^{n-1}_{|\partial B_r(\overline{x})}$$

in the distributional sense, where the density in front of the Hausdorff-measure is obviously non-positive. Testing this inequality with radially symmetric test functions and using the special form of $\Delta^{\mathbb{M}_{n,k}}$ we obtain (5.12). Hence, if we mollify $(t,s) \to h_t(s)$ with respect to s by a non-negative smooth kernel we obtain a family $(t,s) \to h_t^\rho(s)$ of smooth functions, non-increasing in s, satisfying (5.12) in a pointwise sense and which converge to h locally uniformly on $\mathbb{R}_+ \times \mathbb{R}_+$ for $\rho \to 0$, which in particular implies that for all $T > 0$ and ρ sufficiently small (depending on T) the function

$$G \ni y \to \psi_t^\rho(y) = h_t^\rho(d(x,y))$$

belongs to $D_c(\mathcal{E}(G))$ for all $t \in (0,T)$. Using the weak Laplacian comparison inequality (5.6) and (5.12) one deduces that $(t,y) \to \psi_t^\rho(y)$ satisfies (5.11) for $0 < \sigma \leq \tau$ for sufficiently small ρ and all $0 \leq \xi \in D_c(\mathcal{E}_G)$. If we assume also that $\xi \in D(\Delta^G)$ then we may integrate by parts on the left hand side of (5.12), pass to the limit for $\rho \to 0$ and integrate by parts again which yields for the function $\psi_t(y) = h_t(d(x,y))$

$$\int_\sigma^\tau \mathcal{E}(\psi_t,\xi)dt \leq -\langle\psi_\tau,\xi\rangle_m + \langle\psi_\sigma,\xi\rangle_m \quad \forall\, 0 \leq \xi \in D(\Delta^G). \tag{5.13}$$

Standard arguments of Dirichlet form theory now show that for $\xi \in D(\mathcal{E})$ the sequence $\xi_\lambda = \lambda R_\lambda \xi \in D(\Delta^G)$ converges to ξ in the Dirichlet space $(D(\mathcal{E}), \|.\|_1)$ for λ tending to infinity, where R_λ is the $\lambda-$resolvent associated to $(\mathcal{E}, D(\mathcal{E}))$. From the representation $R_\lambda = \lambda\int_0^\infty e^{-\lambda s}P_s ds$ and the fact that the heat semigroup is positivity preserving $\xi \geq 0$ implies $\xi_\lambda \geq 0$. Hence we obtain (5.13) for arbitrary $0 \leq \xi \in D_c(\mathcal{E}_G)$ by approximation.

For $\delta > 0$ let $\psi_t^\delta(y) = (P_\rho^G \psi_t)(y) = \int_G q_\rho^G(y,z)\psi_t(z)m(dz)$ and $q_t^{G,\delta}(y) = q_{t+\delta}^G(x,y)$. Then ψ_t^δ and $q_t^{G,\delta}$ belong to $D_c(\mathcal{E}_G)$, where $q_t^{G,\delta}$ obviously satisfies (5.11) with equality sign, whereas for $\psi_t^\delta(y)$ we find

$$\int_\sigma^\tau \mathcal{E}_G(\psi_t^\delta,\xi)dt = \int_\sigma^\tau \mathcal{E}_G(\psi_t, P_\delta^G \xi)dt$$
$$\leq -\langle \psi_\tau, P_\delta^G\xi\rangle_m + \langle \psi_\sigma, P_\delta^G \xi\rangle_m$$
$$= -\langle \psi_\tau^\delta, \xi\rangle_m + \langle \psi_\sigma^\delta,\xi\rangle_m \quad \forall\, 0 \leq \xi \in D_c(\mathcal{E}(G)).$$

Hence the function $(t,y) \to f_t(y) = \psi^\delta(y) - q_t^{G,\delta}(y)$ satisfies (5.11). As for the initial boundary value $f_0(.)$ we study the behaviour of $\psi_t(.)$ when t tends to zero: as before the Alexandrov convexity implies that there is a radially isometric (i.e. $d(x,y) = \overline{d}(\Lambda_x(x),\Lambda_x(y))$ for all $y \in X$) and non-expanding map $\Lambda_x : (X,d) \mapsto (\mathbb{M}_{n,k},\overline{d})$. Thus

$$\lim_{t\to 0}\int_X h_t(d(x,y))m(dy) = \lim_{t\to 0}\int_{B_r(\Lambda_x(x))} h_t(\overline{d}(\Lambda_x(x),\overline{y}))\Lambda_{x*}m(d\overline{y})$$
$$= \frac{d(\Lambda_{x*}m)}{d\,vol_{\mathbb{M}_{n,k}}}(\Lambda_x(x)) = \lim_{r\to 0}\frac{m(B_r(x))}{b_{n,k}(r)} =: q_0(x)$$

(with $q_0(x) \leq 1$ and $m(\{q_0 < 1\}) = 0$). From this and the standard Gaussian estimates for the heat kernel on smooth manifolds it follows that for any $\zeta \in C_c^0(X)$

$$\lim_{t\to 0}\int_X h_t(d(x,y))\zeta(y)m(dy) = C_x\zeta(x)$$

with $C_x = 1/q_0(x)$, i.e. the function $y \mapsto h_t(d(x,y))$ converges weakly to $C_x\delta_x$ for t tending to zero. From the local $(L^1,1)$-volume regularity of (X,d) it follows that $C_x = 1$ for m-almost every X. Hence let us assume that $C_x = 1$ for the previously fixed x. Then it follows that in fact $f_t = \psi_t^\delta - q_t^{G,\delta}$ converges to $f_0(y) = q_\delta^G(x,y) - q_\delta^G(x,y) = 0$ in $L^2(G,dm)$ for $t \to 0$ and we may apply lemma 5.3 which yields $f_t(.) \leq 0$ m-almost everywhere in G. Since $\delta > 0$ was arbitrary and $P_\delta^G \to 1$ in $L^2(G)$ for $\rho \to 0$ this implies also $\psi_t(y) \leq q_t^G(y)$ for m-almost every $y \in G$. Hence from the continuity of $\psi_t(.) = h_t(d(x,.))$ and $q_t^G(.) = q_t(x,.)$ we may conclude

$$q_t^{k,r}(\overline{x},\overline{y}) = h_t(d(x,y)) \leq q_t^G(x,y) \quad \forall\, y \in B_r(x). \tag{5.14}$$

Since the set $R = \{x \in G | C_x = 1\}$ is dense in G for general $x \in G$ with $\overline{B_r(x)} \subset G$ and $y \in B_r(x)$ we may find an approximating sequence (x_l, y_l) with $d(x_l,y_l) = d(x,y) = \overline{d}(\overline{x},\overline{y})$ such that $y_l \in B_r(x_l) \Subset G$ and $C_{x_l} = 1$, which by the continuity of $q_t^G(.,.)$ establishes (5.14) also in the case $C_x > 1$. Finally, we can offset the assumption $B_r(x) \Subset G$ by considering $B_{r-\epsilon}(x)$ first which gives $(5.14)_{r'}$ for $r' = r - \epsilon$ and fixed $y \in B_{r-\epsilon}(x)$. Using the continuity of the Dirichlet heat kernel $q_t^r(\overline{x},\overline{y})$ on $B_r(\overline{x}) \subset \mathbb{M}_{n,k}$ with respect to r

(which follows from the parabolic maximum principle on $\mathbb{M}_{n,k}$) we may pass to the limit for $\epsilon \to 0$ in the left hand side of (5.14), obtaining

$$q_t^{k,r}(\overline{x}, \overline{y}) = h_t(d(x,y)) \leq q_t^G(x,y) \quad \forall\, y \in B_{r-\epsilon}(x)$$

where $\epsilon > 0$ is arbitrary. Hence, for general $y \in B_r(x)$ the claim follows from the continuity of $q_t^G(x,.)$ by an approximation with $B_{r-1/l}(x) \ni y_l \to y$ for $l \to \infty$. □

5.5 Eigenvalue Comparison

An immediate corollary to the heat kernel comparison theorem is the analogue of Cheng's eigenvalue comparison theorem [Che75]:

Theorem 5.4. *Let (X,d) be a $(L^1,1)$-locally volume regular n-dimensional Alexandrov space with lower curvature bound k and $\dim_r \mathbb{S}_X^{1,1} \leq n-2$. Then for any ball $B_r(x) \subset X$ its first Dirichlet eigenvalue $\lambda_1(B_r(x))$ is bounded from above by*

$$\lambda_1(B_r(x)) \leq \lambda_1^k(r)$$

where $\lambda_1^k(r)$ denote the first Dirichlet eigenvalue for $\overline{B}_r(\overline{0}) \subset \mathbb{M}_{n,k}$.

Proof. This follows directly from the heat kernel comparison theorem and the eigenfunction expansion of the heat kernel on X ([KMS01]) and on $\mathbb{M}_{n,k}$, c.f. [SY94]. □

Corollary 5.2. *Let (X,d) be as above and for $r>0$ let $p \in X$ such that $X_0 := B_r(p) \subset X$. If $\lambda_j(X_0)$ denotes the j-th (counted with multiplicity) Dirichlet eigenvalue of $B_j(p)$ with $0 = \lambda_0(X_0) < \lambda_1(X_0) \leq \lambda_2(X_0) \leq \cdots$. Then*

$$\lambda_j(X_0) \leq \lambda_1^k(\operatorname{diam}(X_0)/2j).$$

This is a consequence of the Max-Min-principle for the higher eigenvalues of the Laplace operator. For further standard results in this direction see [Cha84], chapter III.

5.6 Distance and Short Time Asymptotics

Another consequence of the heat kernel comparison theorem is the extension of Varadhan's formula [Var67], which relates the distance on a manifold to the asymptotic behaviour of the heat kernel for short time, to Alexandrov spaces (X,d):

Proposition 5.1. *Let $G \subset X$ and $q_t^G(.,.)$ be the (Dirichlet of Neumann) heat kernel on G under the same conditions of theorem 5.3. The for all x, $y \in G$*

$$\lim_{t\to 0} 2t \log q_t^G(x,y) = -d^2(x,y). \tag{5.15}$$

Proof. This follows from the Davies' sharp upper Gaussian estimate which persists on local Dirichlet spaces with Poincaré inequality and doubling base measure (cf. [Stu95]). The other inequality follows from (5.10) and (5.15) on manifolds with lower Ricci bound (cf. [Dav89]) applied to $q_t^{k,r}$ on $\mathbb{M}_{n,k}$. □

Remark 5.5. In [Nor96] Norris extends Varadhan's formula to the case of Lipschitz (Riemannian) manifolds with measurable and uniformly elliptic metric tensor (g_{ij}), which rules out the most general Alexandrov spaces. However, since he does not impose any curvature condition on the resulting metric space (X, d) his result is rather complementary and not just a special case of ours.

5.7 Diffusion Process Comparison

In the classical situation where $\Xi \in M$ is some semi-martingale on a smooth Riemannian manifold (M^d, g) and $f \in C^2(M;\mathbb{R})$ the geometric Ito-formula yields for the composite process $f(\Xi)$ the representation

$$d(f \circ \Xi) = \sum_{i=1}^{d} df(\Xi)(Ue_i)dZ^i + \frac{1}{2}\sum_{i,j=1}^{d}(\nabla df)(\Xi)(Ue_i, Ue_j)d[Z^i, Z^j]$$

where $U \in \mathcal{O}(M)$ is the horizontal lift of Ξ onto the orthonormal frame bundle of (M, g) and $Z \in R^d$ is the stochastic anti-development of Ξ (see, for instance, [HT94]). In particular if Ξ is a Brownian Motion on M this formula reduces to

$$d(f \circ \Xi) = df(Ue_i)dW^i + \frac{1}{2}\Delta f(\Xi)dt$$

with $W \in \mathbb{R}^d$ being a Brownian Motion on $\mathbb{R}^d$ and Δ the Laplace-Beltrami Operator on (M, g). In this section a decomposition of the same type will be established for the process $\rho_p(\Xi)$, where Ξ is the Hunt process generated by the Dirichlet form $(\mathcal{E}, D(\mathcal{E}))$ on the Alexandrov space (X, d). We start with two general observations concerning the martingale and zero energy part in the Fukushima decomposition of the Dirichlet process $df(\Xi)$:

Lemma 5.4. *Let $(\mathcal{E}, D(\mathcal{E}))$ be a strongly local symmetric Dirichlet form on some Hilbert space $L^2(X, \sigma, m)$ and $f \in D(\mathcal{E})$. Then for arbitrary g, $h \in L^2 \cap L^\infty(X, m)$ such that $gf^2 \in D(\mathcal{E})$*

$$\mathbb{E}_{h\cdot m}\left(\tilde{g}(\Xi_S)\langle M^{[f]}\rangle_S\right) = \int_0^S \int_X P_t h \cdot P_{S-t} g \mu_{\langle f\rangle} dt$$

where $\langle M^{[f]}\rangle$ is the quadratic variation process of the martingale additive functional part of $df(\Xi)$ and $\mu_{\langle f\rangle}$ is the energy measure associated to f.

Proof. This fact essentially follows from [FOT94], theorem 5.2.3. and lemma 5.1.10. However, we present here an almost self contained proof which requires only a certain familiarity with the concept of energy measures. For $\Delta \in \mathbb{R}$ small let S, T be some fixed positive numbers such that $S > T + \Delta$. For general sufficiently regular g, h the Markov property of Ξ yields

$$\begin{aligned}
&\mathbb{E}_{h\cdot m}(g(\Xi_S)[f(\Xi_{T+\Delta}) - f(\Xi_T)]^2) \\
&= \mathbb{E}_{h\cdot m}(\mathbb{E}_{\Xi_{T+\Delta}}(g(\Xi_{S-(T+\Delta)}))[f(\Xi_{T+\Delta}) - f(\Xi_T))]^2) \\
&= \mathbb{E}_{h\cdot m}((P_{S-(T+\Delta)}g)(\Xi_{T+\Delta})[f(\Xi_{T+\Delta}) - f(\Xi_T))]^2) \\
&= \mathbb{E}_{P_T h\cdot m}((P_{S-(T+\Delta)}g)(\Xi_{\Delta})[f(\Xi_{\Delta}) - f(\Xi_0))]^2).
\end{aligned} \tag{5.16}$$

We want to divide in (5.16) by Δ and pass to the limit for Δ tending to zero. Before doing so one notices that in general for $f, g \in D(\mathcal{E})$

$$\lim_{\Delta\to 0}\frac{1}{\Delta}\mathbb{E}_{h\cdot m}(g(\Xi_\Delta)[f(\Xi_\Delta) - f(\Xi_0)]^2) = \int_X h \cdot g\mu_{\langle f\rangle}. \tag{5.17}$$

If in (5.17) the term $g(\Xi_\Delta)$ was replaced by $g(\Xi_0)$ this would just be the well known coincidence of the energy measure $\mu_{\langle f\rangle}$ and the Revuz-measure $\mu_{\langle M^{[f]}\rangle}$ ([FOT94], lemma 5.3.3.), but in the given form (5.17) can be verified as a consequence of the chain rule for the energy measure, which holds true by the strong locality of $\mathcal{E}$. In a second step one has to verify that whenever hf^2, $g \in D(\mathcal{E})$

$$\begin{aligned}
&\lim_{\Delta\to 0}\frac{1}{\Delta}\mathbb{E}_{h\cdot m}(P_\Delta g(\Xi_\Delta)[f(\Xi_{T+\Delta}) - f(\Xi_T)]^2) \\
&\qquad - \lim_{\Delta\to 0}\frac{1}{\Delta}\mathbb{E}_{h\cdot m}(P_\Delta g(\Xi_\Delta)[f(\Xi_{T+\Delta}) - f(\Xi_T)]^2) \\
&= \lim_{\Delta\to 0}\frac{1}{\Delta}\langle P_\Delta h, [P_\Delta g - g]f^2\rangle_m \\
&\quad - \lim_{\Delta\to 0} 2\langle P_\Delta(fh), [P_\Delta g - g]f\rangle_m + \lim_{\Delta\to 0}\langle P_\Delta(f^2 h), [P_\Delta g - g]\rangle_m \\
&= \mathcal{E}(hf^2, g) - 2\mathcal{E}(hf^2, g) + \mathcal{E}(hf^2, g) = 0.
\end{aligned}$$

Due to these two assertions taking the limit in (5.16) gives

$$\lim_{\Delta\to 0}\frac{1}{\Delta}\mathbb{E}_{h\cdot m}(g(\Xi_S)[f(\Xi_{T+\Delta}) - f(\Xi_T)]^2) = \int_X P_T h \cdot P_{S-T} g\mu_{\langle f\rangle}.$$

Now it is easy to compute the quadratic variation of $M^{[f]}$ because we know that by general Ito theory

$$\begin{aligned}\mathbb{E}_{h\cdot m}(g(\Xi_S)\langle M^{[f]}\rangle_S) &= \lim_{\Delta\to 0}\mathbb{E}_{h\cdot m}(g(\Xi_S)\langle M^{[f]}\rangle_{S-\Delta})\\ &= \lim_{\Delta\to 0}\mathbb{E}_{h\cdot m}\big(g(\Xi_S)\sum_{i=0}^{\lfloor\frac{S}{\Delta}\rfloor-2}[f(\Xi_{(i+1)\Delta}-f(\Xi_{i\Delta})]^2\big)\\ &= \lim_{\Delta\to 0}\sum_{i=0}^{\lfloor\frac{S}{\Delta}\rfloor-2}\big(\Delta\int_X P_{i\Delta}h\cdot P_{S-i\Delta}g\mu_{\langle f\rangle}+O(\Delta)\big)\\ &= \int_0^S\int_X P_th\cdot P_{S-t}g\mu_{\langle f\rangle}dt.\end{aligned}$$

□

Corollary 5.3. *Let $(\mathcal{E},D(\mathcal{E}))$ be a strongly local Dirichlet form defined on $L^2(X,\sigma,m)$ such that the associated semigroup has the Feller property. Then for $f\in D(\mathcal{E})$ the following implication holds:*

$$\big\{\mu_{\langle f\rangle}=m\big\}\Longrightarrow M^{\langle f\rangle}\text{ is a real }P_x\text{-Brownian Motion for all }x\in X.$$

Proof. For $\mu_{\langle f\rangle}=m$ the previous lemma gives

$$\begin{aligned}\mathbb{E}_{h\cdot m}\left(\tilde g(\Xi_S)\langle M^{[f]}\rangle_S\right) &= \int_0^S\langle P_th,P_{S-t}g\rangle_m dt\\ &= S\langle h,P_Sg\rangle_m = S\mathbb{E}_{h\cdot m}(g(\Xi_S)).\end{aligned}$$

By a monotone class argument this implies $\langle M^{[f]}\rangle_S=S$ $P_{h\cdot m}$-almost surely and thus for all $x\in X$ also P_x-almost surely, since we can let tend $h\cdot m$ tend to δ_x when utilizing the Feller property of P_t. Levy's characterization of Brownian Motion then yields the claim. □

Lemma 5.5. *Let the $(\mathcal{E},D(\mathcal{E}))$ and Ξ be as in lemma 5.4 and f in $D(\mathcal{E})$. Then for the CAF of zero energy $A^{[f]}$ belonging to $(f(\Xi_t)-f(\Xi_0))_{t\geq 0}$ one has*

$$\mathbb{E}_{h\cdot m}(\tilde g(\Xi_S)A_S^{[f]}) = -\int_0^S\mathcal{E}(P_thP_{S-t}g,f)dt\quad\forall\, h,g\in L^\infty(X,m)\cap D(\mathcal{E}),$$

where P_t is the semigroup generated by $(\mathcal{E},D(\mathcal{E}))$ and $\tilde\zeta$ is a quasi-continuous version of ζ.

Proof. We proceed as in the proof of lemma 5.4 by using the Markov property and the additivity of $A^{[f]}$ to obtain for $S>T+\Delta$

$$\mathbb{E}_{h\cdot m}(\tilde g(\Xi_S)(A_{T+\Delta}^{[f]}-A_T^{[f]}) = \mathbb{E}_{P_Th\cdot m}((P_{S-(T+\Delta)}\tilde g)(\Xi_\Delta)A_\Delta^{[f]})).$$

For $f = R_1\eta$ with $\eta \in L^2$ the right hand side equals

$$\langle P_T h, \mathbb{E}_{.}\Big(P_{S-(T+\Delta)}\tilde{g}(\Xi_\Delta)\int_0^\Delta f(\Xi_u) + \eta(\Xi_u)du\Big)\rangle_m$$

and from this representation and the continuity of $\tilde{g}(\Xi)$ it is clear that

$$\begin{aligned}\lim_{\Delta\to 0}\frac{1}{\Delta}\mathbb{E}_{h\cdot m}(\tilde{g}(\Xi_S)(A^{[f]}_{T+\Delta} - A^{[f]}_T) &= \langle P_T h \cdot P_{S-T} g, f + \eta\rangle_m \\ &= -\mathcal{E}(P_T h \cdot P_{S-T} g, f).\end{aligned}$$

For general $f \in D(\mathcal{E})$ one establishes this result by the usual approximation argument (compare [FOT94], thm. 5.2.4.). As before we now can compute

$$\begin{aligned}\mathbb{E}_{h\cdot m}(\tilde{g}(\Xi_S)A^{[f]}_S) &= \lim_{\Delta\to 0}\sum_{i=0}^{\lfloor\frac{S}{\Delta}\rfloor -1}\Delta\,\frac{1}{\Delta}\mathbb{E}_{h\cdot m}(\tilde{g}(\Xi_S)(A^{[f]}_{(i+1)\Delta} - A^{[f]}_{i\Delta})) \\ &= -\lim_{\Delta\to 0}\Big(\sum_{i=0}^{\lfloor\frac{S}{\Delta}\rfloor -1}\Delta\;\mathcal{E}(P_{i\Delta}h\cdot P_{S-i\Delta}g, f) + O(\Delta)S\Big) \\ &= -\int_0^S \mathcal{E}(P_t h\cdot P_{S-t} g, f)dt.\end{aligned}$$

□

Lemma 5.6. *Let (X, d) be an n-dimensional locally $(L^1, 1)$-volume regular Alexandrov space with lower curvature bound k and $(\mathcal{E}, D(\mathcal{E}))$ the canonical intrinsic Dirichlet form on $L^2(X, \mathcal{B}(X), m)$, where m is the n-dimensional Hausdorff-measure. Then for each $p \in X$ the distance function $\rho_p(.) = d(p, .) : X \mapsto \mathbb{R}$ has the energy measure $\mu_{\langle\rho_p\rangle} = m$.*

Proof. From the fact that $\mathcal{E}^r \to \mathcal{E}$ for $r \to 0$ pointwise on the set of Lipschitz functions on (X, d), which serves as a common core for the forms $\mathcal{E}^r$ and $\mathcal{E}$, it follows that

$$\mu^r_{\langle f\rangle} \to \mu_{\langle f\rangle} \text{ for } r \to 0 \text{ weakly in the sense of Radon measures.}$$

Now obviously

$$\mu^r_{\langle f\rangle}(dx) = \int_{B^*_r(x)}\left(\frac{f(x) - f(y)}{d(x,y)}\right)^2 m_r(dy)m_r(dx)$$

and in the special case $f = d_p$ the first variation formula for the distance function on Alexandrov spaces ([OS94], thm 3.5) says that for fixed p, $x \in X$ any choice of segments γ_{xy} for $y \in X$ the formula

$$\rho_p(x) - \rho_p(y) = d(x,y)\cos\inf_{\gamma_{px}}\sphericalangle pxy + o_x(d(x,y))$$

obtains, where the infimum is taken over all possible choices of segments γ_{px} connecting p with x. Furthermore, for fixed p the cut locus C_p has measure zero (ibid., prop. 3.1), which implies that ρ_p is differentiable in m-a.e. $x \in X$. Moreover, using $\Sigma_x \simeq R^d$ for m-a.e. $x \in X$, the volume regularity and the weak Riemannian structure of (X, d) we find that

$$\lim_{r\to 0} \int_{B_r^*(x)} \left(\frac{f(x)-f(y)}{d(x,y)}\right)^2 m_r(dy) \frac{1}{\sqrt{m(B_r(x))}} = \fint_{B_1(0_x)} \langle \gamma'_{xp}, z\rangle_{g_x}^2 dz = 1 \text{ for } m\text{-a.e. } x \in X,$$

which yields the claim by Lebesque's dominated convergence theorem. □

Proposition 5.2. *Let Ξ be the diffusion proceess generated by the canonical intrinsic Dirichlet Form $(\mathcal{E}, D(\mathcal{E}))$ on an n-dimensional locally $(L^1, 1)$-volume regular Alexandrov space with lower curvature bound k. Then for any $p \in X$ the process $\rho_p(\Xi)$ satisfies the stochastic differential inequality*

$$d\rho_p(\Xi) \leq dB_t + (n-1)(\ln S_k)' \circ \rho_p(\Xi)dt \quad P_x\text{-a.s.} \tag{5.18}$$

for all $x \in X$, where B_t is a real-valued standard P_x-Brownian Motion.

Proof. Due to corollary 5.3 and lemma 5.6 we know that the MCAF-part in the Ito decomposition of $\rho_p(\Xi)$ is a real-valued Brownian motion. As for the CAF part $A^{[\rho_p]}$ of zero energy we apply lemma 5.5 to $f = \rho_p$ for arbitrary nonnegative $h, \zeta \in L^\infty(X, m) \cap D(\mathcal{E})$ and the weak laplacian comparison (thm. 5.2), which gives

$$\begin{aligned}
\mathbb{E}_{h\cdot m}(\tilde{\zeta}(\Xi_s)A_s^{[f]}) &= -\int_0^s \mathcal{E}(P_t h \cdot P_{s-t}\zeta, \rho_p)dt \\
&\leq (n-1)\int_0^s \langle P_t h \cdot P_{s-t}\zeta, (\ln S_k)' \circ \rho_p\rangle_m dt \\
&= (n-1)\mathbb{E}_{h\cdot m}\left[\tilde{\zeta}(\Xi_s)\int_0^s (\ln S_k)' \circ \rho_p(\Xi_t)dt\right].
\end{aligned}$$

Letting tend $h \cdot m$ to δ_x and using the monotone class theorem this means that $A_s^{[f]} \leq (n-1)\int_0^s (\ln S_k)' \circ \rho_p(\Xi_t)dt$ P_x-a.s. and thus the proof is complete. □

Since the radial process of a Brownian Motion on the Model space satisfies (5.18) with "$\leq$" replaced by "$=$" one immediately obtains the

Theorem 5.5 (Brownian Motion Comparison Principle). *Brownian Motion on X is slower than on $\mathbb{M}_{n,k}$ in the following sense: for any $x \in X$ let Ξ_x be the canonical diffusion process on X starting in x then*

$$\mathbb{E}\left[\rho_x(\Xi_{x,t})\right] \leq \mathbb{E}\left[\overline{\rho}_{\overline{x}}(\overline{\Xi}_{\overline{x},t})\right] \quad \forall t \geq 0,$$

where $\overline{\Xi}_{\overline{x}}$ is Brownian Motion on $\mathbb{M}_{n,k}$ starting in some $\overline{x} \in \mathbb{M}_{n,k}$ and $\overline{\rho}_{\overline{x}}$ is the distance function of $\overline{x}$ on $\mathbb{M}_{n,k}$.

6 Appendix

6.1 A - Remark on Coupling by Dirichlet Forms

In the Euclidean situation the coupling process for Brownian motion can also be constructed via a suitable Dirichlet form: the coupling process is characterized by its generator $L_c^{\mathbb{R}^d}$ on $C^2(\mathbb{R}^d \times \mathbb{R}^d; \mathbb{R})$

$$(L_c^{\mathbb{R}^d} F)(x,y) = \langle A(x,y) : \mathrm{Hess}\, F(x,y)\rangle$$

with coefficients $A = (\alpha_{ij})$ given by

$$(\alpha_{ij})(x,y) = \frac{1}{2}\begin{pmatrix} \mathbf{1}_{\mathbb{R}^d} & \mathbf{1}_{\mathbb{R}^d} - 2e_{xy}\otimes e_{xy} \\ \mathbf{1}_{\mathbb{R}^d} - 2e_{xy}\otimes e_{xy} & \mathbf{1}_{\mathbb{R}^d} \end{pmatrix} \text{ and } e_{xy} = \frac{x-y}{\|x-y\|}$$

(cf. [LR86]), for which one obtains the decomposition $(\alpha_{ij})(x,y) = \sigma(x,y)\sigma(x,y)^t$ with

$$\sigma(x,y) = \begin{pmatrix} I & -I \\ I & I \end{pmatrix}\begin{pmatrix} M_{xy} & \\ & M_{xy} \end{pmatrix}\begin{pmatrix} 0 & & & & & \\ & 1 & & & & \\ & & \ddots & & & \\ & & & 1 & & \\ & & & & 1 & \\ & & & & & 0 \\ & & & & & & \ddots \\ & & & & & & & 0 \end{pmatrix} = \begin{pmatrix} 0 & b_2 & \dots & b_d & b_1 & 0 & \dots & 0 \\ 0 & b_2 & \dots & b_d & -b_1 & 0 & \dots & 0 \end{pmatrix}$$

where $M_{xy} = (b_1, b_2, \ldots, b_d)(x,y)$ is a matrix of orthonormal basis vectors of $\mathbb{R}^d$ depending smoothly on $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d \setminus \{x = y\}$ such that $b_1(x,y) = e_{xy}$. Obviously M_{xy} cannot be extended smoothly across $\{x = y\}$. So the coupling operator on $\mathbb{R}^d \times \mathbb{R}^d$ is degenerate (i. e. not strongly elliptic) and has discontinuities on the diagonal $\{x = y\}$. In order to apply the theory of Dirichlet forms for the construction of the coupling process one needs to find a measure m on $\mathbb{R}^d \times \mathbb{R}^d$ such that L_c is m-symmetric. If $m = e^{\beta}(x)dx$ then

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \phi L_c\psi \cdot dm = \int_{\mathbb{R}^d\times\mathbb{R}^d} \phi(x,y)\, \langle A(x,y) : \mathrm{Hess}\psi(x,y)\rangle\, e^{\beta(x,y)} dxdy$$

$$\stackrel{!}{=} \int_{\mathbb{R}^d\times\mathbb{R}^d} \langle \nabla\phi, A\nabla\psi\rangle_{\mathbb{R}^d\times\mathbb{R}^d}\, dm \quad \forall \phi,\psi \in C_0^\infty(\mathbb{R}^d \times \mathbb{R}^d)$$

is equivalent to

$$\mathrm{div} A + A \cdot \nabla\beta = 0$$

in the weak sense, where $(\,\mathrm{div}A)_i = \alpha_{ij,j}$. In the case given above one computes

$$\mathrm{div}A(x,y) = \frac{(d-2)}{\|x-y\|^2}\begin{pmatrix} (x-y) \\ (y-x) \end{pmatrix} = A \cdot v \text{ with } v = \frac{(d-2)}{\|x-y\|^2}\begin{pmatrix} (x-y) \\ (y-x) \end{pmatrix}$$

such that one obtains as one particular solution to the previous equation

$$\beta(x,y) = -(d-2)\ln\|x-y\|.$$

Consequently one defines the coupling pre-Dirichlet form on $L^2(\mathbb{R}^{2d}, \frac{1}{\|x-y\|^{(d-2)}} dx \otimes dy)$ by

$$\mathcal{E}_c(u,v) = \frac{1}{2} \int\limits_{\mathbb{R}^{2d}} \Big[\langle \nabla_x u, \nabla_x v \rangle + \langle \nabla_y u, \nabla_y v \rangle + \langle \nabla_x u, \nabla_y v \rangle + \langle \nabla_y u, \nabla_x v \rangle$$
$$-2(\langle \nabla_x u, e_{xy} \rangle \langle \nabla_y v, e_{xy} \rangle + \langle \nabla_y u, e_{xy} \rangle \langle \nabla_x v, e_{xy} \rangle)\Big] \frac{1}{\|x-y\|^{(d-2)}} dxdy$$

for all $u, v \in C_0^\infty(\mathbb{R}^{2d})$ such that $\mathcal{E}_c(u,v) < \infty$. Note that even if the generator of $\mathcal{E}_c$ can be considered as an extension of the Laplacian on each component $\mathbb{R}^d$ this is not true for the corresponding Dirichlet forms.

On a general Riemannian manifold it is much harder to find a symmetrizing measure for the coupling operator L_c which is determined by a choice of $\Phi(.,.)$ as in section 2.1. From equation (2.4) one obtains the representation of L_c on smooth functions $F : M \times M \to \mathbb{R}$

$$L_c F(x,y) = \frac{1}{2} \langle \Sigma_\Phi : \text{Hess}\, F \rangle_{T^2_{(x,y)}(M\times M)} (x,y)$$

where one defines the bilinear form Σ_Φ by

$$\Sigma_\Phi((U_1^x, V_1^y), (U_2^x, V_2^y)) = \Big\langle \Phi^{-1}(x,y)(U_1^x, V_1^y), \begin{pmatrix} \mathbf{1}_{\mathbb{R}^d} & \mathbf{1}_{\mathbb{R}^d} \\ \mathbf{1}_{\mathbb{R}^d} & \mathbf{1}_{\mathbb{R}^d} \end{pmatrix} \Phi^{-1}(x,y)(U_2^x, V_2^y) \Big\rangle_{\mathbb{R}^d \times \mathbb{R}^d}.$$

As before the symmetry of L_c with respect to the volume measure with exponential density $e^{\beta(x,y)} m(dx) \otimes m(dy)$ on $M \times M$ is guaranteed by the condition

$$\text{div}(\Sigma_\Phi) + A_{\Sigma_\Phi} \cdot \nabla\beta \overset{!}{=} 0 \tag{6.1}$$

where we used the identification $T^{(p,r)}(M \times M) \cong T^{p+r}(M \times M)$: $A_\Sigma \cong \Sigma$ and the definition $\text{div}\, T = \sum_i (\nabla_{e_i} T)(e_i, .)$ for the divergence for a $(1,r)$-Tensor T with some orthonormal basis $\{e_i, i = 1, \ldots, 2d\}$ of $T(M \times M)$. Now let us assume that Φ satisfies conditions $(*)$ and $(**)$ of section 2, i.e. setting $S_{yx} = \Phi_1 \circ \Phi_2^{-1}(x,y) : T_y(M) \to T_x(M)$ and $S_{xy} = \Phi_2 \circ \Phi_1^{-1}(x,y)$ accordingly, the requirement is

$$S_{xy} = T_{xy} \circ R_{xy}$$

where T_{xy} is parallel translation of $T_x M$ to $T_y M$ along the curve γ_{xy} and R_{xy} is the linear isometry on T_x by reflecting the $\dot\gamma_{xy}$-direction. Then we obtain the representation for Σ_Φ

$$\begin{aligned} \Sigma_\Phi((U_1^x, V_1^y), (U_2^x, V_2^y)) &= \Big\langle (U_1^x, V_1^y), \begin{pmatrix} \mathbf{1}_{T_x M} & T_{yx} \circ R_{yx} \\ T_{xy} \circ R_{xy} & \mathbf{1}_{T_y M} \end{pmatrix} (U_2^x, V_2^y) \Big\rangle_{T_x M \times T_y M} \\ &= \Big\langle (U_1^x, V_1^y), \begin{pmatrix} \mathbf{1}_{T_x M} & T_{yx} \\ T_{xy} & \mathbf{1}_{T_y M} \end{pmatrix} + 2 \begin{pmatrix} & \dot\gamma_{xy} \otimes \dot\gamma_{yx} \\ \dot\gamma_{yx} \otimes \dot\gamma_{xy} & \end{pmatrix} (U_2^x, V_2^y) \Big\rangle_{T_x M \times T_y M}, \end{aligned}$$

i.e. $\Sigma_\Phi = \Sigma_T + 2\Sigma_\Gamma$. Computation shows that

$$\begin{aligned} \text{div}\, \Sigma_\Gamma(x,y) &= \begin{pmatrix} (\Delta_y^M d_x(y) - 1/d(x,y))\dot\gamma_{xy} \\ (\Delta_x^M d_y(x) - 1/d(x,y))\dot\gamma_{yx} \end{pmatrix} \\ \text{div}\, \Sigma_T(x,y) &= \begin{pmatrix} \text{div}_y T_{yx} \\ \text{div}_x T_{xy} \end{pmatrix}. \end{aligned}$$

For the computation of $\text{div}_y T_{yx}$ fix some normal basis $\{e_i\}$ of T_yM, i.e. $\nabla_{e_i} e_j = 0$ and $\langle e_i, e_j\rangle = \delta_{ij}$. Then for some $V \in T_xM$

$$\langle \text{div}_y T_{yx}, V\rangle_x = \sum_i e_i(\langle T_{yx} e_i, V\rangle_x) = \sum_i \frac{\partial}{\partial s}_{|s=0} \langle V, T_{\sigma_i(s)x}\dot{\sigma}_i(s)\rangle_x$$

where σ_i is a (w.l.o.g. geodesic) integral curve of e_i near y

$$= \sum_i \frac{\partial}{\partial s}_{|\substack{s=0\\t=1}} \langle V_i(s,t), /\!\!/_{(0,t)}^{c_i(s,.)}(\dot{\sigma}_i(s))\rangle_x$$

where $c(s,t) : [-\epsilon, \epsilon[\times[0,1] \to M$ is the geodesic variation of $c_i(0,.) = \gamma_{yx}(.)$ with fixed end point x induced from $c_i(.,o) = \sigma_i(.)$ and $V_i(.,.)$ is the vector field over $c_i(.,.)$ obtained by transporting V parallely along $c_i(s,t)$ first with respect to the t and then with respect to s. Hence the summands equal

$$\frac{\partial}{\partial s}_{|s=0} \int_0^1 \frac{\partial}{\partial t}\langle \cdots, \cdots\rangle dt - \frac{\partial}{\partial s}_{|\substack{s=0\\t=0}} \langle V_i(s,t), /\!\!/_{(0,t)}^{c_i(s,.)}(\dot{\sigma}_i(s))\rangle_x$$

where the second term vanishes, since $V(.,0)$ is parallel and σ_i was chosen to be a geodesic.

$$= \int_0^1 \langle \nabla^{c^*}_{\frac{\partial}{\partial s}} \nabla^{c^*}_{\frac{\partial}{\partial t}} V_i(s,t), /\!\!/_{(0,t)}^{c_i(s,.)}(\dot{\sigma}_i(s))\rangle_{x|s=0} dt$$
$$+ \int_0^1 \langle \nabla^{c^*}_{\frac{\partial}{\partial t}} V_i(s,t), \nabla^{c^*}_{\frac{\partial}{\partial s}} /\!\!/_{(0,t)}^{c_i(s,.)}(\dot{\sigma}_i(s))\rangle_{x|s=0} dt$$

where the second term vanishes, again since $V(.,0)$ is parallel

$$= \int_0^1 \langle \nabla^{c^*}_{\frac{\partial}{\partial t}} \nabla^{c^*}_{\frac{\partial}{\partial s}} V_i(s,t), /\!\!/_{(0,t)}^{c_i(s,.)}(\dot{\sigma}_i(s))\rangle_{x|s=0} dt$$
$$+ \int_0^1 \langle R(c_* \frac{\partial}{\partial t}, c_* \frac{\partial}{\partial s}) V(s,t), /\!\!/_{(0,t)}^{c_i(s,.)}(\dot{\sigma}_i(s))\rangle_{x|s=0} dt$$

where the first integral is zero because $\nabla^{c^*}_{\frac{\partial}{\partial s}} V(s,t) = 0$

$$= \int_0^1 \langle R(\dot{\gamma}_{yx}(t), J_i(t)) /\!\!/_{(0,1-t)}^{\gamma_{xy}} V, /\!\!/_{(0,t)}^{\gamma_{yx}} e_i\rangle_x dt$$
$$= -\Big\langle \int_0^1 /\!\!/_{(t,1)}^{\gamma_{yx}} \big\{ R(\dot{\gamma}_{yx}(t), J_i(t)) /\!\!/_{(0,t)}^{\gamma_{yx}} e_i \big\} dt, V\Big\rangle_x$$

with the Jacobi-field J_i along γ_{yx} obtained from the geodesic variation $c_i(.,.)$.
Hence the the equation (6.1) for $\nabla\beta = (\nabla_x\beta, \nabla_y\beta)$ becomes

$$\begin{pmatrix} \nabla_x\beta + T_{yx}\nabla_y\beta + 2\dot\gamma_{xy}\langle\dot\gamma_{yx}, \nabla_y\beta\rangle \\ \nabla_y\beta + T_{xy}\nabla_x\beta + 2\dot\gamma_{yx}\langle\dot\gamma_{xy}, \nabla_x\beta\rangle \end{pmatrix} \stackrel{!}{=} - \begin{pmatrix} 2(\Delta_y^M d_x(y) - 1/d(x,y))\dot\gamma_{xy} + \mathrm{div}_y T_{yx} \\ 2(\Delta_x^M d_y(x) - 1/d(x,y))\dot\gamma_{yx} + \mathrm{div}_x T_{xy} \end{pmatrix},$$

and hence we are left with the question under which conditions on (M, g) this equation for β is solvable.

Conversely one might try the analogue of the the Euclidean coupling form in the manifold case by defining

$$\mathcal{E}_c(u,v) = \frac{1}{2}\int\limits_{M\times M} \Big[\langle\nabla_x u, \nabla_x v\rangle_x + \langle\nabla_y u, \nabla_y v\rangle_y + \langle\nabla_x u, T_{yx}\nabla_y v\rangle_x + \langle\nabla_y u, T_{yx}\nabla_x v\rangle_y$$
$$+2(\langle\nabla_x u, \dot\gamma_{xy}\rangle_x\langle\nabla_y v, \dot\gamma_{yx}\rangle_y + \langle\nabla_y u, \dot\gamma_{yx}\rangle_y\langle\nabla_x v, \dot\gamma_{xy}\rangle_x)\Big] e^{\beta(x,y)} m(dx)m(dy)$$

or more general

$$\mathcal{E}_c(u,v) = \frac{1}{2}\int\limits_{M\times M} \begin{pmatrix}\nabla_x u \\ \nabla_y u\end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{pmatrix} \begin{pmatrix}\nabla_x v \\ \nabla_y v\end{pmatrix} e^{\beta(x,y)} m(dx)m(dy)$$

where $a_{ij} = a_{ij}(x,y)$ are linear maps between the corresponding tangent spaces, such that the associated operators on functions of the form $u = f \otimes 1$ ($\Leftrightarrow \nabla_y u = 0$) is

$$L_c(f\otimes 1)(x,y) = \langle a_{11}, \mathrm{Hess}\, f\rangle_x + \langle\nabla_x f, \mathrm{div}_x a_{11} + a_{11}\nabla_x\beta + \mathrm{div}_y a_{12} + a_{12}\nabla_y\beta\rangle_x. \quad (6.2)$$

If $\mathcal{E}_c$ is to generate a coupling $\overline{X} \in M \times M$ between the marginal processes $X_1 = \pi_1\overline{X}$ and $X_2 = \pi_2\overline{X}$ then the corresponding condition for the generators L_c, L_1 and L_2 reads as follows

$$L_c(f\otimes 1) = (L_1 f)\otimes 1 \text{ and } L_c(1\otimes g) = 1\otimes(L_2 g) \quad (6.3)$$

(equivalently $L_c(f\circ\pi_1) = (L^1 f)\circ\pi_1$ and $L_c(g\circ\pi_2) = (L^2 g)\circ\pi_2$) which leads to the requirement that the operator in equation (6.2) should not depend on y, i.e.

$$a_{11}(x,y) \stackrel{!}{=} a_{11}(x) \text{ and } V(x,y) := a_{11}\nabla_x\beta + \mathrm{div}_y a_{12} + a_{12}\nabla_y\beta \stackrel{!}{=} V(x)$$

that is, if we have fixed a diffusion process on each factor of $M\times M$ with generator L_1 and L_2 respectively then the condition (6.3) gives first order equations for some appropriate choice of a_{ij} and β in terms of the data L_1 and L_2.

6.2 B - More about the Geometry of Alexandrov Spaces

Rough Dimension and Natural Coordinates

Definition 6.1. *The rough dimension of a bounded subset $V \subset X$ in a metric space is defined as*

$$\dim_r V = \inf \left\{ \alpha > 0 \,\middle|\, \lim_{\epsilon \to 0} \epsilon^\alpha \beta_V(\epsilon) = 0 \right\}$$

where $\beta_V(\epsilon)$ is the cardinality of a maximal ϵ-net in V, i.e. the cardinality of a maximal subset $\{x_i \,|\, i \in I\} \subset V$ such that $d(x_i, x_j) \geq \epsilon$.

Note that the rough dimension is an upper bound for the Hausdorff dimension. If (X, d) is an Alexandrov space the convexity of geodesic triangles implies that the rough dimension is locally constant. This follows from the observation that for fixed p, $q \in X$ with $d(p, q) \leq \delta$ and a given ϵ-net $\{x_i\} \subset V_p$ of a neighborhood $V_p \ni p$ the image points $\{x_i'\}$ of $\{x_i\}$ under a geodesic contraction with respect to q will be an ϵ'-net of V_q where V_q is some appropriately chosen neighbourhood of q and where $0 < \epsilon' \to 0$ for $\delta \to 0$. This argument can be made rigorous to yield the estimate $\dim_r V_p \leq \dim_r V_q$ for q sufficiently close to p and V_p and V_q sufficiently small. Thus $n = \dim_r X = \lim_{V_p \to \{p\}} \dim_r V_p$ is well defined, i.e. independent of p.

Definition 6.2. *For $p \in X$ a system of tuples $\{(a_{-i}, a_{+i} \,|\, i = 1, \ldots, m\}$ is called a (m, δ)-strainer at p iff*

$$\overline{\sphericalangle} a_i p a_j \geq \begin{cases} \pi/2 - \delta & \text{if } i \neq \pm j \\ \pi - \delta & \text{if } i = -j. \end{cases}$$

In this case the map

$$\phi_{\{a_i\}} : X \supset V_p \to \mathbb{R}^m, \quad q \to (d(a_1, q), \ldots, d(a_m, q))$$

is called the natural coordinate function with base $\{a_i \,|\, i = 1, \ldots, m\}$.

The idea about a strainer in a point p is that if p admits a system of tuples $\{(a_{-i}, a_{+i}) \in X \times X \,|\, i = 1, \ldots, m\}$ such that the corresponding broken geodesic segments $\gamma_{a_{-i}p} * \gamma_{pa_{+i}}$ are almost geodesic in p and are also almost mutually perpendicular then the endpoints $\{a_i\}$ naturally define a coordinate system on X in a neighborhood of p, provided m is the maximal number of such points for given p. In fact, if $\{(a_{-i}, a_{+i}) \,|\, i = 1, \ldots, m\}$ is an (m, δ)-strainer at p with $\delta \leq 1/2m$ and such that there is no $(m + 1, 4\delta)$-strainer at p then $\phi_{\{a_i\}}$ provides a bi-Lipschitz homeomorphism of some neighborhood U_p of p onto an open subset of $\mathbb{R}^m$ (equipped with the metric stemming from the maximum-norm), see thm. 5.4. in [BGP92]. Since in that case the Hausdorff-dimension, the topological and the rough dimension of U_p are equal to n, the *maximal strain index* $Ind_s(p)$ at p is, in fact, constant on X, which is used for the definition of $\dim(X)$.

Weak Riemannian Structure ([OS94])

For a given (n,δ)-strainer $\{(a_{-i},a_{+i})\}$ at $p \in X$ and sufficiently small δ let $U_{\{a_i\}}$ be some neighborhood of p which is mapped (bi-Lip) homeomorphically onto an open set in $\mathbb{R}^n$ ([BGP92]), the set $V_{\{a_i\}} = U_{\{a_i\}} \cap \bigcap_{i=1,\dots,n} C_{a_i}$ with $C_{a_i} = \{y \in X \,|\exists! \gamma_{ya_i}\}$ is the set where all geodesic segments from a point in $x \in U_{\{a_i\}}$ to the base points a_i are unique. Then the function

$$g_{\{a_i\}} : U_{\{a_i\}} \to \mathbb{R}^{n\times n}_{sym}, \quad x \to (\cos \sphericalangle(\gamma_{xa_i}, \gamma_{xa_j}))_{ij}$$

is uniquely defined and continuous on $V_{\{a_i\}}$ ([OS94]). Starting from this observation one can show that the system

$$\Big\{\phi_{\{a_i\}}, g_{\{a_i\}}, U_{\{a_i\}}, V_{\{a_i\}}\Big\}_{\{\{(a_{-i},a_{+i})\} \text{ is a } (n,\delta)\text{-strainer at } p,\, p\in X\setminus S_X,\, \delta\leq 1/2n\}}$$

is *weak* C^1*-Riemannian structure on* $X\setminus S_X \subset X$, i.e. it possesses the following properties

- for all $\{a_i\}$, $V_{\{a_i\}} \subset U_{\{a_i\}}$, $U_{\{a_i\}} \subset X$ open.
- every $\phi_{\{a_i\}}$ maps $U_{\{a_i\}}$ homeomorphically onto an open set in $\mathbb{R}^n$
- $\{V_{\{a_i\}}\}$ is a covering of $X \setminus S_X$.
- if two maps $\phi_{\{a_i\}}$ and $\phi_{\{b_i\}}$ satisfy $V_{\{a_i\}} \cap V_{\{b_i\}} \neq \emptyset$ then $\phi_{\{a_i\}} \circ \phi^{-1}_{\{b_i\}}$ is differentiable on $\phi_{\{b_i\}}\big(V_{\{a_i\}} \cap V_{\{b_i\}} \cap (X \setminus S_X)\big)$
- for each $\{a_i\}$ the map $g_{\{a_i\}} \circ \phi^{-1}_{\{a_i\}}$ is continuous on $\phi_{\{a_i\}}\big(V_{\{a_i\}} \cap (X \setminus S_X)\big)$
- for any $x \in V_{\{a_i\}} \cap V_{\{b_i\}} \cap X \setminus S_X$

$$g_{\{b_i\}}(x) = \frac{d}{dx}\big(\phi_{\{a_i\}} \circ \phi^{-1}_{\{b_i\}}\big)^t (\phi_{\{b_i\}}(x)) \cdot g_{\{a_i\}}(x) \cdot \frac{d}{dx}\big(\phi_{\{a_i\}} \circ \phi^{-1}_{\{b_i\}}\big)(\phi_{\{b_i\}}(x)).$$

Thus the difference between a Riemannian and a weak Riemannian structure is basically that the topological types of $X \setminus S_X$ and $V_{\{a_i\}}$ are undetermined. In particular $X \setminus S_X$ need not be a manifold.

Excess measure on Alexandrov Surfaces ([Ale55, Mac98])

The excess of a geodesic triangle Δ in an Alexandrov space (X,d) is defined by

$$e(\Delta) = \sum_{i=1}^{3} \sphericalangle_i^{\text{int}}(\Delta) - \pi$$

where $\sphericalangle_i^{\text{int}}(\Delta)$ denote the interior angles of Δ. The following observation is the basis for the definition of the excess measure for general subsets in X. Let D be an open polygonal domain in a two-dimensional Alexandrov space, i.e. the boundary ∂D is given by a set of geodesic curves γ_i, $i = 1, \dots, I$, and let Δ_j, $j = 1, \dots, J$ be a decomposition of D into geodesic triangles, then

$$\sum_{j=1}^{J} e(\Delta_j) + \sum_{k=1}^{K} (2\pi - L(\Sigma_{p_k})) = 2\pi\chi(D) - \sum_{i=1}^{I} (\pi - \beta_i)$$

where p_k, $k = 1, \dots K$ are the vertices of the decomposition $D = \bigcup \Delta_j$ lying in (the interior of) D and $L(\Sigma_{p_k})$ denotes the length of their space of directions ($\dim \Sigma_{p_k} = 1$), $\chi(D)$ is the Euler-Poincaré-characteristic of D in X and β_i, $i = 1, \dots, I$ denotes the interior angle between the geodesics γ_{i-1} and γ_i. Since in this equality the right hand side does not depend on the decomposition $\{\Delta_j\}$ the total excess of D is defined by

$$e(D) = \sum_{j=1}^{J} e(\Delta_j) + \sum_{k=1}^{K} (2\pi - L(\Sigma_{p_k}))$$

This defines an finitely additive measure on the class $\mathcal{D}$ of subsets of X which can be represented as finite unions or set-theoretic differences of geodesic triangles, segments and points. By the usual procedure $e(.)$ is extended as an outer measure on X which is countably additive on the σ-algebra generated by $\mathcal{D}$. Finally, by choosing the Δ_j sufficiently small and by comparing them with the almost congruent geodesic triangles in $\mathbb{R}^2$ obtained from the natural coordinate functions $\phi_{\{a_i\}}$ of the previous paragraphs one can deduce the central estimate (cf. [Mac98])

$$e(D) \geq \kappa m(D) \quad \forall\, D \in \mathcal{D}$$

which implies in particular that $e(.) + |\kappa| m(.)$ extends to a positive outer measure on X.

Bibliography

[AD99] Lars Andersson and Bruce K. Driver. Finite-dimensional approximations to Wiener measure and path integral formulas on manifolds. *J. Funct. Anal.*, 165(2):430–498, 1999.

[Ale55] A. D. Alexandrow. *Die innere Geometrie der konvexen Flächen.* Akademie-Verlag, Berlin, 1955.

[Ale57] Alexander D. Alexandrow. Über eine Verallgemeinerung der Riemannschen Geometrie. *Schr. Forschungsinst. Math.*, 1:33–84, 1957.

[BBI01] Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A course in metric geometry.* American Mathematical Society, Providence, RI, 2001.

[BGP92] Yuri Burago, Misha Gromov, and Grigorij Perel'man. A. D. Alexandrov spaces with curvature bounded below. *Russ. Math. Surv.*, 47(2):1–58, 1992.

[Bil68] Patrick Billingsley. *Convergence of probability measures.* John Wiley & Sons Inc., New York, 1968.

[Blu84] Gilles Blum. A note on the central limit theorem for geodesic random walks. *Bull. Austral. Math. Soc.*, 30(2):169–173, 1984.

[BNR93] V. N. Berestovskiĭ, I. G. Nikolaev, and Yu. G. Reshetnyak. *Geometry. IV.* Springer-Verlag, Berlin, 1993. Nonregular Riemannian geometry, Encyclopaedia of Mathematical Sciences, 70.

[CH98] Jingyi Chen and Elton P. Hsu. Gradient estimates for harmonic functions on manifolds with Lipschitz metrics. *Canad. J. Math.*, 50(6):1163–1175, 1998.

[Cha84] Isaac Chavel. *Eigenvalues in Riemannian geometry.* Academic Press Inc., Orlando, Fla., 1984.

[Che75] Shiu Yuen Cheng. Eigenvalue comparison theorems and its geometric applications. *Math. Z.*, 143(3):289–297, 1975.

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

[Dav89] Edward B. Davies. *Heat kernels and spectral theory.* Cambridge University Press, Cambridge, 1989.

[Dur96] Richard Durrett. *Stochastic calculus.* CRC Press, Boca Raton, FL, 1996.

[EK86] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes.* John Wiley & Sons Inc., New York, 1986. Characterization and convergence.

[Éme89] Michel Émery. *Stochastic calculus in manifolds.* Springer-Verlag, Berlin, 1989. With an appendix by P.-A. Meyer.

[FOT94] Masatoshi Fukushima, Yōichi Oshima, and Masayoshi Takeda. *Dirichlet forms and symmetric Markov processes.* Walter de Gruyter & Co., Berlin, 1994.

[Gro99] Misha Gromov. *Metric structures for Riemannian and non-Riemannian spaces.* Progress in Mathematics (Boston, Mass.). Birkhäuser, 1999.

[HT94] Wolfgang Hackenbroch and Anton Thalmaier. *Stochastische Analysis.* B. G. Teubner, Stuttgart, 1994.

[Ito50] Kiyosi Ito. Stochastic differential equations in a differentiable manifold. *Nagoya Math. J.*, 1:35–47, 1950.

[IW89] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic differential equations and diffusion processes.* North-Holland Publishing Co., Amsterdam, 1989.

[Jør75] Erik Jørgensen. The central limit problem for geodesic random walks. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 32:1–64, 1975.

[JS87] Jean Jacod and Albert N. Shiryaev. *Limit theorems for stochastic processes.* Springer-Verlag, Berlin, 1987.

[Ken86] Wilfrid S. Kendall. Nonnegative Ricci curvature and the Brownian coupling property. *Stochastics*, 19(1-2):111–129, 1986.

[KMS01] Kazuhiro Kuwae, Yoshiroh Machigashira, and Takashi Shioya. Sobolev spaces, Laplacian and heat kernel on Alexandrov spaces. *Math. Z.*, 238(2):269–316, 2001.

[KS88] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus.* Springer-Verlag, New York, 1988.

[KS98] Kazuhiro Kuwae and Takashi Shioya. Sobolev spaces and Dirichlet spaces over maps between metric spaces. Preprint, 1998.

[KS01] Kazuhiro Kuwae and Takashi Shioya. On generalized measure contraction property and energy functionals over Lipschitz maps. *Potential Anal.*, 15(1-2):105–121, 2001.

[Lia88] Ming Liao. Hitting distributions of small geodesic spheres. *Ann. Probab.*, 16(3):1039–1050, 1988.

[Lin97] Fang Hua Lin. Analysis on singular spaces. In *Collection of papers on geometry, analysis and mathematical physics*, pages 114–126. World Sci. Publishing, River Edge, NJ, 1997.

[LR86] Torgny Lindvall and L. C. G. Rogers. Coupling of multidimensional diffusions by reflection. *Ann. Probab.*, 14(3):860–872, 1986.

[LZ95] M. Liao and W.A. Zheng. Radial part of Brownian motion on a Riemannian manifold. *Ann. Probab.*, 23(1):173–177, 1995.

[Mac98] Yoshiroh Machigashira. The Gaussian curvature of Alexandrov surfaces. *J. Math. Soc. Japan*, 50(4):859–878, 1998.

[Mos94] Umberto Mosco. Composite media and asymptotic Dirichlet forms. *J. Funct. Anal.*, 123(2):368–421, 1994.

[Nor96] James Norris. Small time asymptotics for heat kernels with measurable coefficients. *C. R. Acad. Sci. Paris Sér. I Math.*, 322(4):339–344, 1996.

[OS94] Yukio Otsu and Takashi Shioya. The Riemannian structure of Alexandrov spaces. *J. Diff. Geom.*, 39(3):629–658, 1994.

[Ots98] Yukio Otsu. Almost everywhere existence of second differentiable structure of Alexandrov spaces. 1998. Preprint.

[Pau99] Gustav Paulik. Anwendungen des Couplings von Flußprozessen auf Riemanschen Mannigfaltigkeiten. 1999. Diploma Thesis, Friedrich-Wilhelms-Universität, Bonn.

[Per98] Gregorij Perelman. DC structure on Alexandrov spaces. 1998. Preprint.

[Pet86] Stefan Peters. *Konvergenz Riemannscher Mannigfaltigkeiten.* Universität Bonn Mathematisches Institut, Bonn, 1986. Dissertation, Friedrich-Wilhelms-Universität, Bonn, 1985.

[Pet97] Anton Petrunin. Applications of quasigeodesics and gradient curves. In *Comparison geometry (Berkeley, CA, 1993–94)*, pages 203–219. Cambridge Univ. Press, Cambridge, 1997.

[Pet98] Anton Petrunin. Parallel transportation for Alexandrov space with curvature bounded below. *Geom. Funct. Anal.*, 8(1):123–148, 1998.

[Pin79] Mark A. Pinsky. Stochastic Riemannian geometry. In *Probabilistic analysis and related topics, Vol. 1, 199-236* . 1979.

[RS75] Michael Reed and Barry Simon. *Methods of modern mathematical physics. II. Fourier analysis, self-adjointness.* Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1975.

[She93] Zhong Min Shen. A regularity theorem for Alexandrov spaces. *Math. Nachr.*, 164:91–102, 1993.

[Shi93] Katsuhiro Shiohama. *An introduction to the geometry of Alexandrov spaces.* Seoul National University Research Institute of Mathematics, Global Analysis Research Center, Seoul, 1993.

[Stu95] Karl-Theodor Sturm. Analysis on local Dirichlet spaces. II. Upper Gaussian estimates for the fundamental solutions of parabolic equations. *Osaka J. Math.*, 32(2):275–312, 1995.

[Stu96] Karl-Theodor Sturm. Analysis on local Dirichlet spaces. III. The parabolic Harnack inequality. *J. Math. Pures Appl. (9)*, 75(3):273–297, 1996.

[Stu98] Karl-Theodor Sturm. Diffusion processes and heat kernels on metric spaces. *Ann.-Probab.*, 26(1):1–55, 1998.

[Sun84] Toshikazu Sunada. Geodesic flows and geodesic random walks. In *Geometry of geodesics and related topics (Tokyo, 1982)*, pages 47–85. North-Holland, Amsterdam, 1984.

[SV79] Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes.* Springer-Verlag, Berlin, 1979. Grundlehren.

[SY94] Richard M. Schoen and Shing Tung Yau. *Lectures on differential geometry.* International Press, Cambridge, MA, 1994.

[Var67] S. R. Srinivasa Varadhan. On the behavior of the fundamental solution of the heat equation with variable coefficients. *Comm. Pure Appl. Math.*, 20:431–455, 1967.

[Yam96] Takao Yamaguchi. A convergence theorem in the geometry of Alexandrov spaces. In *Actes de la Table Ronde de Géométrie Différentielle (Luminy, 1992)*, pages 601–642. Soc. Math. France, Paris, 1996.

[Yau75] Shing Tung Yau. Harmonic functions on complete Riemannian manifolds. *Comm. Pure Appl. Math.*, 28:201–228, 1975.