From Stochastic Parallel Transport to Harmonic Maps

Euroconference in Mathematics

Dirichlet Forms and their applications in Geometry and Stochastics

Anogia, Crete, June 22-28, 1997

From Stochastic Parallel Transport to Harmonic Maps

Wilfrid S Kendall,
Statistics,
University of Warwick

April 28, 1998

Chapter 1

Introduction

1.1 Aim and objectives.

The aim of this account is to show how to do stochastic differential geometry: how one can work effectively and clearly with random processes which live on geometric objects, particularly by using semimartingale stochastic calculus. We shall do this by:

- explaining how one can carry out semimartingale stochastic calculus effectively in an intrinsic way;
- studying the interplay between particular semimartingales (such as Brownian motion on a manifold) and geometry;
- developing a probabilistic approach to the nonlinear elliptic theory of *harmonic maps*.

1.2 History and motivation

We begin with three motivations for stochastic differential geometry.

1. **Transition densities for geometric statistics.** Prompted by the study of datasets of directions of remanent magnetism in rock flows, and also by intellectual curiosity, Fisher [48] introduced the study of probability and statistical inference for distributions on the sphere. In his introduction he remarked that "Any topological framework requires the development of a theory of errors of characteristic and appropriate form". This comment is an excellent early expression of

one of the motivations for work in stochastic differential geometry: it can guide us in the choice of families of probability densities to be used in geometric statistical problems (see [97] for example).

Fisher's work here did not directly involve Brownian motion, but Roberts and Ursell [123] took up some probabilistic aspects of Fisher's paper, particularly concerning limiting behaviour of random walks on the sphere (hence Brownian motion on the sphere, $BM(S^2)$), and connected them with the geometric theory of harmonic spaces. (It is interesting to note in passing that the subsequent development of harmonic spaces has seen at least one substantial contribution using a probabilistic proof [111].)

2. **The stochastic approach to harmonic functions.** It has been known for a long time [21, 68] that there is an intimate and fruitful relationship between, on the one hand, probability (in particular Brownian motion) and, on the other hand, harmonic functions $f : \mathbb{M} \to \mathbb{R}$, stationary values of the Dirichlet form

$$\mathcal{E}(f) \quad = \quad \int_{\mathbb{M}} |\nabla f|^2 \, \mathrm{d}\mu \, . \tag{1.1}$$

In particular Brownian motion provides a intuitive and fertile interpretation of the integral representation for bounded harmonic functions using harmonic measure, delivering a stochastic mean-value formula. Many probabilists have used the sample path properties of Brownian motion to derive new and remarkable results for harmonic functions and Euclidean heat kernels. For a simple but effective example see [9, 84].

In the early 1980's Eells and Lemaire [40] asked whether one could apply probabilistic methods to harmonic maps $F : \mathbb{M} \to \mathbb{N}$ (where $\mathbb{N}$ is a curved Riemannian manifold), stationary values of the *nonlinear* Dirichlet form

$$\mathcal{E}(F) \quad = \quad \int_{\mathbb{M}} \| \operatorname{grad} F\|^2 \, \mathrm{d}\mu \, . \tag{1.2}$$

The difficulty here is that the relationship between Brownian motion and Eq. (1.1) depends crucially on the image $\mathbb{N}$ having the linear structure of the real line, which supplies the tools of the mean value formula and conditional expectation. Curvature intervenes for general $\mathbb{N}$ and destroys this relationship. For a considerable time it was quite unclear how one might establish a relationship in the nonlinear case, and particularly unclear whether such a relationship could ever deliver a way of constructing harmonic maps using stochastic methods.

3. **Intrinsic geometry of diffusions.** In several areas of applied probability one is presented with a diffusion defined by some statistical problem, and it is required to find some succinct expression of the diffusion, whether to deduce a statistical distribution theory of "characteristic errors", as in Fisher's quotation above, or whether simply to present the diffusion in a clear and conceptual way. In particular this issue arises in the *statistical theory of shape*, where one needs to find out about diffusions related to the *shape* (equivalence class under similarities) of configurations of finitely many points diffusing in a Euclidean space. The machinery of stochastic differential geometry does good service in dealing with this kind of task, particularly in combination with implementations of stochastic calculus in computer algebra packages (see for example [72, 83, 86, 97]).

Of course these are not the only reasons for studying random processes on manifolds - for example the idea of stochastic development described below is central to the very exciting recent growth of stochastic analysis for spaces of loops on manifolds. However our treatment will start from the ideas and questions described above, and we will focus in particular on extending to harmonic maps the successful relationship between probability and harmonic functions.

1.3 Stochastic calculus formalism

The approach to stochastic differential geometry which we describe here is based on semimartingales and Stratonovich calculus. Good introductions to the basic theory at varying levels of sophistication are to be found in [113, 124, 70].

Recall the *Doob-Meyer decomposition*: a real-valued random process with continuous sample paths is a *semimartingale* if it possesses a "signal-plus-noise" decomposition

$$X = V + M \tag{1.3}$$

where V is an adapted continuous process whose sample paths are of locally bounded variation (the "signal") and M is a continuous local martingale (the "noise"). The Doob-Meyer decomposition asserts that the *drift part* V and the *noise part* M of the process X are unique, if X has been identified as possessing such a decomposition. Notice that latent in the above definition is reference to an underlying filtration $\{\mathfrak{F}_t : t \in [0, \infty)\}$ (required to give meaning to the terms "adapted" and "local martingale"): we fix this filtration once and for all and assume that it supports sufficiently many Brownian motions for all our geometric purposes.

In a slight variation on conventional notation, we shall include in the family of semimartingales those random processes which admit a decomposition of the form of Eq. (1.3) up to a random stopping time at which they become unbounded.

We also note here that the theory of stochastic calculus extends to include random processes whose sample paths have jumps, so long as the paths remain left-continuous with right-hand limits. This extension has also been pressed into service in stochastic differential geometry [37] and indeed in the probabilistic study of harmonic maps [19, 116], but space constraints do not permit us to give this much attention here.

It is convenient to write equations involving semimartingales using the notation of *stochastic differentials*, for example:

$$\mathrm{d}_I X \quad = \quad \mathrm{d}V + \mathrm{d}_I M\,. \tag{1.4}$$

The suffix "$_I$" stands for "Itô", to signal that d_I is no ordinary differential. (This is important in the context of the concise notation of differential geometry, where otherwise the symbol "d" would be overloaded to a quite unintelligible extent!) We shall see many such equations below. They are to be interpreted by integration against integrands which are bounded *predictable* functions (limits of continuous random functions which do not depend on the future). Itô [62] gives an algebraic formalism in terms of modules over bounded predictable integrands, which is of importance in understanding *symbolic Itô calculus*, the computer algebra implementation [87, 93] of stochastic calculus mentioned above.

We need to refer to a couple of operations which can be performed on stochastic differentials to yield classical differentials: using the Doob-Meyer decomposition (1.3) we define the *Drift* operator

$$\mathtt{Drift}\, \mathrm{d}_I X \quad = \quad \mathrm{d}V \tag{1.5}$$

and the *Bracket* or *quadratic covariation*

$$\mathrm{d}\,[\,X_1, X_2\,] \quad = \quad \mathrm{d}\,[\,M_1, M_2\,]\,. \tag{1.6}$$

Thus $\mathtt{Drift}(\,\mathrm{d}_I X)$ isolates the "trend" V from the Doob-Meyer decomposition $X = M + V$: if the drift vanishes then the semimartingale X is actually a local martingale. The bracket $[\,M, M\,]$ of a local martingale M can be defined as the unique continuous increasing process such that

$$M^2 - [\,M, M\,]$$

is a continuous local martingale, and $[M_1, M_2]$ can be defined by polarization. However a good intuitive appreciation of the bracket is to think of $\mathrm{d}[X, X]$ as measuring the amount of randomness in the infinitesimal increment $\mathrm{d}_I X$ of the semimartingale X.

Both operations are appropriately module-linear: for example

$$\mathrm{d}[\int f \, \mathrm{d}X_1, X_2] = \mathrm{d}[\int f \, \mathrm{d}M_1, M_2] = \int f \, \mathrm{d}[M_1, M_2] = \int f \, \mathrm{d}[X_1, X_2]$$

whenever f is a bounded predictable function.

Intrinsic formulations of stochastic calculus must take account of the second-order nature of the *Itô formula*, stochastic calculus's fundamental theorem:

$$\mathrm{d}_I f(X) = f'(X) \, \mathrm{d}_I(X) + \frac{1}{2} f''(X) \, \mathrm{d}[X, X] \, .$$

This holds for twice-differentiable f, though there is an important extension to the case of convex f.

It is immediate from the Itô formula that without further structure one can give no meaningful intrinsic notion of "(local) martingale" on a smooth manifold (though we *can* define a local martingale M on a vector space $\mathbb{V}$ simply by requiring M to satisfy the condition, $\ell(M)$ to be a local martingale whenever $\ell : \mathbb{V} \to \mathbb{R}$ is linear). For general manifolds Schwartz [125, 126] noted the following fundamental point: the Itô formula shows that semimartingales *are* well-defined on smooth manifolds:

Definition 1.1 (Semimartingale on a manifold): *A continuous random process X on a manifold $\mathbb{M}$ is a* semimartingale *if and only if its composition $f(X)$ with any C^2 function $f : \mathbb{M} \to \mathbb{R}$ is a real-valued semimartingale.*

As Itô's formula shows, Itô calculus is second-order and so does not fit well to the demands of conventional differential geometry. However *Stratonovich calculus* is first-order: if we define

$$\int Y \, \mathrm{d}_S X = \int Y \, \mathrm{d}_I X + \frac{1}{2} [Y, X]$$

for semimartingales X and Y then an application of the Itô formula shows that

$$\mathrm{d}_S f(X) = f'(X) \, \mathrm{d}_S X$$

whenever f is C^3. Again this can be interpreted in terms of modules, but this time using multiplication by semimartingales rather than by bounded predictable functions. As a consequence of the above formula we *can* define semimartingales intrinsically on manifolds with no particular effort, so long as we use Stratonovich differentials instead of Itô differentials, since the coefficients then transform contravariantly as they should do.

The typical procedure for use of stochastic calculus in stochastic differential geometry is therefore as follows: we define a semimartingale on a manifold using a stochastic differential equation written in terms of Stratonovich differentials. This of course relies on the existence and uniqueness theory for stochastic differential equations (*sde*): typically the coefficients are at least locally Lipschitz, so we may assume the solution is uniquely defined at least up to a so-called explosion time, when the semimartingale first exits the domain of definition of the sde.

It is worth noting that sde solutions are defined up to an event of probability zero. We may therefore choose to interpret an sde for a manifold-valued semimartingale in terms of coordinate patches, in which case we can piece together a countable number of coordinate-patch solutions without danger. However we can alternatively interpret the solution in terms of an extrinsic construction: an embedding of the manifold in high-dimensional Euclidean space, for example. Both approaches are equally valid, and define the same process up to an event of probability zero, so it is a matter of taste and convenience which one might choose in any specific case.

However, as we will see, in practice it is often unnecessary to descend to specific coordinate representations. Given a semimartingale X on a smooth manifold $\mathbb{M}$, and a smooth function $f : \mathbb{M} \to \mathbb{R}$, and some extra geometric structure, we shall in the next chapter see how to write down an expression for the Itô differential $\mathrm{d}_I f(X)$ in terms of invariant quantities which can be related directly to the geometry of the manifold. (In the event that it *is* necessary to work with a specific coordinate representation, much labour and time can be saved, and the risk of computational errors can be reduced, by using the computer algebra technique of symbolic Itô calculus mentioned above).

The organization of this account is as follows. In Chapter 2 we describe the intrinsic approach to semimartingales on manifolds afforded by stochastic development and parallel transport. We use this to introduce a "zoo" of manifold-valued semimartingales, and a geometric version of the fundamental Itô formula. We also introduce the relationship with convexity and harmonic maps which is the recurrent theme of the article, and summarize important geometric inequalities for variations of geodesic distance. This sets the scene for an analysis in Chapter 3 of the interaction between distance and semimartingales (particularly Γ-martingales and Brownian motion), which in turn allows us to derive a number of elegant probabilistic proofs of results about harmonic maps. In Chapter 4 the distance-based analyses are developed into considerations of *coupling* (when can two random processes meet?) which allow significant development of the previous harmonic map results, and which finally in Chapter 5 deliver fully probabilistic treatment of the small-image Dirichlet problem for harmonic maps, resolving the open question alluded to in the discussion surrounding Eq. (1.2).

Before moving on to the main part of the exposition we should comment on the literature of expositions and approaches to random processes (particularly semimartingales) on manifolds. The original construction of Brownian motion on a manifold was either by "patching" Itô stochastic differential equations (working from coordinate patch to coordinate patch as in [61, 51, 67]) or by using semigroup methods (it is hard to trace an original reference for this, but [123] is interesting). The use of Dirichlet forms directly to define Brownian motion on a manifold is an immediate consequence of the more general work required to construct a diffusion once one has specified a Dirichlet form, and is covered elsewhere in this volume. An attractive probabilistic development of the semigroup approach is to regard Brownian motion and more general processes as solutions to a *martingale problem*: see [129]. We have mentioned the extrinsic construction, which builds Brownian motion on a manifold by solving a conventional stochastic differential equation in the ambient Euclidean space of an (isometric) embedding of the manifold: [128] describes this for a sphere while [121, 134, 124] develop it for the general case. The approach taken below is to develop the Stratonovich approach using stochastic parallel transport and the frame bundle of the manifold to solve parallelization problems and indeed to supply a highly useful representation: this has its origins in work by Eells and Elworthy and Malliavin (see [41, 109]: see also [37, 38, 64, 119]). Finally, previous expositions and reports by the author on the subject of semimartingale stochastic differential geometry include [78, 80, 81, 92].

Before closing this introductory section, it is a pleasure to acknowledge the thoughtful and careful comments of Dr. Huiling Le of Nottingham University, who read an early draft of this article.

Chapter 2

Stochastic differential geometry

2.1 Intrinsic approach to semimartingales

Given the above definition of semimartingale on a manifold, it is immediate that a rather trivial example is given by $\gamma(B)$, where $\gamma : \mathbb{R} \to \mathbb{M}$ is a geodesic on the Riemannian manifold $\mathbb{M}$ (or more generally any smooth curve on a general smooth manifold $\mathbb{M}$) and B is a real-valued Brownian motion (or indeed more generally any continuous semimartingale). This follows by applying the Itô formula to $f(\gamma)$ when $f : \mathbb{M} \to \mathbb{R}$ is a smooth function.

In order to remove non-stochastic complications, we will assume in the following that all our Riemannian manifolds are *complete*, which is to say that their geodesics can be indefinitely extended.

Although the above example is geometrically trivial, it is fundamental to our ability to do stochastic calculus in a geometrically invariant and elegant way.

Less trivial examples are given by Brownian motion $\mathrm{BM}(\mathbb{M})$ on a Riemannian manifold $\mathbb{M}$. If we choose temporarily to define such a Brownian motion as a process X satisfying the *martingale problem*:

$$f(X_t) - \frac{1}{2}\int_0^t \Delta f(X_s) ds \qquad \text{is a martingale},$$

whenever f is a smooth function of compact support and Δ is the Laplace-Beltrami operator on a Riemannian manifold $\mathbb{M}$, then it is immediate that X is a semimartingale on $\mathbb{M}$.

The reader is reminded that the conventional definition of a semimartingale has been extended above to include processes which are semimartingales up to random stopping times. This allows us to include in our class of manifold-valued semimartingales those Brownian motions which can "explode" off the manifold in finite time (Riemannian manifolds with "exploding" Brownian motions are said to be *stochastically incomplete*).

Combining these two ideas in a generalization, note that $F(X)$ is a semimartingale on $\mathbb{N}$ whenever $F : \mathbb{M} \to \mathbb{N}$ is a smooth map and X is $\mathrm{BM}(\mathbb{M})$.

2.1.1 Lifting to the frame bundle

The martingale problem approach referred to above can be generalized to treat a wide range of semimartingales on smooth manifolds in a manner which is more-or-less intrinsic. However it is more appealing to the intuition, and indeed more rewarding in further applications, to work out how to use concepts of differential geometry to "flatten out" a manifold-valued semimartingale X, viewing it as the solution to a system of (Stratonovich) stochastic differential equations driven by some flat semimartingale Z (that is to say, a semimartingale living in a reference Euclidean space $\mathbb{R}^m$ of the same dimension as the manifold).

If one wishes to do this then the first step must be to relate the differential of Z to the differential of X. *Formally* the differential of Z lives in Euclidean space $\mathbb{R}^m$ while the differential of X lives in the tangent space $T_X\mathbb{M}$ to $\mathbb{M}$ at X (though this has no proper mathematical meaning unless in the exceptional case when X is a differentiable function of time). So we need to relate the two formal differentials to each other using a linear isometry

$$\Xi : \mathbb{R}^m \quad \to \quad T_X\mathbb{M}\,.$$

To make proper sense of this, we should somehow ensure that Ξ is itself a semimartingale, because then we can write down the Stratonovich differential equation

$$\Xi\,\mathrm{d}_S Z \quad = \quad \mathrm{d}_S X\,.$$

Thus we must view Ξ itself as a manifold-valued semimartingale in the *orthonormal frame bundle*, the smooth manifold $\mathrm{O}(\mathbb{M})$ of *orthonormal frames* for $\mathbb{M}$.

In order for this definition to succeed, we must give a rule for the evolution of the semimartingale Ξ. It is clear what must be done: the only possible way to drive Ξ is to use Z or equivalently X, and therefore what is needed is a way to lift the differential of X up to the frame bundle so as to drive Ξ. Thus we need to supply a linear map

$$H_\xi : T_{\pi\xi}\mathbb{M} \quad \to \quad T_\xi \,\mathrm{O}(\mathbb{M})\,,$$

where $\pi : \mathrm{O}(\mathbb{M}) \to \mathbb{M}$ is the obvious projection map sending $\xi : \mathbb{R}^m \to T_x\mathbb{M}$ to the point $x \in \mathbb{M}$ at which its image space is based.

Clearly not every such map H will be satisfactory. We must be sure to maintain the relationship $\pi(\Xi) = X$, since otherwise the differential $\Xi\,\mathrm{d}_S Z$ will cease to be relevant for X. Differentiating, it follows that we must be sure (whatever the behaviour of the flat semimartingale Z and hence whatever $\mathrm{d}_S X$ might be) that $(\mathrm{d}\pi)H_\xi\,\mathrm{d}_S X = \mathrm{d}_S X$. Applying this to the case when Z is a smooth function of time, we find

$$(\mathrm{d}\pi)H_\xi \quad = \quad \text{identity map on } T_{\pi\xi}\mathbb{M}\,.$$

A further constraint is imposed by a natural requirement of equivariance under rotations of the initial frame Ξ_0 and the driving semimartingale Z. Together these two constraints exactly comprise the requirement that the map H be the *horizontal lift map* for an *orthonormal connection.* (This will be no surprise to geometers: the above arguments specialize in the smooth case to the Cartan path-based approach to a connection!).

We now impose an additional condition of *zero-torsion*, less pleasant to discuss from our stochastic point of view, (Zero-torsion essentially means that the resulting connection is *symmetric*: if used to differentiate one vector-field U by another vectorfield V then essentially the same result is obtained as if the rôles of U,V were reversed.) A well-known result of geometry is that there is one and only one such orthonormal connection H for each Riemannian manifold, the *Levi-Civita connection.* Since this map H is all that is required to relate our manifold-valued semimartingale X to a flat semimartingale Z, and since H turns out to be a smooth map, we can apply the theory of stochastic differential equations to arrive at the following:

Theorem 2.1 (Equations of stochastic development and parallel transport): *To each manifold-valued semimartingale X in a Riemannian manifold $\mathbb{M}$ we can associate a* stochastic anti-development *Z in $\mathbb{R}^m$ and a* stochastic parallel transport *Ξ in $O(\mathbb{M})$ such that*

$$\begin{aligned} \mathrm{d}_S X &= \Xi\,\mathrm{d}_S Z\,, \\ \mathrm{d}_S \Xi &= H_\Xi\,\mathrm{d}_S X\,. \end{aligned} \tag{2.1}$$

Note furthermore that Z, Ξ are unique when the initial frame Ξ_0 is prescribed (and depend equivariantly on this choice); and on the other hand X can be defined by prescribing the initial condition Ξ_0 (and hence X_0) and the driving flat semimartingale Z.

We make three remarks here.

Remark 1 *Why* anti-*development? Because the classical notion of development is that the curve Z on the reference space is* developed *onto a curve X on the manifold.*

Remark 2 *In fact the semimartingale X is defined by the stochastic anti-development only up to its* explosion time, *the random time at which X leaves all compact subsets of $\mathbb{M}$. It is immediate from stochastic differential equation theory that Ξ (hence X also) is defined up to the random time at which Ξ leaves all compact subsets on $O(\mathbb{M})$; because of the linearity of the stochastic differential system in the frame directions one can show that Ξ will not explode before X.*

Remark 3 *The above construction also works for smooth manifolds which are not Riemannian, but for which a smooth connection has been defined. However one must then use the full general linear frame bundle $GL(\mathbb{M})$ rather than the frame bundle of isometries (which of course is not defined for a non-Riemannian manifold!).*

Thus the notion of stochastic development and parallel transport allow us to "flatten out" a manifold-valued semimartingale. This provides an attractive route to the definition of different kinds of semimartingales, which we will now discuss. We will see later that the use of stochastic development goes far beyond this.

2.1.2 The semimartingale zoo

It is now easy to categorize whole families of manifold-valued semimartingales: simply use the familiar flat-space categorizations on their stochastic anti-developments!

1. *Geodesic motion.* Consider the straight-line Euclidean motion $Z_t = v \cdot t$, for a fixed velocity vector $v \in \mathbb{R}^m$. The corresponding semimartingale $X = \gamma$ on $\mathbb{M}$ is the geodesic $\gamma : [0, \infty) \to \mathbb{M}$ with initial tangent vector $\Xi_0(v)$: the stochastic differential system for stochastic development converts to the classical Cartan system

$$\begin{aligned} \mathrm{d}\gamma/\,\mathrm{d}t &= \xi(v)\,, \\ \mathrm{d}\xi/\,\mathrm{d}t &= H_\xi\,\mathrm{d}\gamma/dt = H_\xi\xi(v)\,. \end{aligned} \tag{2.2}$$

   Here we use lower-case symbols γ, ξ to represent *non-random* processes (which is to say, non-random functions of time).

It is convenient here to introduce notation for *left-invariant vector-fields* on $\mathrm{O}(\mathbb{M})$:

$$L_v(\xi) \quad = \quad H_\xi \xi v$$

so that

$$(\,\mathrm{d}\pi)L_v(\xi) \quad = \quad (\,\mathrm{d}\pi)H_\xi \xi v \quad = \quad \xi v$$

and

$$L_v g(\xi) \quad = \quad \left[\frac{\mathrm{d}}{\mathrm{d}u} g(\xi(u))\right]_{u=0}$$

so that the Cartan system can be written

$$\begin{aligned} \mathrm{d}\gamma/\,\mathrm{d}t &= \xi(v)\,, \\ \mathrm{d}\xi/\,\mathrm{d}t &= L_v \xi\,. \end{aligned} \tag{2.3}$$

2. *Brownian geodesic.* In a mild generalization, we can replace the straight-line motion by a scalar Brownian motion B along a straight line, $Z_t = v \cdot B_t$, to obtain

$$\begin{aligned} \mathrm{d}_S X &= (\Xi(v))\,\mathrm{d}_S B\,, \\ \mathrm{d}_S \Xi &= H_\Xi\,\mathrm{d}_S X\,. \end{aligned} \tag{2.4}$$

It is easily checked that the above equations are solved by $\Xi_t = \xi_{B_t}$, $X_t = \gamma(B_t)$ with ξ, γ as given in the first example. So the stochastic system Eq. (2.4) for Brownian geodesics can be integrated; this is indeed a rather trivial example! However the following observation will be useful later: if $g : \mathrm{O}(\mathbb{M}) \to \mathbb{R}$ is smooth then the above Stratonovich system can be applied to $g(\Xi)$ to yield the *scalar* stochastic differential equation

$$\mathrm{d}_S g(\Xi) \quad = \quad (L_v g)\,(\Xi)\,\mathrm{d}_S B\,. \tag{2.5}$$

We can convert this to Itô form: using the above equation applied in a further iteration to the smooth function $L_v g$ we can expand the drift term $\frac{1}{2}\,\mathrm{d}\,[\,(L_v g)\,(\Xi), B\,]$ of the Itô stochastic differential equation to find

$$\mathrm{d}_I g(\Xi) \quad = \quad (L_v g)\,(\Xi)\,\mathrm{d}_I B + \frac{1}{2}\,(L_v L_v g)\,(\Xi)\,\mathrm{d}t\,.$$

This specializes when $g = f \circ \pi$ is the lift of a smooth function f on $\mathbb{M}$ to yield

$$\mathrm{d}_I f(X) \quad = \quad \langle \operatorname{grad} f(X), \Xi v\rangle\,\mathrm{d}_I B + \frac{1}{2}\operatorname{Hess} f(X)(\Xi v, \Xi v)\,\mathrm{d}t\,.$$

We note in passing that one can also define a (local) martingale geodesic, by replacing the Brownian motion B with a local martingale M. Then Eq. 2.5 converts to

$$\mathrm{d}_S g(\Xi) = (L_v g)(\Xi)\,\mathrm{d}_S M\,, \tag{2.6}$$

expanding to

$$\mathrm{d}_I g(\Xi) = (L_v g)(\Xi)\,\mathrm{d}_I M + \frac{1}{2}(L_v L_v g)(\Xi)\,\mathrm{d}\,[\,M, M\,]\,. \tag{2.7}$$

3. *General semimartingale.* The same idea can be applied to obtain a *geometric Itô formula* for a general semimartingale X with stochastic anti-development Z and stochastic parallel transport Ξ. Working first with a smooth $g : \mathrm{O}(\mathbb{M}) \to \mathbb{R}$, we apply the stochastic differential system for stochastic development to find a scalar Stratonovich stochastic differential equation for $g(\Xi)$:

$$\mathrm{d}_S g(\Xi) = \sum_i (L_{v_i} g(\Xi))\;\mathrm{d}_S Z^i$$

where the sum is over an orthonormal basis (v_i) of the reference space $\mathbb{R}^m$, and $Z = \sum_i v_i Z^i$. Following the method described for Brownian and martingale geodesics, we can iterate this once for the smooth function $L_{v_i} g$ and so convert to Itô form

$$\mathrm{d}_I g(\Xi) = \sum_i (L_{v_i} g(\Xi))\;\mathrm{d}_I Z^i + \frac{1}{2}\sum_i\sum_j \left(L_{v_j} L_{v_i} g(\Xi)\right)\;\mathrm{d}\left[\,Z^i, Z^j\,\right]\,.$$

When $g = f \circ \pi$ this reduces to the following concise form.

Theorem 2.2 (Geometric Itô formula): *Suppose that $f : \mathbb{M} \to \mathbb{R}$ is a smooth function, and that X is a semimartingale with stochastic anti-development Z and stochastic transport Ξ. Then*

$$\mathrm{d}_I f(X) = \langle \mathit{grad}\, f(X), \Xi\rangle\,\mathrm{d}_I Z + \frac{1}{2}\,\mathit{trace\,Hess}\, f(X)(\Xi,\Xi)\,\mathrm{d}\,[\,Z, Z\,]\,. \tag{2.8}$$

Here

$$\langle \mathit{grad}\, f(X), \Xi\rangle\,\mathrm{d}_I Z = \sum_i \langle \mathit{grad}\, f(X), \Xi v_i\rangle\,\mathrm{d}_I Z^i\,,$$

$$\mathit{trace\,Hess}\, f(X)(\Xi,\Xi)\,\mathrm{d}\,[\,Z, Z\,] = \sum_i \mathit{Hess}\, f(X)(\Xi v_i, \Xi v_i)\,\mathrm{d}\left[\,Z^i, Z^i\,\right]\,.$$

where the v^i form an orthonormal basis.

Thus the apparently trivial notion of a Brownian geodesic leads directly to a fundamental formula for all manifold-valued semimartingales! Notice that there is the possibility of choosing the basis (v_i) adaptively, depending on the past of X, so as to diagonalize the stochastic differential of the matrix-valued bracket process $[Z, Z]$: we take this up below in the important special case of Γ-martingales.

Remark 4 *Notice that $\langle \operatorname{grad} f(X), \Xi\rangle$ defines a random time-varying (indeed, predictable) linear map from $\mathbb{R}^m$ to $\mathbb{R}$, where m is the dimension of the manifold $\mathbb{M}$. Thus the Itô differential $\langle \operatorname{grad} f(X), \Xi\rangle \, \mathrm{d}_I Z$ makes intrinsic sense.*

4. *Brownian motion.* In the case when the stochastic anti-development is a Euclidean Brownian motion B on $\mathbb{R}^m$ we **define** the resulting X to be Brownian motion BM($\mathbb{M}$) on the manifold $\mathbb{M}$. The geometric Itô formula now tells us that this agrees with our previous provisional definition using the martingale problem:

$$\begin{aligned}
\mathrm{d}_I f(X) &= \langle \operatorname{grad} f(X), \Xi\rangle \, \mathrm{d}_I B + \frac{1}{2} \operatorname{trace} \operatorname{Hess} f(X)(\Xi, \Xi) \, \mathrm{d}\,[B, B] \\
&= \langle \operatorname{grad} f(X), \Xi\rangle \, \mathrm{d}_I B + \frac{1}{2} \sum_i \operatorname{Hess} f(X)(\Xi v_i, \Xi v_i) \, \mathrm{d}t \\
&= \langle \operatorname{grad} f(X), \Xi\rangle \, \mathrm{d}_I B + \frac{1}{2} \sum_i \left(L_{v_i} L_{v_i} f\right)(X) \, \mathrm{d}t \\
&= \langle \operatorname{grad} f(X), \Xi\rangle \, \mathrm{d}_I B + \frac{1}{2} \Delta f(X) \, \mathrm{d}t \,,
\end{aligned} \tag{2.9}$$

using the fact that different coordinates of a Euclidean Brownian motion have zero bracket with each other (a consequence of these coordinates being independent), and the formula for the Laplace-Beltrami operator as the sum of second derivatives along an orthonormal set of geodesics.

Note that there is an issue here concerning whether or not the equations of stochastic development and parallel transport will define BM($\mathbb{M}$) for all time. It is possible for BM($\mathbb{M}$) to explode in finite (random) time, though simple arguments using comparison results described below show that this cannot happen unless the manifold $\mathbb{M}$ has rapidly growing unbounded negative curvatures. If such an explosion has zero probability of occurring then we say that $\mathbb{M}$ is *stochastically complete.* The corresponding issue for geodesics is of course exactly the question of whether the manifold is *complete* (for emphasis we would say, *geodesically complete*). A criterion for stochastic completeness using this approach is given in [56], which also shows the same criterion excludes *implosion*, which is to say the possibility that the Brownian motion X can be defined over the time interval $(0, T)$ for some positive T but cannot be extended back to time zero because $\lim_{t\downarrow 0} X_t$ diverges to infinity on the manifold. (Implosion is equivalent to the Feller property [95].)

In the case of Brownian motion we can obtain sharper results on explosion and implosion using Dirichlet forms rather than stochastic calculus; however the stochastic calculus approach comes into its own when dealing with more general semimartingales such as the Γ-martingales discussed below.

5. *Γ-martingale.* If the stochastic anti-development is a local martingale $Z = M$ then we *define* the resulting semimartingale to be a *Γ-martingale*. (Note that Brownian motion is the special case when M is a Euclidean Brownian motion.) In that case the geometric Itô formula takes the form

$$\mathrm{d}_I f(X) \;=\; \langle \operatorname{grad} f(X), \Xi\rangle \,\mathrm{d}_I M + \frac{1}{2} \operatorname{trace} \operatorname{Hess} f(X)(\Xi, \Xi)\, \mathrm{d}\,[\, M, M\,]\,. \tag{2.10}$$

The "Γ" in Γ-martingale refers to the traditional notation (using Christoffel symbols) for a connection. Γ-martingales actually depend only on the choice of connection.

Eq. (2.10) now corresponds to the Doob-Meyer decomposition, since

$$\langle \operatorname{grad} f(X), \Xi\rangle \,\mathrm{d}_I M$$

is the stochastic differential of a local martingale. Therefore, in the Γ-martingale case, if (v_i) is an orthonormal basis for the reference Euclidean space $\mathbb{R}^m$ and if $M = \sum_i v_i M^i$ then

$$\texttt{Drift}\ \mathrm{d}_I f(X) \;=\; \frac{1}{2} \sum_i \sum_j \operatorname{Hess} f(X)(\Xi v_i, \Xi v_j)\, \mathrm{d}\left[\, M^i, M^j \,\right] \tag{2.11}$$

and as noted above in the discussion on semimartingales we can diagonalize here. Indeed, define the *intrinsic time* τ of the Γ-martingale X in terms of its stochastic anti-development by

$$\mathrm{d}\tau \;=\; \sum_i \mathrm{d}\left[\, M^i, M^i \,\right] \tag{2.12}$$

Then we can choose the orthonormal basis (v_i) adaptively so as to diagonalize (a predictable version of) the density

$$A^{ij} \quad = \quad \frac{\mathrm{d}\left[\, M^i, M^j \,\right]}{\mathrm{d}\tau} \tag{2.13}$$

for $\mathrm{d}\tau$-almost all time. (This uses the fact that one can define a diagonalizing basis for a symmetric matrix which depends measurably on the matrix.)

Recall that a function $f : \mathbb{M} \to \mathbb{R}$ is said to be *convex* if it is convex when composed with geodesics. From the above diagonalization, and from the fact that $\operatorname{Hess} f(\xi v, \xi v) = \mathrm{d}^2 f(\gamma(t))/\,\mathrm{d}t^2$ is nonnegative for smooth convex functions, we deduce that if f is a smooth convex function and X is a Γ-martingale then $\texttt{Drift}\ \mathrm{d}_I f(X) \geq 0$, so a convex function of a Γ-martingale is a submartingale. Note that often a manifold supports *no* nonconstant convex functions, for example if it is compact: however the above relationship can be localized for convex functions defined only within given regions of a manifold, so long as the Γ-martingale is stopped when first it exits the region of definition.

In fact this relationship with convexity characterizes Γ-martingales [26], and extends to non-smooth convex functions [46].

Once again there are issues of explosion (which is to say, *martingale-completeness*) and implosion, for which results can be found in [43, 95].

The interplay between Γ-martingales and convexity is in fact at the heart of this essay, since it combines with the remark (established by considering the Euler-Lagrange equations of extremality for the energy form (1.2)) that smooth harmonic maps $F : \mathbb{M} \to \mathbb{N}$ are characterized by the property (again interpreted locally!) that $\phi(F)$ is subharmonic whenever the smooth real-valued function ϕ defined on $\mathbb{N}$ is convex. As a consequence we can verify an observation due to Bismut in about 1981, which is fundamental to the relationship between harmonic maps and probability:

Theorem 2.3 (Γ-martingales and harmonic maps): *The smooth map*

$$F : \mathbb{M} \to \mathbb{N}$$

is a harmonic map if and only if

$$F(BM(\mathbb{M})) \ \textit{is a } \Gamma\textit{-martingale}\,.$$

Note that the natural question, what kind of map carries Brownian motion over into a time-changed *Brownian motion* (not just a Γ-martingale) has also been answered: the map must be a *harmonic morphism* (pull back harmonic functions to harmonic functions). This is equivalent to the requirement that it be a harmonic map which is also horizontally semi-conformal [10], a rather rigid requirement except in dimension 2, where it reduces to the holomorphic or anti-holomorphic property. (But note that nevertheless there has been some interesting work on probabilistic approaches to such maps: see [35, 36].)

We shall have occasion later to use a natural generalization to the case of non-smooth harmonic maps:

Definition 2.4 (Finely harmonic maps): *Suppose a map* $F : \mathbb{M} \to \mathbb{N}$ *is such that* $F(BM(\mathbb{M}))$ *is a* Γ*-martingale for quasi-every starting point of the Brownian motion* $BM(\mathbb{M})$*. Then we say* F *is a* finely harmonic map*.*

As we shall see, there is a beautiful three-way interaction between harmonic maps, convexity, and Γ-martingales which will form the theme of this essay. Aided by the notion of stochastic development, this will allow us to "flatten out" the nonlinear harmonic map theory, and to tie it closely to the linear potential theory of Brownian motion.

6. *Γ-martingales of bounded quasi-conformality and dilatation.* We continue discussion of the Γ-martingale X and its stochastic anti-development M. The matrix $A = \mathrm{d}\,[\,M, M\,]\,/\,\mathrm{d}\tau$ has eigenvalues

$$0 \leq \lambda_m \leq \lambda_{m-1} \leq \cdots \leq \lambda_2 \leq \lambda_1$$

and the extent to which these differ measures the extent to which X fails to be a (possibly time-changed) Brownian motion. Two summary measures turn out to be of importance to us:

Definition 2.5 (Quasi-conformal Γ-martingales): *We say* X *is* K^2-quasi-conformal *if*

$$\lambda_1 \quad \leq \quad K^2 \lambda_m \,.$$

Definition 2.6 (Bounded-dilatation Γ-martingales): *We say* X *is of* K^2-bounded dilatation *if*

$$\lambda_1 \quad \leq \quad K^2\lambda_2 \,.$$

Note that a harmonic map F of quasi-conformality or bounded dilatation carries these properties over to the associated Γ-martingale $F(\mathrm{BM}(\mathbb{M}))$.

A Γ-martingale of bounded dilatation has "randomness of at least two substantial degrees of freedom" when viewed in its intrinsic time-scale $\mathrm{d}\tau$, while a quasi-conformal Γ-martingale has "substantial randomness in every direction"when viewed in its intrinsic time-scale. (Reference to this intrinsic time-scale is important: working in the original time-scale $\mathrm{d}t$ we may see either kind of Γ-martingale coming to a complete stop.)

It is intuitively clear, and we shall soon substantiate this, that a Γ-martingale of bounded quasi-conformality is very similar to a Brownian motion. However care must be exercised: there *are* differences. For example, if $K^2 > m - 1$ then a K^2-quasi-conformal Γ-martingale can hit prescribed points at positive time, whereas (if $m > 2$) Brownian motion has probability zero of so doing. (This example is due to Krylov [100, §1.3].)

7. *Other possibilities.* Clearly the above possibilities do not exhaust the inventory of the semimartingale zoo. For example, if the manifold $\mathbb{M}$ in question carries a complex structure then we can define *holomorphic* and *anti-holomorphic martingales*: semimartingales on $\mathbb{M}$ which are mapped into time-changed complex Brownian motion by holomorphic maps from $\mathbb{M}$ into $\mathbb{C}$. In fact holomorphic martingales are less closely tied to connections than the inhabitants of the semimartingale zoo discussed previously; it is more natural to define them in a sheaf-theoretic way (holomorphic and anti-holomorphic martingales are precisely those semimartingales which are carried into submartingales by the sheaf of plurisubharmonic functions). There has been a considerable amount of work on the resulting relationship between holomorphic martingales and several-complex-variable theory [50, 69, 125], also [60, 2nd edition]; however this lies largely outside the scope of our essay (though we shall note some important points of overlap).

The issue of sheaf-theoretic definitions propagates back to the consideration of Γ-martingales: as we noted above, Darling has characterized Γ-martingales as those semimartingales which are carried into submartingales by the sheaf of convex functions. This thought can be further developed by asking the question, suppose we reverse the stochastic development idea and ask what connections are compatible with a specific class of semimartingales forming Γ-martingales? Emery [44] has addressed this question: two connections have the same Γ-martingales if they give the same connection when symmetrized. It follows from this that much of the theory of Γ-martingales can be very naturally developed using convexity of functions as a primitive notion, as indeed is done in [44]: however the stochastic development approach comes into its own when one wishes to study and to exploit relationships between different Γ-martingales defined on the same probability space and manifold, as we will do below.

2.1.3 Intrinsic geometry

As noted for example in [60], a smooth elliptic semimartingale diffusion X leads to a natural Riemannian geometry on its state space. This is easily seen from the stochastic development construction: starting with an arbitrarily chosen connection we can develop X to deliver a semimartingale $\tilde{Z}$ on the reference space $\mathbb{R}^m$ and the corresponding stochastic parallel transport $\tilde{\Xi}$. This leads (as in our discussion of Γ-martingales) to a random symmetric matrix process (defined as a matrix-valued Radon-Nikodym derivative)

$$\tilde{A} \quad = \quad \frac{\mathrm{d}\left[\tilde{Z},\tilde{Z}\right]}{\text{trace }\mathrm{d}\left[\tilde{Z},\tilde{Z}\right]}$$

and by the Markov property of the diffusion and its regularity we can realize $\tilde{A}$ in terms of a smooth quadratic form Q on the tangent bundle $T(\mathbb{M})$ *via*

$$\tilde{A}(v,v) \quad = \quad \langle\tilde{\Xi}v, Q(X)\tilde{\Xi}v\rangle\,.$$

But now this quadratic form Q yields a new Riemannian metric on $\mathbb{M}$ which is intrinsic to the diffusion. Moreover if we rework the stochastic development procedure using the new intrinsic Riemannian metric and its associated Levi-Civita connection then the resulting stochastic anti-development Z will be Euclidean Brownian motion B plus a superimposed drift $U\,\mathrm{d}t$:

$$\mathrm{d}_I Z \quad = \quad \mathrm{d}_I B + U\,\mathrm{d}t$$

and this drift can be related (again using the Markov property of the diffusion) to an *intrinsic drift vectorfield* χ on $\mathbb{M}$ by

$$\chi(X) \quad = \quad \Xi U$$

where Ξ is the parallel transport of X with respect to the new metric.

Thus each semimartingale which is a smooth elliptic diffusion can be characterized in distribution by an intrinsic Riemannian metric and an intrinsic drift vectorfield.

Note that a much deeper generalization of this phenomenon (intrinsic geometry of diffusion) can be obtained for processes defined by rather general Dirichlet forms: see Sturm's contribution to this volume.

2.1.4 Diffusions and statistical shape

As mentioned in the introduction, a rather non-trivial example of the intrinsic geometry of diffusions arises when one considers the *statistical shape* of a configuration of m points in d-dimensional Euclidean space, where the points are moving with independent Euclidean Brownian motions. The shape here is the equivalence class of a configuration under changes of scale, orientation, and location; and it forms a Riemannian manifold away from a singularity set. The shape diffuses on the resulting *shape-manifold* as a semimartingale. In case $m = 3$, $d = 2$ the semimartingale can be viewed as a random time-change of $\mathrm{BM}(S^2)$, Brownian motion on the sphere [72]; more generally the semimartingale is *not* a time-change of Brownian motion [83, 86].

An alternative shape geometry, at least for three points in the plane, arises if we require the three points to move not independently but as if they are being transformed according to Brownian motion on the special linear group. This is more appropriate in biological applications of the statistics of shape. The resulting geometry is that of the hyperbolic plane of curvature -2 [97]: this will be of no surprise to people who know symmetric space theory! As previously noted, most of these shape theory results were obtained using a mixture of stochastic differential geometry and computer algebra, more specifically symbolic Itô calculus.

2.1.5 Geodesic variational formulae

The formula (2.11) for the drift of a function $f(X)$ of a Γ-martingale X makes it plain that control of $f(X)$ depends on control of the second derivative of f along geodesics, since

$$\operatorname{Hess} f(X)(\Xi v, \Xi v) \quad = \quad L_v L_v f \circ \pi(\Xi) \quad = \quad \left[\frac{\mathrm{d}^2}{\mathrm{d}u^2} f(\gamma(u))\right]_{u=0}$$

for a unit-speed geodesic γ started at X with initial vector Ξv (note that $|v| = 1$ as the geodesic γ is unit-speed). In stochastic differential geometry we typically require to control such second derivatives when f is the distance from something: a point, a geodesic, or another Γ-martingale. We do this using a number of important and classical curvature-related bounds which we summarize for future use below.

Before doing this we note a technical difficulty. Distance functions on manifolds are not everywhere smooth! For example, the distance from a fixed point $\mathbf{o}$ is certainly not smooth at $\mathbf{o}$ itself, though this usually does not matter (either the Γ-martingale is Brownian motion, which will never hit $\mathbf{o}$, or in any case the required control need be exercised only away from $\mathbf{o}$). But also the distance function is not smooth at the *cut-locus* $C(\mathbf{o})$, the locus where geodesics emanating from $\mathbf{o}$ fail to be length-minimizing. If the underlying manifold $\mathbb{M}$ is simply-connected and of non-positive curvature (a *Cartan-Hadamard manifold*) then the cut-locus is empty, but in general it cannot be ignored.

2.1.6 Distance from a fixed point

Suppose then that we are considering $\ell(u) = \operatorname{dist}(\gamma(u), \mathbf{o})$, the distance from $\mathbf{o}$ of a unit-speed geodesic $\gamma(u)$. Suppose further that $\gamma(u)$ does not meet $C(\mathbf{o})$, so the technical difficulty mentioned above does not arise. Set $\ell(0) = \ell$. If the sectional curvatures of the underlying manifold $\mathbb{M}$ are pinched between k_- and k_+,

$$k_- \quad \leq \quad \text{sectional curvatures} \quad \leq \quad k_+ \,,$$

and if $\ell\sqrt{k_+} < \pi$ in the case when $k_+ > 0$, then we have the following bounds on the second derivative of ℓ:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}u^2}\ell(u)\right]_{u=0} \quad \begin{matrix}\geq \\ \leq\end{matrix} \quad |(\gamma')^\perp|^2 \, \frac{C_{k_\pm}(\ell)}{S_{k_\pm}(\ell)} \tag{2.14}$$

where we take $\geq$ for k_+, $\leq$ for k_-, set $(\gamma')^\perp$ to be the component of γ' normal to the radial geodesic from $\mathbf{o}$, and set

$$\begin{aligned}
&\left.\begin{aligned} S_k(\ell) &= \tfrac{\sin(\sqrt{k}\ell)}{\sqrt{k}} \\ C_k(\ell) &= \cos(\sqrt{k}\ell) \end{aligned}\right\} && \text{if } k > 0\,; \\
&\left.\begin{aligned} S_k(\ell) &= \ell \\ C_k(\ell) &= 1 \end{aligned}\right\} && \text{if } k = 0\,; \\
&\left.\begin{aligned} S_k(\ell) &= \tfrac{\sinh(\sqrt{-k}\ell)}{\sqrt{-k}} \\ C_k(\ell) &= \cosh(\sqrt{-k}\ell) \end{aligned}\right\} && \text{if } k < 0\,.
\end{aligned} \tag{2.15}$$

The proof is basically a variational argument using the Index lemma for Jacobi vectorfields, combined with a simple comparison argument for a second-order differential equation. For a modern treatment of these ideas in textbook form see [66]. Also note that the comparison result given above can be obtained as a simple consequence of curvature bounded by k in the sense of Alexandrov, as described in Jost's contribution to this volume.

There is a variation on the above result which will be important: if in addition $k_+ > 0$ then

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}u^2}\frac{1}{k_+}\left(1 - C_{k_+}(\ell(u))\right)\right]_{u=0} \quad \geq \quad C_{k_+}(\ell)\,. \tag{2.16}$$

(In case $k_+ = 0$ then a valid variant replaces $\frac{1}{k_+}\left(1 - C_{k_+}(\ell(u))\right)$ by $\ell(u)^2/2$.)

Remark 5 *The inequality fails for $k_+ < 0$ simply because then we have $|C_{k_+}| > 1$.*

2.1.7 Distance between two moving points (I)

With curvature bounds as above, and in particular supposing that if $k_+ > 0$ then $\ell\sqrt{k_+} < \pi$, let $\ell(u)$ now be the distance between two different geodesics $\gamma_0(u)$ and $\gamma_1(u)$. Suppose that $\gamma_0(u)$ avoids the cut-locus of $\gamma_1(u)$ (this implies that $\gamma_1(u)$ avoids the cut-locus of $\gamma_0(u)$ as well). We find

$$\begin{aligned}\left[\frac{\mathrm{d}^2}{\mathrm{d}u^2}\ell(u)\right]_{u=0} \quad & \begin{matrix}\geq \\ \leq\end{matrix} \\ & \left[|(\gamma_0')^\perp|^2 - 2\frac{\langle\xi_0^{-1}(\gamma_0')^\perp, \xi_1^{-1}(\gamma_1')^\perp\rangle}{C_{k_\pm}(\ell)} + |(\gamma_1')^\perp|^2\right]\frac{C_{k_\pm}(\ell)}{S_{k_\pm}(\ell)}\,.\end{aligned} \tag{2.17}$$

Here ξ is the parallel transport for the geodesic γ, so that the inner product $\langle(\xi_0^{-1}\gamma_0')^\perp, \xi_1^{-1}(\gamma_1')^\perp\rangle$ makes sense.

Once again, we take $\geq$ for k_+, $\leq$ for k_-.

Convenient bounds for applications follow by completing the square ....

The proof is essentially as for the previous bound.

2.1.8 Distance between two moving points (II)

Suppose now we are interesting in bounding the distance between two *Brownian motions.* It turns out that we need to bound a quantity similar to that given in the previous section, but *averaged* over two sets of orthonormal geodesics, $\gamma_0^{(i)} : i = 1, \ldots, m-1$ intersecting at a shared point $b = \gamma_0^{(i)}(0)$, and $\gamma_1^{(i)} : i = 1, \ldots, m-1$ intersecting at a shared point $a = \gamma_1^{(i)}(0)$, such that all geodesics are perpendicular to the distance minimizing geodesic between a and b and such that parallel transport along the intervening geodesic carries $\gamma_0^{(i)}$ to $\gamma_1^{(i)}$. Suppose the curvature condition is replaced by

$$(m-1)k_- \quad \leq \quad \text{Ricci curvatures of } \mathbb{M}$$

(recall the Ricci curvature is the sum of curvatures for $m-1$ planes sharing a common vector and otherwise defined by a perpendicular set of $m-1$ vectors). For $k_- > 0$, suppose additionally that $\ell\sqrt{k_-} < \pi$. Let $\ell^{(i)}(u)$ be the distance between $\gamma_0^{(i)}(u)$ and $\gamma_1^{(i)}(u)$. Then

$$\sum_{i=1}^{m-1} \left[\frac{\mathrm{d}^2}{\mathrm{d}u^2} \ell^{(i)}(u) \right]_{u=0} \quad \leq \quad -2(m-1)k_- \frac{S_{k_-}(\ell/2)}{C_{k_-}(\ell/2)} \,. \tag{2.18}$$

The proof follows the same lines as before, except that a further argument is required exploiting the fact that in the case of constant curvature Jacobi vectorfields which are perpendicular at both end-points of a geodesic are perpendicular across the whole intervening range.

A similar result holds for the sum of second geodesic derivatives of radial distance from a single point, and results in upper bounds on $\operatorname{dist}(\mathrm{BM}, \mathbf{o})$ using lower Ricci curvature bounds.

There are further variations on these comparison results, for example distance from a geodesic [75, 57] or more generally a totally geodesic submanifold; also distance in a manifold with varying radial curvature bounds depending on distance from a fixed reference point $\mathbf{o}$ [54]. However the above bounds suffice for most purposes of stochastic differential geometry (and indeed arguments both for the more general bounds and for their applications are largely technical variations on what is given here.)

Chapter 3

Distance and semimartingales

Distance is a primitive notion in Riemannian geometry. Therefore a natural starting point for stochastic differential geometry is to conduct an Itô analysis of the distance process $\operatorname{dist}(X, \mathbf{o})$, for X a semimartingale (typically a Γ-martingale or Brownian motion), and $\mathbf{o}$ a fixed reference point. Variations include: distance from a totally geodesic hyperplane, distance from another (independent or coupled) Brownian motion or Γ-martingale.

3.1 Distance of a Γ-martingale from a point

The simplest case, which also carries all the main ideas, is that of the Itô analysis of the distance of a Γ-martingale from a reference point $\mathbf{o}$. Consider a Γ-martingale X on a manifold $\mathbb{M}$ all of whose sectional curvatures are bounded above by $-H^2 < 0$ (but which is not necessarily simply connected). Let M be the stochastic anti-development, and let Ξ be the parallel transport, for X. Set

$$\begin{aligned} r(x) &= \operatorname{dist}(x, \mathbf{o}) \\ R &= r(X) \end{aligned} \tag{3.1}$$

and use the equations of stochastic development and the bound Eq. (2.14) on the second derivative of geodesic arc-length to compute the following comparison, valid up to the first time that X hits the cut-locus of the reference point $\mathbf{o}$:

$$\begin{aligned}
\mathrm{d}_I R &= \langle \operatorname{grad} r(X), \Xi\rangle \, \mathrm{d}_I M + \frac{1}{2}\sum_i \operatorname{Hess} r(X)(\Xi v_i, \Xi v_i)\lambda_i \, \mathrm{d}\tau \\
&\geq \langle \operatorname{grad} r(X), \Xi\rangle \, \mathrm{d}_I M + \frac{H}{2}\coth(HR)\left(\sum_i |v_i^\perp|^2 \lambda_i\right) \mathrm{d}\tau
\end{aligned} \tag{3.2}$$

where (v_i) is an adapted basis which diagonalizes the Hessian quadratic form $\operatorname{Hess} r(X)(\Xi, \Xi)$, the (λ_i) are the associated eigenvalues, and $\Xi v^\perp$ is the component of Ξv which is normal to $\operatorname{grad} r(X)$, and the inequality on stochastic differentials is interpreted as meaning that the difference of left-hand side minus right-hand side is a *classical* differential of a non-negative random measure. The above stochastic differential inequality is valid for general semimartingales but is not then so useful, as in the general case the Itô differential term will contribute a further drift term.

In the Γ-martingale case we now change time-scale according to the quadratic variation of R:

$$\mathrm{d}\sigma \quad = \quad \sum_i |\langle \operatorname{grad} r(X), \Xi v_i\rangle|^2 \lambda_i \, \mathrm{d}\tau \quad = \quad \sum_i |v_i^\parallel|^2 \lambda_i \, \mathrm{d}\tau$$

where $v^\parallel$ is defined by requiring $\Xi v^\parallel$ to be the component of Ξv which is *parallel* to $\operatorname{grad} r$. Now a continuous martingale in the time-scale of its quadratic variation differential $\mathrm{d}\sigma$ is a standard Brownian motion (stopped at a random time if $\mathrm{d}\sigma$ integrates up to a finite limit). Therefore the stochastic differential inequality can be reformulated in the $\mathrm{d}\sigma$ time-scale as

$$\mathrm{d}_I R \quad \geq \quad \mathrm{d}_I W + \frac{H}{2}\coth(HR)\left(\frac{\sum_i |v_i^\perp|^2\lambda_i}{\sum_i |v_i^\parallel|^2\lambda_i}\right) \mathrm{d}\sigma \tag{3.3}$$

where W is a standard Brownian motion in the $\mathrm{d}\sigma$ time-scale.

Thus a negative curvature upper bound has a tendency, bounded below in drift by a term with coefficient $\frac{H}{2}\coth(HR)$, to push the Γ-martingale away from the reference point. If in addition we can bound from below the expression

$$\left(\frac{\sum_i |v_i^\perp|^2\lambda_i}{\sum_i |v_i^\parallel|^2\lambda_i}\right) \tag{3.4}$$

(note that this depends implicitly on $\operatorname{grad} r$ and X through the adapted basis (v_i)) then we can use the theory of standard Brownian motion with constant drift to deliver lower bounds on the behaviour of $R = r(X)$. Such bounds on the expression (3.4) are automatic in the case of Brownian motion (for then any basis (v_i) will serve and the expression equals $m-1$): for general Γ-martingales we need a condition such as quasiconformality or bounded dilatation of the Γ-martingale.

It is straightforward to modify this for cases when the upper curvature bound is non-negative, or when there is a lower curvature bound (giving rise to an upper bound on the radial process). We leave such considerations as an exercise for the reader, though we shall shortly have occasion to use the instance of a lower curvature bound.

Notice that the above analysis holds only till the first time that X hits the cut-locus of $\mathbf{o}$. In the case of a Cartan-Hadamard manifold (simply-connected, non-positive curvature) there is no cut-locus and so this is not an issue. In other cases it presents a technical obstacle for the application of stochastic calculus to differential geometry: we shall discuss its solution later.

3.2 Probabilistic approaches to Picard theorems

The main question to be faced now is, how can we use such bounds as are given above to establish links between Brownian motion and harmonic maps? To make a start on this we review Burgess Davies' delightful Brownian motion proof of the Picard little theorem from complex analysis [32].

Suppose $f : \mathbb{C} \to \mathbb{C} \setminus \{\pm 1\}$ is a holomorphic map. The Picard little theorem tells us f must be constant. On the other hand Lévy's celebrated theorem tells us that $f(\mathrm{BM}(\mathbb{C}))$ is a time-changed complex Brownian motion $\mathrm{BM}(\mathbb{C})$. Burgess Davies showed that if we consider $\mathrm{BM}(\mathbb{C} \setminus \{\pm 1\})$, and connect end-point to start-point with a straight-line segment, then with probability one for all sufficiently large times the resulting loop will be topologically entangled in $\mathbb{C} \setminus \{\pm 1\}$ (of course one has to fix-up the construction of the loop when the segment crosses one of the excluded points $\{\pm 1\}$).

Now suppose that f is not constant. Burgess Davies points out that if we consider $f(Z)$ for Z a $\mathrm{BM}(\mathbb{C})$ then by Lévy's theorem we know that $f(Z)$ must be a $\mathrm{BM}(\mathbb{C})$ up to a random time-change. Omitting some technical details, this means we can apply the entangling result and the recurrence of the original $\mathrm{BM}(\mathbb{C})$ to create a null-homotopic loop in $\mathbb{C}$ which is mapped by f into a homotopically non-trivial loop. This is a contradiction, and we are forced to conclude that f is constant.

How can this be generalized to the case of harmonic maps? We should not expect to exploit Brownian entanglement results, as we will want to work in higher dimensional situations. This generalization has already been addressed by differential geometers working in harmonic map theory: for example [52] observes that by lifting to universal covers we can resolve the Picard little theorem into a Liouville theorem for holomorphic maps into the complex unit disk and thus one can formulate a generalization. Since the complex unit disk carries a natural metric of constant negative curvature, the generalization of Picard theorem for harmonic maps should run like this: maps F in the following are always constant!

F	:	$\mathbb{M}$	$\rightarrow$	$\mathbb{N}$
harmonic map and further conditions		*suitable conditions (almost Euclidean)*		*simply connected, enough negative curvature*

In [52] and elsewhere geometric methods (volume and distance comparison) have been used to supply Picard little theorems (for example, with a bounded dilatation condition of F, nonnegative Ricci curvature for $\mathbb{M}$, and curvature bounded above away from zero for $\mathbb{N}$). The question of how a probabilistic approach might be made now becomes clearer: one should try to prove results for suitable Γ-martingales on $\mathbb{N}$ which are incompatible with properties valid for Brownian motion on $\mathbb{M}$ unless there are no non-constant F of the required form.

3.3 Γ-martingale convergence theorem

The first step on this road (though not the first step chronologically!) is to try to prove a Liouville theorem for harmonic maps using Brownian motion. But to do this we need to know more about Γ-martingales. Now the fundamental result for ordinary (that is, real-valued!) martingales is the martingale convergence theorem: in a simple form, bounded martingales always converge at time infinity. It was natural to attempt to generalize this result to Γ-martingales and this was done at the very earliest stages of the theory: see Darling [27, 28] and Zheng [138]. They established the following result:

Theorem 3.1 (Darling-Zheng Γ-martingale convergence theorem): *If X is a Γ-martingale on a manifold $\mathbb{M}$ with intrinsic time τ then*

1. *(Darling): if $\tau(\infty) < \infty$ then X converges in $\mathbb{M} \cup \{\infty\}$, the one-point compactification of $\mathbb{M}$;*

2. *(Zheng): if X converges in $\mathbb{M}$ then $\tau(\infty) < \infty$.*

We shall prove only the first part, which is all we will need below.

Proof: (part (i) only.) A *localization* argument (stopping X at the exit times from progressively larger geodesic balls) shows we may suppose $\tau(\infty) \leq T < \infty$ for some fixed T.

We can cover $\mathbb{M}$ by geodesic balls in a locally finite cover, such that a ball centre x radius ϵ_x supports a uniformly strictly convex function

$$\phi_x(y) \quad = \quad \frac{1}{k_x}\left(1 - \cos(\sqrt{k_x}\,\mathrm{dist}(x,y))\right)$$

We choose k_x to be a nonnegative upper bound on curvature in the ball: this together with Eq. (2.16) means that ϕ_x is strictly convex as long as the ball is sufficiently small.

The strict convexity of ϕ_x in the corresponding ball means that $\phi_x(X)$ behaves as a submartingale in the ball. Together with the finiteness of the intrinsic time $\tau(\infty)$, this allows us to show there can only be a finite number of downcrossings of X from $\phi_x^{-1}(\epsilon_x)$ to $\phi_x^{-1}(\epsilon_x/2)$, using analysis of $\mathrm{d}_I\phi_x(X)$ (the drift being controlled by the quantification of the strict convexity of ϕ_x given by Eq. (2.16)) and the finiteness of T.

Now a *topological* downcrossings argument completes the proof: if X fails to converge in $\mathbb{M}$ then it must eventually leave any compact subset (since it cannot oscillate between ϕ_x levels too much by the above), and this exactly says it must then converge to ∞ in the one-point compactification. $\square$

3.4 An "Ultimate" Liouville Theorem

We get an immediate pay-off from the Γ-martingale convergence theorem: a probabilistically proved Liouville theorem which, when it was proved, superseded most of the then-known Liouville theorems, and which has a very pleasant "categorical" appearance. To state this we need the notion of a geodesic ball which is not too big:

Definition 3.2 (Regular geodesic ball): *A* regular geodesic ball $\mathcal{B}$ *is a geodesic ball (centre* $\mathbf{o}$*, radius* r*) such that*

(1) $\mathcal{B}$ *does not intersect the cut-locus of the centre* $\mathbf{o}$*;*

(2) the sectional curvatures in $\mathcal{B}$ *are all smaller than* $\left(\frac{\pi}{2r}\right)^2$*.*

A good way to visualize this is to think of regular geodesic balls as "small hemispheres" (compact subsets of open hemispheres). Essentially all the geometric difficulties are carried by this specific example.

Theorem 3.3 ("Ultimate" Liouville Theorem *[59, 81], see also [85]***):** *Consider the functions f and F in the following diagram:*

$$\begin{array}{cccc} & & F & \\ & \mathbb{M} & \longrightarrow & \mathcal{B} \\ \textit{(harmonic) } f & \downarrow & \textit{harmonic} & \textit{regular} \\ & [0,1] & \textit{map} & \textit{geodesic ball} \end{array} \tag{3.5}$$

If all such harmonic functions f are constant, then all such harmonic maps F are also constant.

Proof: Consider X, a Brownian motion BM($\mathbb{M}$) started at $x \in \mathbb{M}$ say, and let $Y = F(X)$. If k_+ is the non-negative strict upper bound for the curvatures of the regular geodesic ball $\mathcal{B}$ (as given in Definition 3.2 (2)), then we can use Eq. (2.16) to show that

$$\phi(y) \quad = \quad \frac{1}{k_+}\left(1 - \cos(\sqrt{k_+}\,\mathrm{dist}(\mathbf{o}, y))\right)$$

(for $y \in \mathcal{B}$, and taking $\mathbf{o}$ to be the centre of the regular geodesic ball $\mathcal{B}$) supplies a strictly convex function which delivers a finite upper bound on the expected total intrinsic time $\mathbb{E}\left[\,\tau(\infty)\,\right]$. This allows us to apply the first part of the Darling-Zheng Γ-martingale convergence theorem, since now we know that $\tau(\infty) < \infty$ with probability one. So the limit $Y(\infty)$ exists in $\mathcal{B}$ (no need to add the compactification $\{\infty\}$ since $\mathcal{B}$ is itself compact).

The hypothesis on $\mathbb{M}$ is that bounded harmonic functions are all constant, and of course this means that $Y(\infty)$ must be non-random: we set

$$y \quad = \quad Y(\infty) \quad = \quad \lim_{t\to\infty} F(X(t))\,.$$

(Otherwise we could use $\mathbb{P}\left[\,Y(\infty) \in A | X(0) = x\,\right]$ to generate non-constant bounded harmonic functions on $\mathbb{M}$.)

Now we use some non-trivial geometry of regular geodesic balls. Again we can use the geodesic comparison arguments for distance from a specified point as given in the previous section and so show the following function is bounded non-negative convex on $\mathcal{B}$:

$$\psi_y(z) \quad = \quad \frac{1}{k}\,\frac{1 - \cos(\sqrt{k}\,\mathrm{dist}(z, y))}{\cos(\sqrt{k}\,\mathrm{dist}(y, \mathbf{o}))}\,.$$

(Here $\mathbf{o}$ is the centre of the regular geodesic ball $\mathcal{B}$ of interest, and $\psi_y(z)$ is defined for $y, z \in \mathcal{B}$.)

It follows that $\psi_y(Y)$ is a bounded negative submartingale (convexity properties of ψ_y) which converges to $\psi_y(Y(\infty)) = \psi_y(y) = 0$. It then follows from the property defining a submartingale that $\psi_y(Y) = 0$ always, and so $Y = y$ for all time. But this implies F must be constant, which establishes the result. □

It should be noted that the choice of the form for the function ψ came from the geometry paper [63], which used partial differential equations to establish a special case of the above Liouville theorem. A translation of the above probabilistic proof into non-stochastic (potential-theoretic) terms may be found in [85]. Note that there are Liouville theorems which are *not* special cases of the above result, for example Cheng's Liouville theorem for harmonic maps of linear growth [18] (but Stafford [127] has produced a probabilistic proof of this).

3.5 Picard little theorem for harmonic maps

Goldberg and coworkers have established a Picard little theorem for harmonic maps: any such F in the following has to be constant!

$$\begin{array}{cccc} F & : & \mathbb{M} & \rightarrow & \mathbb{N} \\ \textit{harmonic map of bounded dilatation} & & \textit{nonnegative Ricci} & & \textit{Sectional curvatures bounded above and away from 0} \end{array} \tag{3.6}$$

A standard covering argument shows that we may suppose both $\mathbb{M}$ and $\mathbb{N}$ to be simply-connected.

Bearing in mind Burgess Davies' proof for the holomorphic Picard little theorem, and the Ultimate Liouville Theorem, one might try for a probabilistic proof along the following lines: show that the Γ-martingale of K^2-bounded dilatation given by $F(\mathrm{BM}(\mathbb{M}))$ has a random *limiting direction* if F is nonconstant, and so derive a contradiction in the manner of the proof of the Ultimate Liouville theorem, using the fact that manifolds with nonnegative Ricci curvatures have no non-constant harmonic functions [137]. This program has been carried out with a large degree of success in [73, 74] and generalized in [24, 53]: [4] provides an alternative stochastic approach using gradient estimates. In fact most of the angle-based work *preceded* the conceptually simpler Ultimate Liouville theorem: it was informed and motivated by work by Dynkin and others [39, 98, 120] concerning limiting directions of Brownian motion on Cartan-Hadamard manifolds.

3.5.1 Limiting directions for Brownian motion

A Cartan-Hadamard manifold $\mathbb{N}$ has no cut-locus and so we may extend geodesic polar coordinates (r, θ) over all $\mathbb{N}$. Hence it makes sense to discuss whether there might exist a limit at time infinity for the angle process $\Theta = \theta(\mathrm{BM}(\mathbb{N}))$ of Brownian motion.

In fact it is easy to show (by skew-product representation arguments) that such a limit cannot exist for the angle process for the simplest Cartan-Hadamard manifold, namely Euclidean space of dimension 2 or more. This argument has been extended to cover manifolds of rotational symmetry [110]: skew-product arguments use the rotational symmetry to show that the angle process has a genuinely random limit if the curvature is less than a negative constant (and indeed even if the curvature is negative and does not decay too fast).

If there is no rotational symmetry then the argument becomes much harder. In order to control the angle of Brownian motion one needs an extra condition: either that the manifold is two-dimensional [57, 75], or a non-radial gradient condition on the curvature [98], or a lower bound on the curvature [103, 120, 130]. The lower curvature bound is the most attractive condition: we state it as a theorem and indicate its proof.

Theorem 3.4 (Limiting angle in pinched Cartan-Hadamard manifolds): *Suppose $\mathbb{N}$ is a Cartan-Hadamard manifold with negative curvature pinched between two negative constants:*

$$-L^2 \quad \leq \quad \textit{Sectional curvatures} \quad \leq \quad -H^2$$

Then BM($\mathbb{N}$) has a random limiting direction.

Proof: (Sketch.) We sample the BM($\mathbb{N}$) process X at the times when it moves unit distance from the most recent sampling point:

$$T_0 = 0, \quad T_{k+1} \quad = \quad \inf\{t > T_k : \operatorname{dist}(X_t, X_{T_k}) = 1\}$$

Conditioning on $\mathfrak{F}_{T_k}$, the past at time T_k, an Itô analysis modifying the analysis of Eq. (3.3) to apply to *lower curvature bounds* shows that $u \mapsto \operatorname{dist}(X_{T_k+u}, X_{T_k})$ can be bounded *above* by Brownian motion plus constant positive drift.

Hence the strong law of large numbers can be applied to show that $T_k \to \infty$ and moreover there is $\beta > 0$ such that as $K \to \infty$ so

$$\mathbb{P}\left[\, T_{K+k} \geq \beta k \text{ for all } k > 0 \mid X_0 \,\right] \qquad \to \quad 1$$

uniformly in the starting point X_0.

Now on the other hand the discussion following Eq. (3.3) shows that the *upper curvature bound* implies that R is bounded *below* by Brownian motion plus a constant positive drift, and so elementary Brownian motion theory (related to the strong law of large numbers) shows that for some positive α as $R_0 \to \infty$ so

$$\mathbb{P}\left[R_t \geq \frac{1}{2}R_0 + \alpha t\right] \to 1$$

Together with a comparison of angles with those defined by comparable triangles on the hyperbolic plane,

$$\angle(\Theta(T_{K+k}), \Theta(T_{K+k+1})) \quad \leq \quad c_0 \exp(-c_1 R(T_{K+k}))$$

for c_0, $c_1 > 0$. This allows us to deduce convergence to a limiting angle which is random (since as the starting distance tends to infinity so it is uniformly increasingly probable that the limiting angle is close to the initial angle). □

The obvious question here is, whether one can drop the lower bound on the curvature. This is an open problem! As noted above, the lower bound is not required in the two-dimensional case. Treatments of variable lower curvature bounds are to be found in [58, 103]: it is noteworthy that the best lower curvature bound corresponds approximately to the lower curvature bound required to prevent the possibility of explosion/stochastic incompleteness (Brownian motion leaving all compact sets within a finite random time). Ancona [2] describes an informative simply-connected three-dimensional example using a warped metric of negative curvature bounded above away from zero, in which the limiting angle exists but is constant. It is conceivable that a variation on this example would lead to a case in which the limiting angle does not exist. However the interesting question here attaches to the *randomness* of the limiting angle. Although the limiting angle in Ancona's example is non-random, nevertheless one can identify limiting pseudo-angles which are random, so the invariant σ-algebra of the Brownian motion in question does not satisfy a zero-one law. It is still an open question whether it is possible to construct a Cartan-Hadamard manifold with curvature bounded above away from zero for which the invariant σ-algebra satisfies a zero-one law, equivalently, for which all bounded harmonic functions are constant.

We should also note the use of Gromov hyperbolicity, visibility, and a condition on the Laplacian of the distance function, all to replace the negative upper bound on the curvature of the Cartan-Hadamard manifold: see for example [1, 99].

3.5.2 Probabilistic proofs for Picard little theorems

The way is now clear for a probabilistic proof of a generalized Picard little theorem. We simply generalize the proof of limiting angle from Brownian motion to Γ-martingales. It is easily seen that we need at least a bounded dilatation condition (since for example Brownian geodesics cannot possess a limiting angle!)

Theorem 3.5 (Limiting angle for Γ-martingales): *Let $\mathbb{N}$ be a Cartan-Hadamard manifold with sectional curvatures pinched between two negative constants. If X is a Γ-martingale of bounded dilatation on $\mathbb{N}$ then it possesses a limiting angle at time infinity: moreover this angle is genuinely random.*

Proof: The techniques carry over largely unchanged from the Brownian motion case above. There are two points to watch. First, the intrinsic time τ of X may be finite, so that we cannot exclude the possibility that X converges within $\mathbb{N}$. However this leads immediately to the conclusion that a limiting angle exists: randomness of this angle can then be derived from an Itô analysis of the distance of X from any prescribed geodesic line.

Second, a simple geometric argument is required to verify that the bounded dilatation condition deploys to enable a lower bound of $\text{dist}(X, \mathbf{o})$ (in intrinsic time-scale) by a standard Brownian motion with constant drift. $\square$

A Picard little theorem for harmonic maps is an immediate corollary of the above.

Theorem 3.6 (Picard little theorem for harmonic maps I): *Suppose that $\mathbb{N}$ is a manifold with sectional curvatures pinched between two negative constants, and $\mathbb{M}$ is a manifold whose universal cover supports no non-constant bounded harmonic functions. If the map*

$$F : \mathbb{M} \quad \to \quad \mathbb{N}$$

is harmonic and of bounded dilatation then it must be constant.

Proof: By a covering argument we can suppose that $\mathbb{N}$ is simply connected. But now $F(\text{BM}(\mathbb{M}))$ is a Γ-martingale of bounded dilatation, and therefore has a random limiting angle unless F is constant. Since a random limiting angle would supply non-constant bounded harmonic functions, we conclude that F must be constant. $\square$

By [137] the Ricci-nonnegative condition on $\mathbb{M}$ implies no non-constant harmonic functions on its universal cover, hence the geometric Picard little theorem for harmonic maps, as stated above, is nearly a special case of the probabilistically proved version. However the probability proof requires the lower curvature bound, and that this restriction is non-trivial is a consequence of the open question for Brownian motion angle.

This contrast shows that the probability proof is genuinely *different* from the geometric proof: a price has to be paid in terms of a lower curvature bound, but in return one gains a tight link to (linear) potential theory on the domain manifold. This link is characteristic of probabilistic approaches to harmonic map theory.

Some gain on the lower curvature bound can be achieved by a more careful analysis along the same lines as above but using varying radial curvature bounds [54] in the manner of the approach to the Brownian motion problem described in [58]: see [78] for an indication of the Γ-martingale approach. (It should also be noted that [53] describes how to weaken the *upper* curvature bound: see also [57].) However such an analysis does not add anything substantial to our understanding of the lower curvature bound problem here.

Before leaving the topic of limiting angle, note that in [24] it is shown how to replace the upper curvature bound in the above by a combination of Gromov hyperbolicity, a visibility condition, and requirement that the Laplacian of the distance function is sufficiently negative in sufficiently many balls in $\mathbb{N}$.

3.6 Life on the cut-locus

The previous analysis depends on the Itô analysis, for X a Γ-martingale with stochastic anti-development M and parallel transport Ξ,

$$\mathrm{d}_I r(X) = \langle \operatorname{grad} r(X), \Xi \rangle \, \mathrm{d}_I M + \frac{1}{2} \operatorname{trace} \left(\operatorname{Hess} r(X)(\Xi, \Xi) \, \mathrm{d}\,[\, M, M\,] \right) \quad (3.7)$$

which, as we have noted, holds true only up to the first time that X first hits the cut-locus of the reference point $\mathbf{o}$. In the case of Cartan-Hadamard manifolds this is not a hindrance, since the cut-locus is always empty. However other manifolds typically *do* have cut-locus (for example, the Cheeger-Gromov theorem says that manifolds with non-negative Ricci curvature and no cut-locus must be Euclidean: see [82] for a probabilistic perspective on this result). So a more detailed analysis is required once we move away from Cartan-Hadamard manifolds, as is required even in generalizations of the above (see [5]).

3.6.1 Historical perspective

In fact it is well-known in the analytic and geometric context what has to be done: the Laplacian of the distance function can be extended over the cut-locus so long as it is there interpreted in a distributional sense as a non-positive measure [13, 136, 137]. The question for probabilists is, how to understand this result in terms of the trajectory of Brownian motion.

3.6.2 Cut-locus correction by upper bound

The first point to note is that by one means or another the analytic interpretation of the Laplacian of the distance function translates directly into a simple correction term to be added to Eq. (3.7). In [79] the geometric treatment given in [13] is adapted to the Brownian path in order to show the following.

Theorem 3.7 (The radial part of Brownian motion on a manifold): *Suppose that X is $BM(\mathbb{M})$ with stochastic anti-development B and parallel transport Ξ. If R is its radial part ($R = \operatorname{dist}(X, \mathbf{o})$) then*

$$\mathrm{d}_I r(X) \quad = \quad \langle \mathit{grad}\, r(X), \Xi \rangle \, \mathrm{d}_I B + \frac{1}{2} \Delta\, r(X)\, \mathrm{d}t - \mathrm{d}L \qquad (3.8)$$

where L is an increasing random process which changes only when X visits the cut-locus of the reference point $\mathbf{o}$.

There is nothing surprising in this: L is the potential corresponding to the measure part of $-\Delta\, r$. Informally speaking, this result tells us that the effect of the cut-locus is simply to pull the Brownian motion closer in to the reference point $\mathbf{o}$.

An alternative more analytic derivation of the probabilistic representation is to be found in [56], which applies it to non-explosion and non-implosion (Feller property) results for Brownian motion. The non-implosion results are extended to Γ-martingales (to answer conjectures by Emery) in [95], where the method of proof is also streamlined to link it tightly to convexity and Topogonov's comparison theorem after the analytic approach of Wu [136].

It is reasonable to ask for a more detailed representation of the $\mathrm{d}L$ term, in particular for an explicit relationship in terms of local times on submanifolds and the local geometry of the cut-locus. For Brownian motion this is to be found in [25]: in fact

$$-\,\mathrm{d}L \quad = \quad \frac{1}{2} \left[D_+ r(\nu) + D_- r(-\nu) \right] \, \mathrm{d}L^{C(\mathbf{o})} \qquad (3.9)$$

where

- ν is a measurable unit normal vectorfield to the cut-locus $C(\mathbf{o})$;
- $D_+ r(\pm\nu)$ is the one-sided directional derivative of r in the direction $\pm\nu$;
- $L^{C(\mathbf{o})}$ is "geometric local time on the cut-locus".

These terms make sense because old work on the cut-locus establishes that it is a $(m-1)$-*dimensional rectifiable set*: up to Hausdorff $(m-2)$-measure zero it is the countable union of C^1 hypersurfaces in $\mathbb{M}$. Consequently ν makes sense. Moreover $-r$ can be related to convex functions using Topogonov's theorem, so that the required one-sided derivatives exist. Finally the $(m-1)$-regularity of $C(\mathbf{o})$ means that $L^{C(\mathbf{o})}$ can be interpreted as a σ-finite local time measure.

This representation has been extended to suitable Γ-martingales (those which do not hit the conjugate part of the cut-locus) in [8, 104].

It is noteworthy that currently applications of Eq. (3.8) do not use the detailed nature of $\mathrm{d}L$ as expressed in Eq. (3.9): applications which *did* do this would be of considerable interest. For a typical application of Eq. (3.8) see Le [8, 101], who extends Γ-martingale inequalities of Darling [29] to the case of non-trivial cut-locus.

It should also be noted that the full radial representation Eq. (3.8) fills a gap in the probabilistic proof in [33] of the constancy of bounded harmonic functions on manifolds of non-negative Ricci curvature (special case of a result by Yau [137]).

Chapter 4

Coupling and parallel transport

Stochastic calculus and the probabilistic approach come into their own when considering and applying the Itô analysis of $\mathrm{dist}(X, Y)$, for X, Y two co-adapted semimartingales (typically Γ-martingales or Brownian motions). This analysis is fundamental for applications of the *coupling* idea to stochastic differential geometry.

The probabilistic notion of *coupling* is as follows. One constructs two random processes (for example two Brownian motions started at different initial points) on the same probability space, in such a way that one can draw immediate conclusions about the random process from the nature of the construction. Thus for example a partial domain monotonicity result for Neumann heat kernels in convex regions can be derived from a coupling construction of reflected Brownian motions X_1, X_2 started from two different initial points, in which the coupling is constructed so as to ensure that if X_1 is initially further from a reference point $\mathbf{o}$ than X_2 then this persists for all time [84]. A more involved coupling construction can be used to show that this convex domain monotonicity does not hold in general [9].

The original idea of coupling is due to Doeblin [34], who used it to prove convergence to equilibrium for appropriate Markov chains. It has been used intensively by workers in interacting particle systems (see for example [105]) and epidemics [7, 112]. A systematic exposition has been given by Lindvall [107] who was also the first to introduce the method in the context of Brownian motion [106]. Lindvall's *Brownian motion reflection coupling* is seminal for the theory described below: we can couple two m-dimensional Euclidean Brownian motions so that they almost surely meet at some random time! (In such a case we say that the coupling *succeeds*.) If B_1 is one of the Brownian motions, begun at $B_1(0) = b_1$, and we require to construct the other B_2 to be begun at $B_2(0) = b_2$, then this may be achieved with great simplicity just by constructing B_2 as the reflection of B_1 in the hyperplane which is the perpendicular bisector of the line segment between b_1 and b_2.

As we shall see, coupling constructions for Brownian motion can also be used to derive gradient inequalities for harmonic functions and harmonic maps, and bounds on the second eigenvalue. (Their importance has recently grown still further, with the advent of coupling-based algorithms for "perfect (or exact) simulation" introduced by Propp and Wilson [47, 96, 122], which has motivated recent work on coalescing Brownian flows [11]: however we shall say no more of that here.)

It is fundamental to the above example that B_1 and B_2 are not independent. Notice also that this coupling is *co-adapted*, in the sense that the Brownian motions can be adapted to the same filtration (so that with respect to this filtration both B_1 and B_2 have the correct first- and second-order stochastic calculus characteristics for Brownian motions). There is a whole taxonomy of different kinds of couplings which progressively weaken this latter requirement, for example *raw couplings* which would involve conditional constructions of B_2 arranged to couple with B_1 at specific times as much as possible under the constraint that B_2 *viewed on its own* should be a Brownian motion (this last kind of coupling has been used to great practical effect in [108]). We shall not consider this taxonomy here, but simply make two remarks:

1. In general the behaviour of raw couplings is more closely connected to potential theory, since existence conditions for raw couplings are typically equivalent to potential theoretic conditions such as constancy of all bounded harmonic functions;

2. On the other hand existence and nonexistence results and constructions for co-adapted couplings typically are more directly related to essentially local conditions based on curvature bounds, and adapt easily to the nonlinear theory of harmonic maps.

The second of these remarks gives the reason why we shall here concentrate on co-adapted couplings: using the machinery of stochastic development we can construct such couplings and use them to deduce geometric results arising from curvature conditions.

It is natural to follow on from Lindvall's reflection coupling for Euclidean Brownian motion to ask when the same can be done for Brownian motion on a curved manifold. Similar ideas work for Brownian motion on spheres;

viewing the sphere as embedded in Euclidean space one can use reflection in the perpendicular bisector hyperplane for the initial points b_1, b_2 viewed in the ambient Euclidean space (this is followed up in [11]). However one runs into difficulties considering hyperbolic space. If $\mathbb{H}^2$ is the hyperbolic plane then an Itô analysis of the hyperbolic distance between B_1 and B_2 shows that reflection coupling cannot succeed. In fact a symmetry argument shows that half this distance is the distance $\operatorname{dist}(B_1, \gamma)$ from B_1 to the geodesic γ acting as perpendicular bisector between b_1, b_2, and variations on the methods of the previous section show that $\operatorname{dist}(B_1, \gamma)$ can be bounded below by Brownian motion with constant positive drift, which has positive probability of drifting off to plus infinity.

Actually a potential-theoretic argument (based on the existence of non-constant bounded harmonic functions) shows that even raw couplings for $\mathrm{BM}(\mathbb{H}^2)$ must have positive probability of not succeeding. However we want to work with curvature conditions rather than potential theory, and to derive results for Γ-martingales not just Brownian motion;we therefore address the following question.

If $\mathbb{N}$ is a Cartan-Hadamard manifold with sectional curvatures bounded above by a negative constant $\kappa^2 < 0$, then when can we construct co-adapted couplings for Γ-martingales which succeed almost surely?

4.1 Distance between two Γ-martingales.

Suppose that Γ-martingales X_1, $X_2 \in \mathbb{N}$ have respective stochastic anti-developments M_1, M_2 and respective parallel transports Ξ_1, Ξ_2. Consider the Itô analysis of $S = \operatorname{dist}(X_1, X_2)$. By viewing (X_1, X_2) as a Γ-martingale on the Riemannian product manifold $\mathbb{N} \times \mathbb{N}$, and applying the geometric Itô formula Eq. (2.8) we can deduce

$$\begin{aligned} \mathrm{d}_I S \;=\; & \langle \operatorname{grad} \operatorname{dist}(X_1, X_2), (\Xi_1, \Xi_2) \rangle (\,\mathrm{d}_I M_1,\, \mathrm{d}_I M_2) + \\ & + \tfrac{1}{2} \operatorname{trace} (\operatorname{Hess} \operatorname{dist}(X_1, X_2)((\Xi_1, \Xi_2), (\Xi_1, \Xi_2))\, \mathrm{d}\,[\,(M_1, M_2), (M_1, M_2)\,]) \end{aligned} \tag{4.1}$$

until the first time that the geodesic between X_1 and X_2 is conjugate or non-unique.

Given suitable curvature bounds, the drift term trace(Hess ...) can be bounded using Eq. (2.17). We also need to control the bracket process for the martingale part $\langle \ldots \rangle (\,\mathrm{d}_I M_1,\, \mathrm{d}_I M_2)$, and for this we require some kind of bounded dilatation or quasi-conformality condition on the two martingales M_1, M_2.

To see that some sort of control on the martingale term is necessary, consider the case where X_1, X_2 are Brownian geodesics running along the same shared geodesic. Then S is bounded above by the distance between X_1, X_2 measured along the geodesic, which is a nonnegative martingale up to the time of first coupling. Basic martingale theory then shows either S converges to a positive constant or to zero. The second is easily arranged, for example by constructing X_1, X_2 to be independent, and in this case X_1, X_2 will couple with probability one *whatever* curvature conditions are imposed!

Theorem 4.1 (Γ-martingales of K^2-bounded dilatation and coupling): *Suppose that X_1 and X_2 are Γ-martingales of K^2-bounded dilatation, begun at different initial points $X_1(0) = x_1$, $X_2(0) = x_2$ of a Cartan-Hadamard manifold with negative sectional curvatures bounded above by a negative constant $-L^2$. Let the Γ-martingales have parallel transports Ξ_1, Ξ_2 and stochastic anti-developments M_1, M_2. Then X_1 and X_2 have positive probability of never meeting.*

Proof: (Sketch: see [77] for full details.) We start with the Itô analysis of S described above in Eq. (4.1). There are no conjugate or non-unique geodesics on a Cartan-Hadamard manifold, so Eq. (4.1) holds for all time. Using the bound (2.17) and completing the square, the drift term $\mathrm{d}\Lambda$ is bounded below using

$$\begin{aligned} &\operatorname{trace}\left(\operatorname{Hess}\operatorname{dist}(X_1,X_2)((\Xi_1,\Xi_2),(\Xi_1,\Xi_2))\,\mathrm{d}\,[\,(M_1,M_2),(M_1,M_2)\,]\right) \\ &\geq\; L\tfrac{\cosh(LS)-1}{\sinh(LS)}\left[(\,\mathrm{d}\tau_1 - \mathrm{d}\,[\,\langle\Xi_1 V_1, M_1\rangle\,]) + (\,\mathrm{d}\tau_2 - \mathrm{d}\,[\,\langle\Xi_2 V_2, M_2\rangle\,])\right]. \end{aligned}$$

Here $\mathrm{d}\tau_1$, $\mathrm{d}\tau_2$ are the intrinsic times for the Γ-martingales X_1, X_2; the vectors V_1, V_2 are unit vectors in $\mathbb{R}^m$ such that $\Xi_1 V_1$, $\Xi_2 V_2$ are parallel to the geodesic connecting X_1 to X_2.

The bounded dilatation condition preserves us from the possibility that $\mathrm{d}\tau_1 - \mathrm{d}\,[\,\langle\Xi_1 V_1, M_1\rangle\,]$ and $\mathrm{d}\tau_2 - \mathrm{d}\,[\,\langle\Xi_2 V_2, M_2\rangle\,]$ might both vanish in the $(\,\mathrm{d}\tau_1 + \mathrm{d}\tau_2)$-timescale. But to prove our result we must consider S in its natural time-scale, and therefore must obtain a lower bound on $\mathrm{d}\Lambda/\,\mathrm{d}\,[\,N,N\,]$ where N is the martingale part

$$\mathrm{d}_I N \quad = \quad \langle \operatorname{grad}\operatorname{dist}(X_1,X_2),(\Xi_1,\Xi_2)\rangle(\,\mathrm{d}_I M_1,\,\mathrm{d}_I M_2)\,.$$

The bounded dilatation condition and elementary but tedious arguments from Euclidean vector geometry (detailed in [77]) allow us to obtain an upper bound on $\mathrm{d}\,[\,N,N\,]\,/(\,\mathrm{d}\tau_1 + \mathrm{d}\tau_2)$, and this delivers the required result. $\square$

It is natural to ask whether the constant negative upper bound on curvature may be relaxed. So long as uniqueness of geodesics is preserved it should be straightforward to do this using the variable curvature comparisons described for example in [54]. However *some* negative curvature is required if coupling is not to be almost surely successful: see the results of the next section.

4.1.1 Coupling Brownian motions in the case of nonnegative Ricci curvature

We have seen that the reflection coupling for Brownian motions is almost surely successful in the case of Euclidean space of whatever dimension. When trying to generalize this result to Riemannian manifolds, it is reasonable to impose a condition of nonnegative *Ricci* curvature, since the rotational symmetry of the Brownian stochastic differential lends itself to Ricci curvature bounds rather than sectional curvature bounds; see the bound Eq. (2.18).

In general, excepting the constant curvature cases where there is much symmetry, we cannot make sense of reflection couplings using single fixed perpendicular bisector hyperplanes. However notice that the purpose of the reflection is simply to transform one Brownian differential $\mathrm{d}B_1$ into another $\mathrm{d}B_2$. Instead of doing this using a fixed hyperplane, we can use parallel transport along a minimal geodesic connecting the two Brownian motions, together with reflection in the hyperplane normal to this geodesic. This is the essence of the proof of the following result.

Theorem 4.2 (Brownian coupling and nonnegative Ricci curvature): *Let $\mathbb{M}$ be a Riemannian manifold of nonnegative Ricci curvature. Given any two initial points x_1, x_2, it is possible to construct two co-adapted Brownian motions X_1, X_2 begun at x_1, x_2 respectively, such that they couple successfully with probability one.*

Proof: (Sketch: for full details consult any one of [22, 76, 81].) Let $S = \operatorname{dist}(X_1, X_2)$, and let Ξ_1, Ξ_2 be the stochastic parallel transports and B_1, B_2 be the stochastic anti-developments of X_1, X_2. We can use the Itô analysis of Eq. (4.1), but only till the first time that the minimal geodesic connecting X_1, X_2 is conjugate or non-unique.

Setting aside this difficulty for a moment, we have to consider how to express the transformation of $\mathrm{d}B_1$ into $\mathrm{d}B_2$. Let the *mirror map* $\mathfrak{M}_{(x_1,x_2)}$ map the tangent space $T_{x_1}\mathbb{M}$ to the tangent space $T_{x_2}\mathbb{M}$ using parallel transport along the minimal geodesic from x_1 to x_2 followed by reflection in the hyperplane normal to the geodesic. Then we define $\mathrm{d}_I B_2 = [(\Xi_2)^{-1}\,\mathfrak{M}_{(X_1,X_2)}\,\Xi_1]\,\mathrm{d}_I B_1$. Notice that $[(\Xi_2)^{-1}\,\mathfrak{M}_{(X_1,X_2)}\,\Xi_1]$ is a (random but predictable) linear mapping of $\mathbb{R}^m$ into itself, despite involving maps between tangent spaces on manifolds (compare the use of

$\langle \operatorname{grad} f(X), \Xi \rangle \, \mathrm{d}_I Z$ in Theorem 2.2). Thus it is proper to use it as the integrand for an Itô differential.

We can now define the coupling between X_1 and X_2 up to the first time that the geodesic between the two Brownian motions fails to be unique, using a stochastic differential system which mixes Itô and Stratonovich terms:

$$\begin{aligned}
\mathrm{d}_S X_1 &= \Xi_1 \, \mathrm{d}_S B_1 \\
\mathrm{d}_S \Xi_1 &= H_{\Xi_1} \, \mathrm{d}_S X_1 \\
\mathrm{d}_S X_2 &= \Xi_2 \, \mathrm{d}_S B_2 \\
\mathrm{d}_S \Xi_2 &= H_{\Xi_2} \, \mathrm{d}_S X_2 \\
\mathrm{d}_I B_2 &= \left[(\Xi_2)^{-1} \, \mathfrak{M}_{(X_1, X_2)} \, \Xi_1 \right] \mathrm{d}_I B_1
\end{aligned} \tag{4.2}$$

Substituting in to the Itô analysis Eq. (4.1) of S, we find the effect of the mirror map is to make the martingale part become

$$\langle \operatorname{grad} \operatorname{dist}(X_1, X_2), (\Xi_1, \Xi_2) \rangle (\, \mathrm{d}_I B_1, \, \mathrm{d}_I B_2) \quad = \quad 2 \langle V_1, \, \mathrm{d}_I B_1 \rangle$$

where $\Xi_1 V_1$ is the unit tangent vector at X_1 defining the geodesic from X_1 to X_2. So the martingale part of S is simply a scalar Brownian motion of rate $2^2 = 4$.

Using the mirror map, the drift part of S simplifies to half a sum of $m-1$ second geodesic derivatives of $\operatorname{dist}(y_1, y_2)$, evaluated at $y_1 = X_1$, $y_2 = X_2$, where a typical geodesic variation causes y_1 to move by a geodesic defined by $\Xi_1 V$ and causes y_2 to move by a geodesic defined by $\mathfrak{M}_{(X_1, X_2)} \, \Xi_1 V$, and V ranges through an orthonormal set of unit vectors such that $\Xi_1 V$ is normal to the geodesic from X_1 to X_2. This is exactly the context in which we can apply the bound (2.18), and results in the conclusion that the quantity

$$\begin{aligned}
\frac{1}{2} \sum_i \Big(& \operatorname{Hess} \operatorname{dist}(X_1, X_2)((\Xi_1, \Xi_2), (\Xi_1, \Xi_2)) \times \\
& \times \mathrm{d} \left[\, (M_1, M_2), (M_1, M_2) \, \right] \Big)
\end{aligned} \tag{4.3}$$

is negative.

It follows that, up till the first time that uniqueness or non-conjugacy fails for the minimal geodesic from X_1 to X_2, we may bound S above by a scalar Brownian motion of rate 4. If this bound could persist for all time then we would have almost sure successful coupling, since scalar

Brownian motion will eventually hit zero. The question is, how to deal with breakdown of the stochastic system Eq. (4.2)?

The original approach of [76, 81] involves a complicated accounting for the different ways in which breakdown can happen. Failure of uniqueness of minimal geodesic can be disposed of with relative ease. Failure of non-conjugacy is fixed by allowing the two Brownian motions X_1, X_2 to evolve independently for a short time near the locus of conjugacy; analysis of Jacobi vectorfields shows that for independent evolution near conjugacy the drift of S is nearly negatively infinite. A simpler and more attractive treatment in [22] finesses matters: again one allows the two Brownian motions to evolve independently for brief intervals when the system Eq. (4.2) breaks down, but now one argues using the ideas of the cut-locus correction described above in §. This shows that in the limit, as the brevity of the brief intervals shrinks to zero, so one achieves a construction which gives S the characteristics of a scalar Brownian motion of rate 4 with a nonpositive drift. This then establishes that the coupling is almost surely successful. $\square$

Again one may ask whether the curvature condition may be relaxed, now to allow regions of negative Ricci curvature. This has been achieved recently by Wang [135].

4.2 Applications

We comment briefly on applications of the idea of coupling Brownian motions on manifolds.

We begin by remarking that some caution is appropriate when making simplifying assumptions on the geometry of manifolds. It is tempting to wonder what might be lost in Theorem 4.2 simply by assuming that, in addition to nonnegative Ricci curvature, any two points are connected by a unique geodesic. However the celebrated splitting theorem of Cheeger and Gromoll [15] implies that the only such manifolds are Euclidean spaces! (A proof using Brownian motion and the cut-locus correction may be found in [82].)

Given the work above, we can immediately establish a toy version of Yau's [137] result on constancy of nonnegative harmonic functions on manifolds of $\mathbb{M}$ of nonnegative Ricci curvature. For suppose that $f : \mathbb{M} \to [0, 1]$ is a *bounded* harmonic function. For $x_1, x_2 \in \mathbb{M}$ consider Brownian motions X_1, X_2 constructed as above so as to begin at x_1, x_2 and to couple almost surely. Then $f(X_1) - f(X_2)$ is a bounded martingale which vanishes at time infinity. Basic martingale results tell us that $\mathbb{E}\left[\, f(X_1(t)) - f(X_2(t))\,\right] =$

$f(x_1) - f(x_2)$, and also

$$\mathbb{E}\left[\, f(X_1(t)) - f(X_2(t)) \,\right] \quad = \quad \lim_{s\to\infty} \mathbb{E}\left[\, f(X_1(s)) - f(X_2(s)) \,\right] \quad = \quad 0\,.$$

Hence $f(x_1) = f(x_2)$.

Of course establishing successful coupling involves more work than the alternative probabilistic proof referred to earlier at the end of §3.6.2. However the coupling approach generalizes to allow one to establish probabilistic proofs of gradient inequalities [22]: here a rather careful analysis is required not only of the distance $\operatorname{dist}(X_1, X_2)$ but also, simultaneously, of the two distances $\operatorname{dist}(X_1, \mathbf{o})$ and $\operatorname{dist}(X_2, \mathbf{o})$. Careful estimation bounds the probability that X_1 and X_2 couple before leaving a small geodesic ball centred on $\mathbf{o}$, bounding in terms of the initial distance $\operatorname{dist}(X_1(0), X_2(0))$. The gradient inequality then follows by a simple martingale argument somewhat on the lines above. (See [133, 132] for an alternative probabilistic approach to gradient estimates using differentiation formulae.)

The coupling approach also permits a very satisfactory probabilistic proof of the original generalized Picard little theorem for harmonic maps (see Eq. 3.6). For if $\mathbb{M}$ has nonnegative Ricci curvatures then for $x_1, x_2 \in \mathbb{M}$ we can construct co-adapted Brownian motions X_1, X_2 which begin at x_1, x_2 and which couple almost surely. Suppose that $F : \mathbb{M} \to \mathbb{N}$ is a harmonic map of K^2-bounded dilatation from $\mathbb{M}$ into a Cartan-Hadamard manifold $\mathbb{N}$ of sectional curvatures bounded above by a negative constant. Then $F(X_1)$ and $F(X_2)$ are co-adapted Γ-martingales which couple almost surely. By Theorem 4.1 above it follows $F(x_1) = F(x_2)$, and hence we may deduce that F is constant.

Notice that the above proof does not require the irritating lower bound on curvatures of $\mathbb{N}$ which was a feature of the angle-based methods described in the previous section. The price paid in return is a more restrictive condition on $\mathbb{M}$: the angle-based methods required only that the universal cover of $\mathbb{M}$ support no non-constant bounded harmonic functions. Clearly the coupling proof generalizes beyond nonnegative Ricci curvature (Wang's result [135] demonstrates this): we only require that $\mathbb{M}$ have the *Brownian coupling property* set out below.

Definition 4.3 (Brownian coupling property): *A Riemannian manifold $\mathbb{M}$ has the* Brownian coupling property *(BCP) if for all x_1, $x_2 \in \mathbb{M}$ we can construct co-adapted Brownian motions X_1, X_2 begun at x_1, x_2 which almost surely couple successfully.*

We have just seen, in the proof of the toy version of Yau's result described above, that the BCP implies that all bounded harmonic functions are constant. It would be most interesting to find an example of a manifold for which all bounded harmonic functions are constant but the BCP fails. (Apart from intrinsic interest this might help guide current investigations into the achievable limits of efficient perfect simulation [11, 12].)

We note in passing that the possibilities for co-adapted coupling of processes on manifolds are more extensive than one might at first glance believe: see [5] which demonstrates successful coupling for some basic examples of hypoelliptic diffusions.

Further geometric applications of coupling include analysis of the ends of a manifold of nonnegative Ricci curvature [23], and a factorization theorem for harmonic maps arising from a factorization result for Brownian motion on a vector bundle [42]. In the next section we describe how coupling ideas combined with convexity considerations take one much further in a probabilistic treatment of harmonic maps. Before turning to this we should also note the important bounds on spectral gaps obtained by Chen and others using these coupling ideas [16, 17].

Chapter 5

Harmonic map Dirichlet problem

Probability has much to contribute to the question of how to solve the Dirichlet problem for harmonic functions. If we are required to extend $f : \partial\Omega \to [0,1]$ from the boundary $\partial\Omega$ to the whole of a domain Ω, then the probabilistic approach delivers a stochastic mean-value representation

$$f(x) \quad = \quad \mathbb{E}\left[\, f(X(T)) \mid X(0) = x \,\right]$$

(here X is a Brownian motion run until the random time T when it first leaves the domain Ω) which is valid for "sensible" domains Ω, and also gives a stochastic interpretation of what "sensible" means (domains for which the boundary $\partial\Omega$ corresponds to the different ways in which X can leave Ω) and a way to fix things up when the domain is not "sensible" (redefine the boundary $\partial\Omega$ using the Martin boundary for the diffusion X).

But what about harmonic maps? The above approach is road-blocked: harmonic maps are defined using a *nonlinear* Dirichlet form as in Eq. 1.2. If we attempt to mimic the probabilistic construction for the linear case we come up against problems:

- How should one define expectation (a linear operation) on a nonlinear manifold?
- Even if one can make such a definition, the "law of iterated expectations" fails.

In this chapter it is shown how nevertheless the Γ-martingale approach can be developed into methods for solving Dirichlet problems for harmonic maps.

5.1 The rôle of convexity and coupling

The rôle of convexity in stochastic differential geometry has been strongly elucidated by Emery and Zheng [44, 46]. It is fundamental for harmonic map Dirichlet problems.

First observe that the convexity-based ideas of the Ultimate Liouville Theorem (Theorem 3.3), together with coupling ideas, suggest a strategy for determining when one has *uniqueness* for a harmonic map Dirichlet problem. Consider a fixed codomain $\mathcal{B} \subset \mathbb{N}$, a domain $\mathbb{M}$ which is a compact Riemannian manifold with ("sensible"!) boundary $\partial\mathbb{M}$, and a Dirichlet problem defined by data of a boundary map $\partial F : \partial\mathbb{M} \to \mathcal{B}$. Two different extensions F_1, F_2 of ∂F lead to two different Γ-martingales $Y_1 = F_1(\mathrm{BM}(\mathbb{M}))$, $Y_2 = F_2(\mathrm{BM}(\mathbb{M}))$ which necessarily agree once the $\mathrm{BM}(\mathbb{M})$ in question has hit the boundary (and part of the definition of a "sensible" boundary is that $\mathrm{BM}(\mathbb{M})$ *will* eventually hit the boundary!). So non-uniqueness for the Dirichlet problem implies that one can construct two Γ-martingales in $\mathcal{B}$ with the same end-value.

Now of course in Theorem 3.3 the end-value was a non-random end-point on $\mathcal{B}$, allowing us to compose with a bounded nonnegative convex function vanishing only at the end-point. In general this will not be the case here. However we *will* have a product Γ-martingale (Y_1, Y_2) on $\mathcal{B} \times \mathcal{B}$, which will surely hit the diagonal $\mathcal{D} = \{(x, x) : x \in \mathcal{B}\}$ when the individual Γ-martingales couple, as we know they must do. So if we can exhibit a bounded nonnegative convex function vanishing only on the diagonal $\mathcal{D}$ then the argument of Theorem 3.3 allows us to deduce that (Y_1, Y_2) must lie on the diagonal for all time. This means $Y_1 = Y_2$ for all time, and hence (bearing in mind we can choose the starting point of the $\mathrm{BM}(\mathbb{M})$ arbitrarily) we deduce $F_1 = F_2$ hence uniqueness for the Dirichlet problem.

This simple conceptual argument motivates the next definition and justifies the two results following.

Definition 5.1 (Convex geometry): *A subset $\mathcal{B} \subset \mathbb{N}$ of a Riemannian manifold $\mathbb{N}$ is said to possess* convex geometry *if the following holds:*

(1) $\mathcal{B}$ is convex;

(2) there is a nonnegative convex function

$$\Phi : \mathcal{B} \times \mathcal{B} \to [0, 1]$$

such that Φ vanishes only on the diagonal: $\Phi(x_1, x_2) = 0$ if and only if $x_1 = x_2$.

Theorem 5.2 (Γ-martingale uniqueness): *Suppose that $\mathcal{B} \subset \mathbb{N}$ possesses convex geometry. If Y_1, Y_2 are two Γ-martingales on $\mathcal{B}$ which almost surely couple successfully then they must be almost surely equal for all time.*

Theorem 5.3 (Uniqueness for small-image harmonic map Dirichlet problem): *Suppose that $\mathcal{B} \subset \mathbb{N}$ possesses convex geometry. Then there can be at most one harmonic map extension to any boundary data $\partial F : \partial\mathbb{M} \to \mathcal{B}$.*

These results deliver a probabilistic proof of the existence result for the Dirichlet problem for harmonic maps as expounded in [6, 55].

It should be noted that in Theorem 5.3 the boundary $\partial\mathbb{M}$ must be "sensible" in the undefined sense mentioned above: however it is easy to see in this probabilistic approach that the linear potential theory boundary is exactly the correct boundary for the harmonic map problem.

There is of course the important question of whether the definition above is largely vacuous. Convex geometry is possessed by all sufficiently small geodesic balls, as a consequence of an argument in [44], though this is not quantified. However in [85, 88] an explicit construction is given of a Φ which works for all regular geodesic balls $\mathcal{B}$: exactly the context in which uniqueness has been obtained by analytic means using partial differential equation theory [63]:

Theorem 5.4 (Convex geometry for regular geodesic balls): *Suppose that $\mathcal{B} \subset \mathbb{N}$ is a regular geodesic ball, with nonnegative upper curvature bound k, centre $\mathbf{o}$ and radius $\arccos(h) < \pi/(2\sqrt{k})$. Then convex geometry for $\mathcal{B}$ is provided by*

$$\Phi(x,y) \quad = \quad \left(\frac{1-\cos(\sqrt{k}\,\mathrm{dist}(x,y))}{\cos(\sqrt{k}\,\mathrm{dist}(x,\mathbf{o}))\cos(\sqrt{k}\,\mathrm{dist}(y,\mathbf{o}))-\tilde{h}^2}\right)^{\nu+1} \qquad (5.1)$$

if $\nu \geq 1$ and $2\nu\tilde{h}^2(h^2-\tilde{h}^2) \geq 1$.

Further work in [85] establishes a continuous dependence of the initial point $Y(0)$ of a Γ-martingale Y in $\mathcal{B}$ on its terminal value $Y(\infty)$ (viewed in the L^2 sense): this is a consequence of local Lipschitz regularity of the convex function endowing $\mathcal{B}$ with convex geometry.

5.2 Construction of Γ-martingales

Given the above, when the co-domain $\mathcal{B}$ has convex geometry we can reduce the existence aspect of the harmonic map Dirichlet problem to a purely probabilistic question, a *Γ-martingale reachability problem*:

- Given a random variable Z taking values in a region $\mathcal{B}$ of convex geometry, can one construct a Γ-martingale Y for which Z is a terminal value?

For if we can do this then for each $x \in \mathbb{M}$ we can construct the unique Γ-martingale $Y^{(x)}$ in $\mathcal{B}$ attaining $\partial F(X^{(x)}(t))$, where $X^{(x)}$ is BM($\mathbb{M}$) begun at x. Hence we obtain a map $F : x \mapsto Y^{(x)}(0)$. Arguments involving the Markov property of BM($\mathbb{M}$), and aided by the continuous dependence mentioned above, then show that F is a *finely harmonic map* (see Definition 2.4) and indeed continuous. This solves the harmonic map Dirichlet problem at least up to the question of whether we can translate "finely harmonic" to "harmonic": by the Itô formula this resolves to the regularity question of whether F is actually twice differentiable. Actually the same issue applies for the probabilistic resolution of the ordinary harmonic function Dirichlet problem: we will say more about this below.

We have to be careful here about what is prescribed in the Γ-martingale reachability problem: in the context of harmonic maps we are prescribing the proposed terminal value Z as a function $\partial F(X(T))$ of a Riemannian Brownian motion, consequently Z *must* be viewed as a random variable defined on a probability space $(\Omega, \mathfrak{F}, \{\mathfrak{F}_t : t \geq 0\}, \mathbb{P})$. In fact this filtered probability space may be viewed as generated by the Riemannian Brownian motion, and the stochastic differential system Eq. (2.1) of stochastic development and parallel transport permits a "flattening" simplification of supposing the generating process to be ordinary Euclidean Brownian motion. (The case of construction of the Γ-martingale when the filtration is *not* specified is covered in [45].)

One way of proceeding from this point, as described in detail in [85], is to exploit the "flattening" simplification together with continuous dependence as follows. Viewing Z as defined on a probability space $(\Omega, \mathfrak{F}, \{\mathfrak{F}_t : t \geq 0\}, \mathbb{P})$ generated by a Euclidean Brownian motion X, we may consider it as a function of the Brownian path: $Z = G(X)$. Because of the continuous dependence property it is enough to find a solution when Z depends on only finitely many values of X: $Z = G(X(t_1), \ldots, X(t_k))$; moreover we may suppose G to be a smooth map, constant off a compact set. The Markov property can now be employed to reduce this to the case $k = 1$ and so (by Brownian scaling) to the case $t_1 = 1$: $Z = G(X(1))$.

5.2.1 Existence using barycentres

Finally we must construct a sequence of approximate Γ-martingale solutions to the reachability problem when $Z = G(X(1))$. We do this using *Riemannian centres of mass* or *barycentres*, a concept due originally to

Cartan [14] (see also [49, 71]).

Definition 5.5 (Riemannian barycentres): *Let $\mathcal{B} \subset \mathbb{N}$ be a convex neighborhood in a Riemannian manifold, and let μ be a probability measure on $\mathcal{B}$. We say that $x \in \mathcal{B}$ is a* (Riemannian) barycentre of μ *if it is a local minimum of*

$$Q(x) \quad = \quad \frac{1}{2}\int_{\mathcal{B}} \operatorname{dist}(x,u)^2 \; \mu(\,\mathrm{d}u)\,.$$

If the minimum is unique we write

$$x \quad = \quad \mathcal{E}(\mu)\,.$$

By extension to random variables we write $x = \mathcal{E}(Z)$ if x is the unique barycentre of the distribution of $Z \in \mathcal{B}$, and let $\mathcal{E}(Z|\mathfrak{F}_s)$ be the $\mathfrak{F}_s$-adapted random variable which is the unique barycentre of the conditional distribution of Z given $\mathfrak{F}_s$.

We note in passing that the notion of barycentre can be made to depend only on the connection if we consider extremals, satisfying

$$\int_{\mathcal{B}} \operatorname{dist}(x,u) \operatorname{grad}_1 \operatorname{dist}(x,u) \; \mu(\,\mathrm{d}u) \quad = \quad 0\,.$$

This is discussed in [45]. Furthermore the notion of barycentre has recently found application in statistical geometry [114] and the statistics of shape [102].

Summarizing the argument of [85], view the barycentre as an approximate expectation (albeit *not* obeying the crucial iterated expectation property: $\mathcal{E}(\mathcal{E}(Z|\mathfrak{F}_s)) \neq \mathcal{E}(Z)$ in the presence of curvature [45]) it is now clear how to construct approximate Γ-martingales which reach the required target $Z = G(X(1))$: simply chain together conditional barycentres: set $Y^{[n]}(1) = Z$ and

$$Y^{[n]}(k/n) \quad = \quad \mathcal{E}(Y^{[n]}((k+1)/n)|\mathfrak{F}_{k/n})\,.$$

These constructions can now be used to define "space-time maps" $F^n : \mathbb{R}^m \times [0,1] \to \mathcal{B}$ by piecewise geodesic interpolation from

$$Y^{[n]}(k/n) \quad = \quad F^n(X(k/n), k/n)\,.$$

We can now do simple calculations with Euclidean Brownian motion exploiting the convex function Φ endowing $\mathcal{B}$ with convex geometry; these show that the family of maps F^n is equicontinuous and hence one can apply the Arzela-Ascoli theorem to deduce a converging subsequence with a limit $F : \mathbb{R}^m \times [0,1] \to \mathcal{B}$. But the resulting random process $Y(t) = F(X(t), t)$

yields a submartingale whenever composed with a convex function; hence [26] it is a Γ-martingale, which clearly has the required reachability property!

The attractions of this approach are as follows:

- It offers easy extensions to the case when the domain is furnished not with Riemannian geometry, but with a suitably regular diffusion. Essentially what one requires is that the diffusion can be co-adaptively coupled: this includes smooth elliptic diffusions but also at least some hypoelliptic diffusions, as demonstrated in [5].
- It connects very directly with linear potential theory. For example note that [85] provides a theorem which systematically relates the question of existence of harmonic map Fatou radial limits to the linear case, thus generalizing analytically based work of [6].
- It offers an intriguing set of equivalences, implications, and counterexamples relating harmonic maps to convexity, Γ-martingales, and geodesics: see the last subsection in this account for more details.

5.2.2 Other approaches

We remark briefly on a variety of other probabilistic approaches uncovered in the last five years (indeed this variety contrasts remarkably with the apparent inability of probability to contribute to existence questions for harmonic maps before 1990!).

Picard [117] uses techniques based on Malliavin calculus to derive existence results. In a second approach Picard formalizes the notion of "Γ-martingale with jumps" (resembling the Γ-martingale approximations described above, but more sophisticated) and conscripts this into the service of harmonic maps [116, 118]. See also Cohen [19] and Arnaudon [3].

It is possible to rewrite the equations of stochastic development and parallel transport for a Γ-martingale $Y = F(X)$ and its generating Brownian motion X, when F is a harmonic map, as a *backwards stochastic differential equation*. This idea, introduced by Pardoux and Peng [115] in the context of stochastic control, is of strikingly wide application. Unfortunately the standard method of solution for backwards sde requires the use of the contraction mapping principle, and this breaks down in the harmonic map / Γ-martingale case because of a quadratic dependence arising from the Levi-Civita connection. Notwithstanding, in a *tour de force* involving a clever localization Darling [30] shows how to use backwards sde theory to construct Γ-martingales and hence harmonic maps. In [31] he extends the methods to deal with Γ-martingales on non-compact manifolds.

Finally Thalmaier [131] describes an approach which is different yet again, this time addressing the question of probabilistic approaches to harmonic map Dirichlet problems where uniqueness fails. Clearly Γ-martingale reachability problems also fail to have unique solutions in such cases: Thalmaier analyses these by considering the martingale solving the reachability problem with the least mean intrinsic time, and relates this to harmonic map singularities.

Although not strictly a probabilistic approach, the work of Jost [65], described elsewhere in this volume, has strong links to probability. This approach constructs harmonic maps *directly* using barycentres, in a manner reminiscent of the method [85] discussed above. Notice however that Jost's approach works for metric spaces which are more general than manifolds: the question of extending Γ-martingale theory to encompass Jost's approach raises intriguing questions concerning the "correct" definitions of Γ-martingale and convexity.

5.3 Smoothness *via* gradient estimates

The probabilistic methods described above all deliver (continuous) finely harmonic maps rather than harmonic maps, and as mentioned above it is an interesting question how one can produce a probabilistic approach which delivers regularity. This has been answered in [94], which uses coupling ideas to deliver gradient estimates for finely harmonic maps (from which full regularity is easily deduced). The idea is based on the work of Cranston [22] with two interesting variations:

1. The arguments of Cranston are simplified by using co-adapted couplings which involve *time-changed* Brownian motions. In two dimensions, a Brownian motion subjected to a Markov time-change can be recognized as a Brownian motion in a conformally equivalent metric. This is no longer the case in higher dimensions: one must add a non-vanishing drift. However in the case of gradient estimates it can be shown that this extra drift assists coupling to happen, and therefore does not affect the resulting gradient estimate. This allows a simplification of the Cranston argument, as one needs to deal only with the distance between the two coupled processes and not their distances from some reference point.

2. In the case of harmonic maps whose image may possess positive curvature, these arguments only deliver Hölder($\frac{1}{2}$) estimates, not sufficient for regularity. However a careful modification of reflection coupling, which actually introduces independence as the two Γ-martingales

come closer together, can be deployed to bootstrap the Hölder($\frac{1}{2}$) estimate up to a full gradient estimate.

5.4 Equivalences

To close this account we mention briefly some work which ties together convexity, coupling of Γ-martingales, harmonic maps, and stochastic control.

The very term "stochastic reachability" suggest control theory ideas, and indeed *any* coupling construction can be viewed as the solution of a (typically non-standard) stochastic control problem, aiming to maximize the probability of a successful coupling. The *value* of such a control problem turns out to be exactly the best candidate for a function to produce convex geometry. In [89, 91] it is shown that a convex domain $\mathcal{B}$ has convex geometry if and only if the Γ-martingale reachability problem has unique solutions depending continuously on the data.

Given the above work it is natural to wonder whether this equivalence extends to include the harmonic map Dirichlet problem with values in $\mathcal{B}$, and though no results have been obtained on this nevertheless one feels it should be the case. Emery [44] made the bold conjecture that the equivalence is implied by a much simpler phenomenon: that any two points in $\mathcal{B}$ are connected by one and only one geodesic (one could view this as a geodesic Dirichlet problem!). In fact this is not the case: [90] provides a two-dimensional counterexample, the "Propeller", using Riemannian barycentres. (A further related counterexample has been derived which is of significance in statistical geometry: [20].)

Bibliography

[1] A. Ancona. Théorie du potentiel sur les graphes et les variétés. In *Ecole d'Eté de Probabilités de Saint-Flour XVIII-1988*, volume 1427 of *Lecture Notes in Mathematics*, pages 5–109, New York–Heidelberg–Berlin, 1990. Springer-Verlag.

[2] A. Ancona. Convexity at infinity and Brownian motion on manifolds with unbounded negative curvature. *Rev. Mat. Iberoamericana*, 10:189–220, 1994.

[3] M. Arnaudon. Sur l'existence dans une variété de martingales de valeur terminale donné. Research report, IRMA and University Louis Pasteur, Strasbourg, 1993.

[4] M. Arnaudon and A. Thalmaier. Complete lifts of connections and stochastic Jacobi fields. Research report, IRMA and University Louis Pasteur, Strasbourg, 1997.

[5] G. Ben Arous, M. Cranston, and W.S. Kendall. Coupling constructions for hypoelliptic diffusions: Two examples. In M. Cranston and M. Pinsky, editors, *Stochastic Analysis: Summer Research Institute July 11-30, 1993.*, volume 57, pages 193–212, Providence, RI Providence, 1995. American Mathematical Society.

[6] P. Aviles, H. Choi, and M. Micallef. Boundary behavior of harmonic maps on non-smooth domains and complete negatively curved manifolds. *Journal of Functional Analysis*, 99:293–331, 1991.

[7] F.G. Ball, D. Mollison, and G. Scalia-Tomba. Epidemics with two levels of mixing. *Annals of Applied Probability*, 7:46–89, 1997.

[8] D. Barden and H.L. Le. Some consequences of the nature of the distance function on the cut locus in a Riemannian manifold. Research report, Mathematics Department, University of Nottingham, Nottingham, 1995.

[9] R. Bass and K. Burdzy. On domain monotonicity of the Neumann heat kernel. *Journal of Functional Analysis*, 116:215–224, 1993.

[10] A. Bernard, E.A. Campbell, and A.M. Davie. Brownian motion and generalized analytic and inner functions. *Ann. Institute Fourier*, 29:207–228, 1979.

[11] K. Burdzy and W.S. Kendall. Brownian coalescence on high-dimensional spheres. In preparation, 1997.

[12] K. Burdzy and W.S. Kendall. Efficient Markovian couplings. In preparation, 1997.

[13] E. Calabi. An extension of E. Hopf's maximum principle with an application to geometry. *Duke Mathematical Journal*, 25:45–56, 1958.

[14] H. Cartan. *Leçons sur la géometrie des espaces de Riemann*. Gauthiers-Villars, Paris, 1928.

[15] J. Cheeger and D. Gromoll. The splitting theorem for manifolds of nonnegative Ricci curvature. *J. Differential Geometry*, 6:119–128, 1971.

[16] M.-F. Chen. Optimal couplings and application to Riemannian geometry. In B. Grigelionis *et al.*, editor, *Probability theory and Mathematical Statistics*, volume 1, page 15 pp. VPS / TEV, 1994.

[17] M.-F. Chen and F.-Y. Wang. Application of coupling method to the first eigenvalue on manifold. *Science in China (Ser. A)*, 37:1–14, 1994.

[18] S.-Y. Cheng. Liouville theorem for harmonic maps. *Proc. Symp. Pure Math.*, 36:147–151, 1980.

[19] S. Cohen. Géometrie différentielle avec sauts. *Comptes Rendus des Séances de l'Académie des Sciences. Série I. Mathématique*, 314:767–770, 1992.

[20] J.M. Corcuera and W.S. Kendall. Riemannian barycentres and geodesic convexity. Research report 325, Department of Statistics, University of Warwick, 1998. Submitted for publication.

[21] R. Courant, K. Freidrichs, and H. Lewy. Über die partiellen differenzengleichungen der mathematischen physik. *Mathematische Annalen*, 100:32–74, 1928.

[22] M. Cranston. Gradient estimates on manifolds using coupling. *Journal of Functional Analysis*, 99:110–124, 1991.

[23] M. Cranston. A probabilistic approach to Martin boundaries for manifolds with ends. *Probability Theory and Related Fields*, 96:319–333, 1993.

[24] M. Cranston, W.S. Kendall, and Yu. Kifer. Gromov's hyperbolicity and Picard's little theorem for harmonic maps. In I.M. Davies, A. Truman, and K.D. Elworthy, editors, *Stochastic Analysis and Applications, Gregynog, (*9^{th}-14^{th} *July 1995)*, pages 139–164, Singapore, 1996. World Scientific Press.

[25] M. Cranston, W.S. Kendall, and P. March. The radial part of Brownian motion II: Its life and times on the cut locus. *Probability Theory and Related Fields*, 96:353–368, 1993.

[26] R.W.R. Darling. Martingales in manifolds - definitions, examples, and behaviour under maps. In *Séminaires des Probabilités XVI supplément, Springer Lecture Notes in Mathematics*, volume 921, pages 217–236. Springer-Verlag, New York-Heidelberg-Berlin, 1982.

[27] R.W.R. Darling. Convergence of martingales on a Riemannian manifold. *Kyoto University. Research Institute for Mathematical Sciences. Publications*, 19:753–763, 1983.

[28] R.W.R. Darling. Convergence of martingales on manifolds of negative curvature. *Annales de l'Institut Henri Poincaré. Probabilités et Statistique*, 21:157–175, 1985.

[29] R.W.R. Darling. Exit probability estimates for martingales in geodesic balls, using curvature. *Probability Theory and Related Fields*, 93:137–152, 1992.

[30] R.W.R. Darling. Constructing gamma-martingales with prescribed limit, using backwards sde. *The Annals of Probability*, 23:1234–1261, 1995.

[31] R.W.R. Darling. Martingales on non-compact manifolds: maximal inequalities and prescribed limits. *Annales de l'Institut Henri Poincaré. Probabilités et Statistique*, 32:431–454, 1996.

[32] B. Davies. Picard's theorem and Brownian motion. *Trans. Amer. Math. Soc.*, 213:353–362, 1975.

[33] A. Debiard, B. Gaveau, and E. Mazet. Théorèmes de comparaison en geometrié Riemannienne. *Kyoto University. Research Institute for Mathematical Sciences. Publications*, 12:391–425, 1976.

[34] W. Doeblin. Exposé de la theorie des chaines simples constantes de Markov à un nombre fini d'etats. *Rev. Math. Union Interbalkanique*, 2:77–105, 1938.

[35] F. Duheille. On the range of $\mathbb{R}^2$ or $\mathbb{R}^3$-valued harmonic morphisms. *Comptes rendus de l'academie des sciences serie I - mathematique*, 320:1495–1500, 1995.

[36] F. Duheille. A probabilistic elementary proof of a result of P. Baird and J.C. Wood. *Annales de l'institut Henri Poincaré - probabilités et statistiques*, 33:283–291, 1997.

[37] T.E. Duncan. Estimation for jump processes in the tangent bundle of a Riemannian manifold. *Applied Mathematics and Optimization*, 4:265–274, 1978.

[38] T.E. Duncan. Stochastic systems in Riemannian manifolds. *Journal of Optimization Theory and Applications*, 27:399–426, 1979.

[39] E.B. Dynkin. Non-negative eigenfunctions of the Laplace-Beltrami operator and Brownian motion in certain symmetric spaces. *Doklady Akademii Nauk SSSR*, 141:288–291, 1961.

[40] J. Eells and L. Lemaire. *Selected Topics in Harmonic Maps.* American Mathematical Society, Providence, RI, 1983.

[41] K.D. Elworthy. *Stochastic differential equations on manifolds.* Cambridge University Press, Cambridge, 1982.

[42] K.D. Elworthy and W.S. Kendall. Factorizations of Brownian motions and harmonic maps. In K.D. Elworthy, editor, *From local times to global geometry, control, and physics*, volume 150 of *Pitman Research Notes in Mathematics*, pages 75–83. Longmans, 1986.

[43] M. Emery. Convergence des martingales dans les variétés. In *Colloque en l'Honneur de Laurent Schwartz (vol 2)*, volume 132, pages 47–62. Société Mathématique de France, 1985.

[44] M. Emery and P.A. Meyer. *Stochastic Calculus in Manifolds, with an Appendix by P.A. Meyer.* Springer-Verlag, New York-Heidelberg-Berlin, 1989.

[45] M. Emery and G. Mokobodzki. Sur le barycentre d'une probabilité dans une variéte. In *Séminaire de Probabilités XXV, Lecture Notes in Mathematics*, volume 1485, pages 220–233. Springer-Verlag, New York-Heidelberg-Berlin, 1991.

[46] M. Emery and W.A. Zheng. Fonctions convexes et semimartingales dans une variété. In *Séminaire de Probabilités XVIII, Lecture Notes in Mathematics*, volume 1059, pages 501–518. Springer-Verlag, New York-Heidelberg-Berlin, 1984.

[47] J. Fill. An interruptible algorithm for exact sampling via Markov Chains. *Annals of Applied Probability*, to appear, 1997. *WWW* URL: `www.mts.jhu.edu/~fill/`.

[48] R.A. Fisher. Dispersion on a sphere. *Proc. Roy. Soc. London*, A217:295–305, 1953.

[49] M. Fréchet. Les éléments aléatoires de nature quelconque dans un espace distancié. *Ann. Inst. Henri Poincaré*, 10:215–310, 1948.

[50] M. Fukushima and M. Okada. On Dirichlet forms for plurisubharmonic functions. *Acta Math.*, 159:171–214, 1984.

[51] R. Gangolli. On the construction of certain diffusions on a differentiable manifold. *Zeitschrift für Wahrscheinlichkeitstheorie und verwe Gebiete*, 2:406–419, 1964.

[52] S. Goldberg, T. Ishihara, and N.C. Petridis. Mappings of bounded dilatation of Riemannian manifolds. *J. Diff. Geom.*, 10:619–630, 1975.

[53] S. Goldberg and C. Mueller. Brownian motion, geometry, and generalizations of Picard's little theorem. *The Annals of Probability*, 11:833–846, 1983.

[54] R.E. Greene and H. Wu. *Function Theory on Manifolds Which Possess a Pole*, volume 699 of *Lecture Notes in Mathematics.* Springer-Verlag, New York, 1979.

[55] S. Hildebrandt. Harmonic mappings of Riemannian manifolds. In E. Giusti, editor, *Harmonic mappings and minimal immersions*, volume 1161 of *Lecture Notes in Mathematics*, pages 1–117, New York, 1985. Springer-Verlag.

[56] E.P. Hsu. Heat semigroup on a complete Riemannian manifold. *The Annals of Probability*, 17:1248–1254, 1989.

[57] E.P. Hsu and W.S. Kendall. Limiting angle of Brownian motion in certain two-dimensional Cartan-Hadamard manifolds. *Annales de la Faculté des Sciences de Toulouse Mathématiques (Série 6)*, 1:169–186, 1992.

[58] E.P. Hsu and P. March. The limiting angle of certain Riemannian Brownian motions. *Communications on Pure and Applied Mathematics*, 38:755–768, 1985.

[59] H. Huang and W.S. Kendall. Correction note to 'Martingales on manifolds and harmonic maps'. *Stochastics*, 37:253–257, 1991.

[60] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes (First Edition).* North-Holland / Kodansha, Amsterdam / Tokyo, 1981.

[61] K. Itô. Stochastic integral. *Proc. Imp. Acad. Tokyo*, 20:519–524, 1944.

[62] K. Itô. Stochastic differentials. *Applied Mathematics and Optimization*, 1:374–381, 1975.

[63] W. Jager and H. Jaul. Uniqueness and stability of harmonic maps and their Jacobi fields. *manuscript mathematica*, 28:269–291, 1979.

[64] E. Jørgensen. Construction of the Brownian motion and the Ornstein-Uhlenbeck process in a Riemannian manifold on basis of the Gangolli-McKean injection scheme. *Zeitschrift für Wahrscheinlichkeitstheorie und verwe Gebiete*, 44:71–87, 1978.

[65] J. Jost. Equilibrium maps between metric spaces. *Calc. Var.*, 2:173–204, 1994.

[66] J. Jost. *Riemannian Geometry and Geometric Analysis.* Springer-Verlag, New York-Heidelberg-Berlin, 1995.

[67] H.P. McKean Jr. *Stochastic Integrals.* Academic Press, New York, 1969.

[68] S. Kakutani. Markoff process and the Dirichlet problem. *Proceedings of the Japanese Academy. Series A. Mathematical Sciences*, 21:227–233, 1945.

[69] H. Kaneko. Liouville theorems based on symmetric diffusions. *Bulletin de la Société Mathématique de France*, 124:545–557, 1996.

[70] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus.* Springer-Verlag, New York, 1988.

[71] H. Karcher. Riemannian centre of mass and mollifier smoothing. *Communications in Pure and Applied Mathematics*, 30:509–541, 1977.

[72] D.G. Kendall. The diffusion of shape. *Advances in Applied Probability*, 9:428–430, 1977.

[73] W.S. Kendall. Brownian motion, negative curvature, and harmonic maps. In D.Williams, editor, *Stochastic Integrals, Proceedings, LMS Durham Symposium, 1980*, volume 851 of *Lecture Notes in Mathematics*, pages 479–491, New York-Heidelberg-Berlin, 1981. Springer-Verlag.

[74] W.S. Kendall. Brownian motion and a generalised little Picard's theorem. *Transactions of the American Mathematical Society*, 275:751–760, 1983.

[75] W.S. Kendall. Brownian motion on 2-dimensional manifolds of negative curvature. *Lecture Notes in Mathematics*, 1059:70–76, 1984.

[76] W.S. Kendall. Nonnegative Ricci curvature and the Brownian coupling property. *Stochastics*, 19:111–129, 1986.

[77] W.S. Kendall. Stochastic differential geometry, a coupling property, and harmonic maps. *The Journal of the London Mathematical Society (Second Series)*, 33:554–566, 1986.

[78] W.S. Kendall. Stochastic differential geometry. In Yu.V. Prohorov and V.V. Sazonov, editors, *Proceedings, First World Congress of the Bernoulli Society*, volume 1, pages 515–524, Utrecht, 1987. VNU Press.

[79] W.S. Kendall. The radial part of Brownian motion on a manifold: Semimartingale properties. *The Annals of Probability*, 15:1491–1500, 1987.

[80] W.S. Kendall. Stochastic differential geometry: An introduction. *Acta Applicandae Mathematicae. An International Journal on Applying Mathematics and Mathematical Applications*, 9:29–60, 1987.

[81] W.S. Kendall. Martingales on manifolds and harmonic maps. In R. Durrett and M. Pinsky, editors, *The Geometry of Random Motion*, volume 73, pages 121–157, Providence, RI, 1988. American Mathematical Society.

[82] W.S. Kendall. Busemann functions and Brownian motion. In J. Norris, editor, *Stochastic Calculus in Applications, Symposium Proceedings, Cambridge 1986*, volume 197 of *Pitman Research Notes in Mathematics*, pages 63–68. Longman, 1988.

[83] W.S. Kendall. Symbolic computation and the diffusion of shapes of triads. *Advances in Applied Probability*, 20:775–797, 1988.

[84] W.S. Kendall. Coupled Brownian motions and partial domain monotonicity for the Neumann heat kernel. *Journal of Functional Analysis*, 86:226–236, 1989.

[85] W.S. Kendall. Probability, convexity, and harmonic maps with small image I: Uniqueness and fine existence. *Proceedings of the London Mathematical Society (Third Series)*, 61:371–406, 1990.

[86] W.S. Kendall. The Euclidean diffusion of shape. In D. Welsh and G. Grimmett, editors, *Disorder in Physical Systems*, pages 203–217, Oxford, 1990. Oxford University Press.

[87] W.S. Kendall. Symbolic Itô calculus: An overview. In N. Bouleau and D. Talay, editors, *Probabilités Numeriques*, pages 186–192, Rocquencourt, 1991. INRIA.

[88] W.S. Kendall. Convexity and the hemisphere. *The Journal of the London Mathematical Society (Second Series)*, 43:567–576, 1991.

[89] W.S. Kendall. Convex geometry and nonconfluent Γ-martingales I: Tightness and strict convexity. In M.T. Barlow and N.H. Bingham, editors, *Stochastic Analysis, Proceedings, LMS Durham Symposium, 11th - 21st July 1990*, pages 163–178, 1991.

[90] W.S. Kendall. The Propeller: A counterexample to a conjectured criterion for the existence of certain convex functions. *The Journal of the London Mathematical Society (Second Series)*, 46:364–374, 1992.

[91] W.S. Kendall. Convex geometry and nonconfluent Γ-martingales II: Well-posedness and Γ-martingale convergence. *Stochastics and Stochastic Reports*, 38:135–147, 1992.

[92] W.S. Kendall. Brownian motion and partial differential equations: From the heat equation to harmonic maps. In *Proc. ISI* 49^{th} *session, Firenze (August 1993)*, volume 3, pages 85–101, 1993.

[93] W.S. Kendall. Doing stochastic calculus with *Mathematica*. In H. Varian, editor, *Economic and Financial Modeling with* Mathematica, pages 214–238. Springer-Verlag, New York, 1993.

[94] W.S. Kendall. Probability, convexity, and harmonic maps II: Smoothness via probabilistic gradient inequalities. *Journal of Functional Analysis*, 126:228–257, 1994.

[95] W.S. Kendall. The radial part of a Γ-martingale and a non-implosion theorem. *The Annals of Probability*, 23:479–500, 1995.

[96] W.S. Kendall. Perfect simulation for the area-interaction point process. In C.C. Heyde and L. Accardi, editors, *Probability Perspective*, Singapore, 1998. World Scientific Press.

[97] W.S. Kendall. A diffusion model for Bookstein triangle shape. *Advances in Applied Probability*, 30(2), 1998.

[98] Yu. Kifer. Brownian motion and harmonic functions on manifolds of negative curvature. *Theory of Probability and its Applications. An English translation of the Soviet journal Teoriya Veroyatnosteĭ i ee Primeneniya*, 21:81–95, 1976.

[99] Yu. Kifer. Spectrum, harmonic functions, and hyperbolic metric spaces. *Israel Journal of Math.*, 89:377–428, 1995.

[100] N.V. Krylov. *Controlled Diffusion Processes.* Springer-Verlag, New York, 1980.

[101] H.L. Le. Exit probability estimates for Γ-martingales on manifolds. *Stochastic and Stochastics Reports*, 54:199–209, 1995.

[102] H.L. Le. Mean size-and-shapes and mean shapes: a geometric point of view. *Advances in Applied Probability*, 27:44–55, 1995.

[103] H.L. Le. Limiting angle of Brownian motion on certain manifolds. *Probability theory and related fields*, 106:137–149, 1996.

[104] H.L. Le and D. Barden. Itô correction terms for the radial parts of semimartingales on manifolds. *Probability Theory and Related Fields*, 101:133–146, 1995.

[105] T.M. Liggett. *Interacting Particle Systems.* Springer-Verlag, New York, 1985.

[106] T. Lindvall. On coupling of Brownian motions. Research report, Department of Mathematics, Chalmers University of Technology & University of Göteborg, Göteborg, 1982.

[107] T. Lindvall. *Lectures on the Coupling Method.* John Wiley & Sons, Chichester and New York, 1992.

[108] T. Lyons and D. Sullivan. Function theory, random paths, and covering spaces. *Journal of Differential Geometry*, 19:299–323, 1984.

[109] P. Malliavin. *Géométrie Differentielle Stochastique*. Séminaire de Mathématiques Supérieures, Université de Montréal, 1978.

[110] P. March. Brownian motion and harmonic functions on rotationally symmetric manifolds. *Annals of Probability*, 14:793–801, 1986.

[111] D. Michel. Comparaison des notions de variétés faiblement et fortement harmoniques. *Comptes Rendus Acad. Sci. Paris*, 282:1007–1010, 1976.

[112] D. Mollison. Spatial contact models for ecological and epidemic spread (with discussion). *J. Roy. Statist. Soc.*, B39:283–326, 1977.

[113] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer-Verlag, New York - Berlin, 1985.

[114] J.M. Oller and J.M. Corcuera. Intrinsic analysis of the statistical estimation. *The Annals of Statistics*, 23:1562–1581, 1995.

[115] E. Pardoux and S.G. Peng. Adapted solutions of a backward stochastic differential equation. *Systems & Control Letters*, 14:55–61, 1990.

[116] J. Picard. Calcul stochastique avec sauts sur une variété. In *Séminaire de Probabilités XXV, Lecture Notes in Mathematics*, volume 1485, pages 196–219. Springer-Verlag, New York-Heidelberg-Berlin, 1991.

[117] J. Picard. Martingales on Riemannian manifolds with prescribed limit. *Journal of Functional Analysis*, 99:223–261, 1991.

[118] J. Picard. Barycentres and martingales on a manifold. *Annales de l'institut Henri Poincare - probabilites et statistiques*, 30:647–702, 1994.

[119] M. Pinsky. Homogenization and stochastic parallel transport. In D. Williams, editor, *Stochastic Integrals*, volume 851 of *Lecture Notes in Mathematics*, pages 271–284, New York, 1981. Springer-Verlag.

[120] J.J. Prat. Étude asymptotique et convergence angulaire du mouvement brownien sur une variété à courbure négative. *C.R. Acad. Sciences Paris Séries A–B,*, 280A:1539–1542, 1975.

[121] G. Price and D. Williams. Rolling with slipping I. *Séminaire de Probabilités*, XVII:194–197, 1983.

[122] J.G. Propp and D.B. Wilson. Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures and Algorithms*, 9:223–252, 1996.

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

[124] L.C.G. Rogers and D. Williams. *Diffusions, Markov Processes, and Martingales volume 2: Itô Calculus.* John Wiley & Sons, Chichester and New York, 1987.

[125] L. Schwartz. *Semi-martingales sur des varietes et martingales conformes sur des varietes analytiques complexes*, volume 780 of *Lecture Notes in Mathematics.* Springer-Verlag, New York-Heidelberg-Berlin, 1980.

[126] L. Schwartz. *Semimartingales and their stochastic calculus on Manifolds.* Les Presses de l'Université de Montréal, Montréal, 1984.

[127] S. Stafford. A probabilistic proof of S.Y. Cheng's theorem. *Annals of Probability*, 18:1816–1822, 1990.

[128] D. Stroock. On the growth of stochastic integrals. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 18:340–344, 1971.

[129] D. Stroock and S.R.S. Varadhan. *Multidimensional diffusion processes.* Springer-Verlag, New York, 1979.

[130] D. Sullivan. The Dirichlet problem at infinity for a negatively curved manifold. *Journal of Differential Geometry*, 18:723–732, 1983.

[131] A. Thalmaier. Brownian motion and the formation of singularities in the heat flow for harmonic maps. *Probability theory and related fields*, 105:335–367, 1996.

[132] A. Thalmaier. On the differentiation of heat semigroups and Poisson integrals. *Stochastics and Stochastics Reports*, 61:297–321, 1997.

[133] A. Thalmaier and F.-Y. Wang. Gradient estimates for harmonic functions on regular domains in Riemannian manifolds. Research report 32, University of Warwick Department of Maths, 1996.

[134] M. van den Berg and J.T. Lewis. Brownian motion on a hypersurface. *Bull. London Math. Soc.*, 17:144–150, 1985.

[135] F.-Y. Wang. Brownian coupling property for negatively curved Riemannian manifolds. Research report, Department of Math., Beijing Normal University, Beijing, 1997.

[136] H. Wu. An elementary method in the study of nonnegative curvature. *Acta Mathematica*, 142:57–78, 1979.

[137] S.-T. Yau. Harmonic functions on complete Riemannian manifolds. *Communications on Pure and Applied Mathematics*, 28:201–228, 1975.

[138] W.A. Zheng. Sur le théorème de convergence des martingales dans une variété Riemannienne. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 63:511–515, 1983.