Brownian motion and the formation of singularities in the heat flow for harmonic maps

Anton Thalmaier

NWF I – Mathematik, Universität Regensburg, D-93040 Regensburg, Germany
(e-mail: anton.thalmaier@mathematik.uni-regensburg.de)

Summary. We develop a general framework for a stochastic interpretation of certain nonlinear PDEs on manifolds. The linear operation of taking expectations is replaced by the concept of "martingale means", namely the notion of deterministic starting points of martingales (with respect to the Levi-Civita connection) ending up at a prescribed state. We formulate a monotonicity condition for the Riemannian quadratic variation of such martingales that allows us to turn smallness of the quadratic variation into a priori gradient bounds for solutions of the nonlinear heat equation. Such estimates lead to simple criteria for blow-ups in the nonlinear heat flow for harmonic maps with small initial energy.

1. Introduction

The theory of harmonic maps connects nonlinear analysis, geometry and topology for Riemannian manifolds in a rather subtle way, see [9], [10], [11]. Harmonic maps $M \to N$ provide a common generalization of the notion of geodesics (for $M = S^1$ or $\mathbb{R}$) and harmonic functions (for $N = \mathbb{R}$ or $\mathbb{R}^n$). It is well-known that probability theory, namely the theory of Brownian motion, is related to the linear case of harmonic functions. Stochastic analysis provides tools to reduce (linear) partial differential equations (e.g., heat equation, Dirichlet problem) to solutions of ordinary *stochastic* differential equations, in such a way that the solution of the partial differential equation is given by taking expectations of a stochastic solution. Although stochastic analysis on Riemannian manifolds is well-developed [13], [14], [15], [20], an essential difficulty in applying stochastic methods to nonlinear geometric PDE problems comes from the fact that taking expectations of random variables is by definition a linear operation, ruling out, for instance, straightforward generalizations to the harmonic map problem for general target manifolds.

There are established ways to define expectations (means) for random variables or probability distributions on manifolds, like the theory of Riemannian centres of mass [23], [30], or the notion of barycentres of measures, see e.g. [16]. Most applications of these concepts are related to random variables concentrated on domains which can be described geometrically in terms of convex geometry (see [25]). It appears difficult to adapt notions relying on convexity to situations where nontrivial topology and the global nature of the manifold is involved. The appropriate replacement of the linear expectation operators is given by a rather sophisticated nonlinear stochastic concept, namely the notion of (deterministic) starting points of martingales with the given random variable as terminal state. Martingales on manifolds depend on a linear connection, e.g., the Levi-Civita connection induced by the metric for Riemannian manifolds. With respect to this connection martingales provide a natural class of free motions relative to the given geometry.

This paper includes a discussion of the heat flow for harmonic maps in terms of martingales on Riemannian manifolds and investigates the development of singularities out of smooth initial data in finite time created by topological reasons. Our main intention is to clarify the probabilistic background of such phenomena and to establish stochastic methods in the field of global geometric evolution problems.

Throughout this paper, we assume that (M, g) is either a compact Riemannian manifold or $(M, g) = (\mathbb{R}^m, \mathrm{can})$ with the canonical Euclidean metric, and that all target manifolds N are compact. All manifolds are connected, and all maps are supposed to be smooth (C^∞) if not stated otherwise explicitly. Solutions of the heat equation are always understood in the classical sense.

2. The heat flow for harmonic maps

Let (M, g) and (N, h) be smooth Riemannian manifolds and let $f \colon M \to N$ be a smooth map. Associated with these data are the two fundamental forms of f.

(a) (*First fundamental form of f*). The pullback of the metric h under f gives a bilinear form $f^*h \in \Gamma(T^*M \otimes T^*M)$ which is defined by

$$(f^*h)_x(u, v) = h_{f(x)}(f_*u, f_*v) \quad \text{for } u, v \in T_xM.$$

(b) (*Second fundamental form of f*). With respect to the Levi-Civita-connections on M and N one has $\nabla df \in \Gamma(T^*M \otimes T^*M \otimes f^*TN)$ defined as covariant derivative of $df = f_* \in \Gamma(T^*M \otimes f^*TN)$.

By taking traces of the fundamental forms (with respect to the given metrics) we get

(i) $\|df\|^2 = \mathrm{trace}\, f^*h \in C^\infty(M)$, the energy density of f, and

(ii) $\tau(f) = \mathrm{trace}\, \nabla df \in \Gamma(f^*TN)$, the tension field of f.

Maps with vanishing tension field $\tau(f)$ are called *harmonic*. Via the Euler-Lagrange equation they appear as stationary (critical) points of the energy functional

$$E(f) = \int_M \|df\|^2 \, d\mathrm{vol}$$

with respect to compactly supported variations of f. As in Hodge theory, where one seeks to realize de Rham cohomology classes by harmonic differential forms, a fundamental question is whether a given homotopy class of maps has a harmonic representative.

The basic existence problem is concerned with the deformation of a map $f\colon M \to N$ into a harmonic map. A classical approach of determining harmonic representatives in a given homotopy class is the so called "deformation under the heat flow". Here one uses the heat equation

$$\begin{cases} \frac{\partial}{\partial t} u = \frac{1}{2}\tau(u) & \text{on } [0,\infty[\,\times M \\ u|_{t=0} = f \end{cases} \tag{2.1}$$

to establish a homotopy $u_t = u(t, \cdot)$ between the initial map $f = u_0$ and a harmonic map, namely $u_\infty \equiv \lim_{t\to\infty} u_t$. Inherent to this method are several difficulties: one of them is that a priori only existence of local solutions to the heat equation is guaranteed, e.g., see [18], [22].

Theorem 2.1 (Short-term existence of solutions) *Let M and N be compact. Then there exists $T > 0$ depending on the initial map f such that the heat equation has a unique smooth solution $(t,x) \mapsto u(t,x)$ for $(t,x) \in [0,T[\,\times M$.*

Short-term existence of solutions also holds for not necessarily compact manifolds M, provided the energy density $\|df\|^2$ of the initial map is bounded on M; then also $\|du(t,\cdot)\|^2$ is bounded on M for any $t < T$. In their pioneering paper on the subject Eells-Sampson obtain the following global existence result, using strong curvature assumptions for N.

Theorem 2.2 [12] *Let M and N be compact, and suppose the sectional curvature Riem^N of N is non-positive. Then for any $f \in C^\infty(M,N)$ the heat equation admits a unique, global, smooth solution $u\colon [0,\infty[\,\times M \to N$. As $t \to \infty$, the maps $u(t,\cdot)$ converge smoothly to a harmonic map $u_\infty \in C^\infty(M,N)$ homotopic to f.*

If the manifold M has a boundary $\partial M \neq \varnothing$, then it is natural to consider the Dirichlet problem, namely whether or not a given (smooth) $\phi\colon \partial M \to N$ has an extension to a harmonic map $u\colon M \to N$ with $u|_{\partial M} = \phi$. In the heat flow approach to this problem one works with the heat equation for a suitable extension f of ϕ to M, together with the Dirichlet boundary condition

$$u(t, \cdot)|_{\partial M} = \phi.$$

The result of Eells-Sampson was extended to cover the case of manifolds with boundaries $\partial M \neq \varnothing$ by Hamilton [18]. It is well-known that the curvature restriction $\mathrm{Riem}^N \leq 0$ can be weakened if initial and boundary data have small range [21].

The situation turns out to be much more complicated if the curvature assumption in Theorem 2.2 is dropped. In higher dimensions hardly any general global result is known for the homotopy problem unless $\mathrm{Riem}^N \leq 0$. There can be topological restrictions which prevent the heat flow from having any chance of converging or subconverging. For example, as is well-known [9], there exist no harmonic maps $T^2 \to S^2$ with degree ± 1. This implies that solutions of the heat equation for any initial data of degree ± 1 cannot converge or subconverge to a harmonic map. A second even more fundamental problem is the question whether the heat flow exists for all $t \geq 0$ without curvature assumptions on N.

Local existence of solutions implies that there is a maximal interval $[0,T[$ where $0 < T \leq \infty$ such that the solution exists on this interval but cannot be extended beyond. Whether the case $T < \infty$, namely "blow-up in finite time", is possible to occur has been a challenging problem for a long time (see [10], p. 63).

It is meanwhile known [4], [3] that blow-up in finite time is a natural phenomenon if the initial map $u_0 = f$ belongs to a nontrivial homotopy class and the initial energy $E(u_0)$ is sufficiently small.

Theorem 2.3 [4] *Let M and N be compact,* $\dim M \geq 3$, *and let $\mathscr{H}$ be any nontrivial homotopy class in $C(M, N)$ such that*

$$E_{\mathscr{H}} := \inf\bigl\{E(f) : f \in \mathscr{H} \cap C^\infty(M, N)\bigr\} = 0.$$

Then there exists $\varepsilon > 0$ such that if $f \in \mathscr{H} \cap C^\infty(M, N)$ and $E(f) < \varepsilon$ then the solution of the heat equation with initial condition f blows up in finite time T. Moreover $T = T(f) \to 0$ as $E(f) \to 0$.

A result of White [44] guarantees $E_{\mathscr{H}} = 0$ for many nontrivial homotopy classes $\mathscr{H}$ of maps $M \to N$. Namely, let M and N be compact Riemannian manifolds and $u_0 \in C^\infty(M, N)$, then

$$\inf\bigl\{E(f) : f \in C^\infty(M, N), \ f \simeq u_0\bigr\} = 0$$

if and only if u_0 is 2-homotopic to a constant map (i.e. $\pi_i(u_0) = 0$ for $i \leq 2$). This is equivalent to the condition that the restriction of u_0 to the 2-skeleton of some triangulation of M is homotopic to a constant map. Specifically, any map $u_0 \in C^\infty(M, N)$ is homotopic to maps with arbitrary small energy, for instance, if

(a) $\pi_1(M) = 0$ and $\pi_2(M) = 0$, or

(b) $\pi_1(N) = 0$ and $\pi_2(N) = 0$, or

(c) $\pi_1(M) = 0$ and $\pi_2(N) = 0$.

Note that if $\mathrm{Riem}^N \leq 0$ then $\pi_i(N) = 0$ for $i \geq 2$, and any map $M \to N$ that is 2-homotopic to a constant is already homotopic to a constant.

Heat flow is energy decreasing and, if the energy is sufficiently small, the deformation goes towards constant maps; blow-ups occur if this is impossible for topological reasons. The way analysis and topology combine in the heat flow to create singularities in finite time is far from being completely understood. In this paper we like to demonstrate that stochastic analysis provides natural tools to deal with such questions.

3. Brownian motion and harmonic maps

Let $\bigl(\Omega, \mathscr{F}, P; (\mathscr{F}_t)_{t \in \mathbb{R}_+}\bigr)$ be a given filtered probability space fulfilling the usual conditions. An adapted stochastic process X with continuous paths on a Riemannian manifold (M, g) is a *Brownian motion* if for each $\varphi \in C^\infty(M)$ with compact support

$$\varphi \circ X_t - \varphi \circ X_0 - \frac{1}{2}\int_0^t (\Delta_M \varphi) \circ X_s \, ds \quad (t \geq 0) \tag{3.1}$$

defines a real-valued martingale; here Δ_M is the Laplace-Beltrami operator on (M, g). We say X is $\mathrm{BM}(M, g)$, and write $X = X^x$ if $X_0^x = x$ a.s.

There is an intrinsic method of constructing Brownian motion on a Riemannian manifold (M, g) of dimension m. Let W be a flat $\mathrm{BM}(\mathbb{R}^m)$. Denote by $L_1, \ldots, L_m$ the canonical horizontal vector fields on the orthonormal base bundle $\mathrm{O}(M)$ over M, given by $L_i(u) = h_u(ue_i)$ where $h\colon \pi^* TM \to T\mathrm{O}(M)$ is the horizontal lift induced

by the Levi-Civita connection on M. Orthonormal frames u are read as isometries $u: \mathbb{R}^m \to T_{\pi(u)}M$, and e_i stands for the i-th standard basis vector in $\mathbb{R}^m$. Let u_0 be a $\mathscr{F}_0$-measurable random variable with values in $O(M)$. Solving the $O(M)$-valued Stratonovich differential equation (SDE)

$$dU = \sum_{i=1}^{m} L_i(U) * dW^i, \quad U_0 = u_0 , \qquad (3.2)$$

gives horizontal Brownian motion on $O(M)$, and the projection $X = \pi \circ U$ of the solution process U down to M defines a $\mathrm{BM}(M, g)$, started at $x_0 = \pi \circ u_0$; see [14].

This geometric procedure of solving SDEs on the orthonormal base bundle $O(M)$ can be extended to construct larger classes of random motions on M if we replace the Brownian motion W in (3.2) by a continuous semimartingale Z on $\mathbb{R}^m$ (started at 0). In this case, $X = \pi \circ U$ on M is called *stochastic development* of Z [15]. Thus, induced by the Levi-Civita connection on M, one can associate to each continuous semimartingale Z on $\mathbb{R}^m$ a stochastic development X on M, together with a horizontal lift U of X on $O(M)$, and therefore a notion of parallel transport along the paths of X via

$$//_{0,t} = U_t \circ U_0^{-1}: T_{X_0}M \xrightarrow{\sim} T_{X_t}M. \qquad (3.3)$$

Stochastic development can be reversed in order to find the "anti-development" Z (with values in $\mathbb{R}^m$) from the process X on M. More precisely, given a semimartingale X on M, we may construct U as horizontal lift of X on $O(M)$ (starting at some initial frame u_0 above x_0) and get the corresponding anti-development Z as a Stratonovich integral of the canonical connection 1-form ϑ on $O(M)$ along the process U (see [14], [15], or [19]), namely as

$$Z = \int \vartheta \, (*dU) \equiv \int_U \vartheta \qquad (3.4)$$

where $\vartheta \in \Gamma(T^*O(M) \otimes \mathbb{R}^m)$ is the canonical 1-form of the connection, defined by $\vartheta_u(e) = u^{-1}(\pi_* e)$ for $e \in T_u O(M)$. Thus, given one of the three processes Z, X, or U, up to a specification of the starting variables, the two others may be constructed. Note that the Riemannian quadratic variation $[X, X] = \int g(dX, dX)$ of an M-valued semimartingale X depends only on the martingale part of its anti-development, namely $[X, X] = [Z, Z]$ where $[Z, Z] = [Z^1, Z^1] + \ldots + [Z^m, Z^m]$ is the usual (Euclidean) quadratic variation of Z.

An adapted process X with values in the Riemannian manifold (M, g) is called a ∇-*martingale* (with respect to the Levi-Civita connection ∇) if X is the stochastic development of a continuous $\mathbb{R}^m$-valued local martingale Z (see [15]). We only consider ∇-martingales with respect to Levi-Civita connections, thus we omit the specification of the connection in the sequel. The class of all martingales on (M, g) is denoted by $\mathrm{Mart}(M, g)$. Using the functional characterization of manifold-valued martingales, due to Darling [6], which relies on the richness of germs of convex functions, it is easy to see that the martingale property is a local property; hence there is a straightforward meaning of M-valued semimartingales being a martingale on stochastic intervals of the form $[\sigma, \tau[$ or $]\sigma, \tau[$.

$\mathrm{BM}(M, g)$ constructed by stochastic development lives on the canonical Wiener space that carries $\mathrm{BM}(\mathbb{R}^m)$. The standard filtration on the m-dimensional Wiener space will be referred to as (m-dimensional) *Brownian filtration* in the sequel. All manifold-valued processes will be defined on some fixed m-dimensional Wiener space with its Brownian filtration, but with m not necessarily the dimension of the manifold. For

instance, the image $f \circ X$ of an M-valued Brownian motion X under a differentiable map $f \colon M \to N$ is a semimartingale on N adapted to the m-dimensional Brownian filtration with $m = \dim M$.

A consequence of the adaptedness to a Brownian filtration is that M-valued semi-martingales start at deterministic points. Hence, in the correspondence between the continuous $\mathbb{R}^m$-valued semimartingale Z, its M-valued stochastic development X, and the $\mathrm{O}(M)$-valued horizontal lift U of X, there is only a choice for the initial frame $U_0 = u$ above $X_0 = x$. Such a choice of u means selecting an isometry $u \colon \mathbb{R}^m \xrightarrow{\sim} T_{\pi(u)}M$ via $z \mapsto \sum_{i=1}^m z^i u_i$. Under the identification of $\mathbb{R}^m$ and $T_x M$ via $u \in \mathrm{O}(M)_x$ we may consider $\bar{Z} \equiv uZ$ with values in $T_x M$ rather than Z with values in $\mathbb{R}^m$. The horizontal lift U then still depends on u, but not the induced parallel transport $/\!/_{0,t} = U_t \circ U_0^{-1} = U_t \circ u^{-1}$ along X.

Maps between manifolds may be studied under the aspect how they transform certain classes of random motions, for instance Brownian motions, see [6]. The action of a map f on $\mathrm{BM}(M,g)$ is described in terms of the first and second fundamental form of f.

Theorem 3.1 *Let* $f \colon (M,g) \to (N,h)$ *be a smooth map between Riemannian manifolds. Let* X *be a semimartingale on* M *with* $X_0 = x$, *which comes by stochastic development from a flat semimartingale* Z *on* $T_x M$. *The image process* $\tilde{X} = f \circ X$ *is a semimartingale on* N *with* $\tilde{X}_0 = f(x)$, *and hence determined by its anti-development* $\tilde{Z}$ *in* $T_{f(x)}N$. *If* $/\!/_{t,0} = U_0 \circ U_t^{-1}$ *and* $/\!\!\widetilde{/}_{t,0} = \tilde{U}_0 \circ \tilde{U}_t^{-1}$ *denote the parallel transports along paths of* X, *resp.* $\tilde{X}$, *we have*

$$d\tilde{Z} = /\!\!\widetilde{/}_{t,0}\, df\, /\!/_{0,t}\, dZ + \tfrac{1}{2}\, /\!\!\widetilde{/}_{t,0}\, \nabla df(dX, dX)\,. \tag{3.5}$$

Specifically, if X *is a* $\mathrm{BM}(M,g)$, *and correspondingly* Z *a* $\mathrm{BM}(T_x M)$, *we get*

$$d\tilde{Z} = /\!\!\widetilde{/}_{t,0}\, df\, /\!/_{0,t}\, dZ + \tfrac{1}{2} /\!\!\widetilde{/}_{t,0}\, \tau(f) \circ X\, dt\,; \tag{3.6}$$

in this case the Riemannian quadratic variation $[\tilde{X}, \tilde{X}] = \int h(d\tilde{X}, d\tilde{X})$ *reads as*

$$d[\tilde{X}, \tilde{X}] = \sum h(f_* U e_i, f_* U e_i)\, dt = \big(\|df\|^2 \circ X\big)\, dt\,. \tag{3.7}$$

The proof of Theorem 3.1 is based on the geometric Itô formula; see [15], or [19] pp. 442-4, for details. Equations (3.6) and (3.7) show that the martingale part in the anti-development of $f \circ X$ measures the energy of f along the path of X, while the drift part captures the tension field of f along X. Specifically, a smooth map $f \colon M \to N$ is harmonic if and only if it maps $\mathrm{BM}(M,g)$ to $\mathrm{Mart}(N,h)$.

4. Expectations on manifolds

In this section we comment on various approaches to an intrinsic definition of expectations (means) of random variables on Riemannian manifolds; see [30], [23], [16], [25], [36], [1] and [2].

Definition 4.1 *Let* (N,h) *be a Riemannian manifold and* ξ *an* N-*valued random variable defined on a probability space* $(\Omega, \mathscr{F}, P)$; *further fix* $y_0 \in N$.

(i) (Barycentres) The point y_0 is said to be in the *barycentre* of ξ if $\varphi(y_0) \leq \mathbb{E}[\varphi \circ \xi]$ for all bounded convex functions $\varphi \colon U \to \mathbb{R}$ such that $U \subset N$ is open, convex, $y_0 \in U$, and $\mathrm{range}(\xi) \subset U$, a.e.

(ii) (Riemannian centres of mass; "Cartan means") Suppose that ξ is L^2-integrable in the sense that $\mathbb{E}\big[\mathrm{dist}_N(x_0, \xi)^2\big] < \infty$ for some $x_0 \in N$. The point y_0 is said to be a *Cartan mean* of ξ if y_0 is a local minimum of the function $Q \colon N \to \mathbb{R}_+$ given by $Q(x) = \mathbb{E}\big[\mathrm{dist}_N(x, \xi)^2\big]$.

(iii) (Martingale means) Let $(\mathscr{F}_t)_{t \geq 0}$ be a filtration on $(\Omega, \mathscr{F}, P)$ fulfilling the usual conditions such that $\mathscr{F} = \mathscr{F}_\infty$. Then the point y_0 is said to be a *martingale mean* of ξ if there is a uniformly integrable martingale Y on N, adapted to $(\mathscr{F}_t)_{t \geq 0}$, starting at y_0 and ending up at ξ, i.e., $Y_0 = y_0$ and $Y_\infty = \xi$, a.s.

Each of the above concepts relies on a different aspect of the geometric setting, e.g., the notion of convexity, the distance function induced by the Riemannian metric, or the idea of a drift-free random motion. Only the martingale mean requires a filtration; it uses the starting point y_0 of a martingale Y ending up at $Y_\infty = \xi$ as a substitute for the missing expectation "$\mathbb{E}[\xi]$". (In addition, one may take Y_s as replacement of the conditional expectation "$\mathbb{E}^{\mathscr{F}_s}[\xi]$"). Note that if $\mathscr{F}_0$ is trivial, e.g., in case of the Brownian filtration on Wiener space, adapted processes start at deterministic points. By definition, the martingale interpolating between y_0 and ξ is assumed to be uniformly integrable in the sense that its anti-development is an uniformly integrable martingale on $\mathbb{R}^n$. Martingales on manifolds include continuous local martingales on Euclidean space; thus without the restriction to uniformly integrable martingales this notion would be too wide.

Example 4.2 Let $\big(\Omega, \mathscr{F}, P; (\mathscr{F}_t)_{t \in \mathbb{R}_+}\big)$ be a filtered probability space such that $\mathscr{F}_0$ is trivial and $\mathscr{F}_\infty = \mathscr{F}$. Take $N = \mathbb{R}^n$, and let $\xi \in L^1(\Omega, \mathscr{F}, P; \mathbb{R}^n)$. Then $Y_s = \mathbb{E}^{\mathscr{F}_s}[\xi]$, $s \geq 0$, defines a uniformly integrable martingale on $\mathbb{R}^n$ with limit $Y_\infty = \xi$, a.s., and starting point $Y_0 = \mathbb{E}^{\mathscr{F}_0}[\xi] = \mathbb{E}[\xi]$.

In Example 4.2 the martingale itself, but not its starting point, depends on the filtration — a situation that changes if we consider random variables ξ with values in general Riemannian manifolds (N, h). Note that in the example $y_0 = \mathbb{E}[\xi]$ is also a Cartan mean and obviously in the barycentre of ξ by Jensen's inequality.

In the context of the martingale mean it appears natural to keep track of the "size" of the interpolating martingales. As mentioned above, the Riemannian quadratic variation of an N-valued martingale coincides with the quadratic variation of its anti-development in $\mathbb{R}^n$. On the other hand, the size of a martingale on Euclidean space is measured by its quadratic variation process. Note that martingales Y with prescribed end state Y_∞ automatically fulfill $[Y, Y]_\infty < \infty$ a.e., as a consequence of the martingale theorem.

For $p \geq 1$, a martingale Y with values in a Riemannian manifold (N, h) is called an H^p-*martingale* if $\mathbb{E}\big([Y, Y]_\infty^{p/2}\big) < \infty$ where $[Y, Y] = \int h(dY, dY)$ is the Riemannian quadratic variation of Y (see also [8]). A martingale mean y_0 of an N-valued random variable ξ is called an H^p-*martingale mean* if there is an H^p-martingale begun at y_0 and converging to ξ. Then, the H^p-norm

$$\|Y\|_{H^p} = \big\|[Y, Y]_\infty^{1/2}\big\|_{L^p}$$

serves as a specification of the goodness of the mean value $Y_0 = y_0$ for the variable ξ.

Giving a brief comparison of the concepts in Definition 4.3, we remark that (at least on sufficiently small domains) both Cartan means and martingale means are compatible with barycentres (see [25]). For instance, let y_0 be a martingale mean for a random variable ξ where Y is the corresponding uniformly integrable martingale with $Y_0 = y_0$ and $Y_\infty = \xi$. Then, for all $\varphi \in C^\infty(N)$, say such that $\|d\varphi\|$ is bounded, we have:

$$\varphi \circ Y_s = \mathbb{E}^{\mathscr{F}_s}[\varphi \circ \xi] - \tfrac{1}{2}\,\mathbb{E}^{\mathscr{F}_s}\left[\int_s^\infty \nabla d\varphi(dY, dY)\right].$$

As a consequence, $\varphi(y_0)$ and $\mathbb{E}[\varphi \circ \xi]$ differ by the nonlinear correction term

$$\tfrac{1}{2}\mathbb{E}\left[\int_0^\infty \nabla d\varphi(dY, dY)\right]$$

depending on the complete martingale Y. Obviously, the correction is 0 if $\nabla d\varphi = 0$ (i.e., φ affine), and $\varphi \circ y_0 \leq \mathbb{E}[\varphi \circ \xi]$ if $\nabla d\varphi \geq 0$ (i.e., φ convex). The obvious problem with barycentres comes from the fact that in general there are not enough convex functions to specify appropriate means for large range random variables. For example, on a compact Riemannian manifold (without boundary) there are no globally defined non-constant convex functions at all.

We like to stress that there is fundamental conceptual difference between Cartan means and martingale means which makes these notions incompatible in any general setting. For instance, fix an $\mathscr{F}_\infty$-measurable random variable ξ with values in N. For simplicity, suppose that $\mathrm{range}(\xi) \subset B \subset N$ where B is a sufficiently small regular geodesic ball such that $\mathrm{dist}_N = \mathrm{dist}_B$ on $B \times B$; here dist_B is the restricted metric on B defined by taking into account only curves connecting x and y within B. If $y_0 \in B$ is a Cartan mean of ξ, then in particular, y_0 is a critical point for

$$Q \colon N \to \mathbb{R}, \quad Q(x) = \mathbb{E}\left[\mathrm{dist}_N(x, \xi)^2\right],$$

or, in other words, $\mathbb{E}\left[\exp_{y_0}^{-1}(\xi)\right] = 0$. Thus, for a Cartan mean y_0 of ξ, there is a martingale $\hat{Y}$ in $T_{y_0}N$, namely $\hat{Y}_s = \mathbb{E}^{\mathscr{F}_s}\left[\exp_{y_0}^{-1}(\xi)\right]$, such that the N-valued process $Y = \exp_{y_0} \hat{Y}$ starts at y_0 and terminates at ξ, i.e., $Y_\infty = \xi$. In general, Y will not be a martingale unless $\exp_{y_0}$ is totally geodesic. On the other hand, if y_0 is a martingale mean of ξ, then there is a martingale Y on N such that $Y_0 = y_0$ and $Y_\infty = \xi$ a.e. By definition, Y is the stochastic development of a martingale $\hat{Y}$ in $T_{y_0}N \cong \mathbb{R}^m$. In both cases there is a martingale $\hat{Y}$ in $T_{y_0}N$, but dealing with the Cartan mean implies that $\hat{Y}$ is transported onto N via the exponential map (at the fixed point y_0) to give a process Y on N connecting y_0 and ξ (in general only a semimartingale), whereas the concept of the martingale mean uses the more complicated procedure of stochastic development to transport $\hat{Y}$ onto N: the frame in $T_{y_0}N$ (used to identify $T_{y_0}N$ and $\mathbb{R}^m$) is carried along Y by stochastic parallel transport.

Throughout this paper, we restrict ourselves to H^2-martingale means which is an appropriate class for the applications we have in mind. The term "martingale mean" will be used in this stronger sense. As pointed out, for H^2-martingales Y,

$$\mathbb{E}([Y, Y]_\infty) = \mathbb{E}\left[\int_0^\infty h(dY, dY)\right]$$

measures the "size" of Y and specifies a "distance" of the mean $Y_0 = y_0$ to the variable ξ. In case uniqueness for Y fails, the (squared) H^2-norm $\mathbb{E}([Y, Y]_\infty)$ allows one to compare different mean values and to specify a hierarchy of means to the random variable ξ.

In this paper the N-valued random variables ξ will be of the form $\xi = f \circ X_t^x$ for some t, or slightly more general, $\xi = f \circ X_\tau^x$ for some stopping time τ; here X^x is a $\mathrm{BM}(M, g)$ and $f \colon M \to N$ a smooth map. The relevant filtration is the m-dimensional Brownian filtration on Wiener space where $m = \dim M$. Since the class of martingales is invariant under transformations of the time scale, martingale means for random variables of the above type are covered by prescribing terminal values as $t \to \infty$. The H^p-norm of martingales is unaffected by a time change.

Example 4.4 For $n > 2$, let $(S^{n-1}, d\vartheta^2)$ be the $(n-1)$-dimensional standard sphere. The map

$$f \colon \mathbb{R}^n \setminus \{0\} \to S^{n-1}, \quad x \mapsto \tfrac{x}{\|x\|},$$

is harmonic. Hence, for a $\mathrm{BM}(\mathbb{R}^n)$ $X = X^x$ started at $x \neq 0$, the angle process Θ^x:

$$\Theta_t^x = f \circ X_t^x, \quad t \geq 0,$$

defines a martingale on S^{n-1}. Its Riemannian quadratic variation is given by

$$d[\Theta^x, \Theta^x] = (n-1)(R^r)^{-2}\, dt, \quad r = \|x\|,$$

where $R^r = \|X^x\|$ is a Bessel process of dimension n started at $r > 0$. Thus, for $x \neq 0$ and any $t > 0$,

$$\mathbb{E}([\Theta^x, \Theta^x]_t) = \mathbb{E}\big(\textstyle\int_0^t d[\Theta^x, \Theta^x]\big) < \infty,$$

and the point $f(x) \in S^{n-1}$ is a $(H^2\text{-})$martingale mean for the random variable $\xi = \Theta_t^x$. But $\mathbb{E}([\Theta^x, \Theta^x]_t) \to \infty$, as $x \to 0$; hence, for x close to 0, $x \neq 0$, the mean $f(x)$ requires a martingale with large H^2-norm, if described by $(\Theta_s^x)_{0 \leq s \leq t}$. Small variations of Θ_t^x, caused by varying x near 0, give rise to big changes in $f(x)$.

Note that, for $x = 0$ in Example 4.4, the angle process $(\Theta_s^0)_{s > 0}$ provides a martingale on $(S^{n-1}, d\vartheta^2)$ without a starting point. We get $\mathbb{E}\big(\int_\varepsilon^t d[\Theta^0, \Theta^0]\big) < \infty$ for any $0 < \varepsilon < t$, but $\mathbb{E}\big(\int_0^t d[\Theta^0, \Theta^0]\big) = \infty$ for $t > 0$. Thus Θ^0 may be considered as martingale begun at $t = -\infty$ in its "intrinsic time".

5. The stochastic representation of the heat equation

We start discussing some nonlinear PDE problems on manifolds under the aspect of taking expectations on manifolds. Consider the following three types of problems.

(i) (*Nonlinear heat equation*). Let M and N be (compact) Riemannian manifolds and let $f \colon M \to N$ be smooth. The map f is deformed using the heat equation:

$$\tfrac{\partial}{\partial t} u = \tfrac{1}{2}\tau(u), \quad u\big|_{t=0} = f. \tag{5.1}$$

(ii) (*Nonlinear heat equation with a boundary condition*). M is allowed to have a boundary $\partial M \neq \emptyset$ and, as additional data, a smooth map $\phi \colon \partial M \to N$ is given. The problem is solving the heat equation (5.1) as in (i), but together with the Dirichlet boundary condition $u(t, \cdot)\big|_{\partial M} = \phi$. (Note: $f\big|_{\partial M} = \phi$).

(iii) (*Dirichlet problem*). M and N are again (compact) Riemannian manifolds; M with boundary $\partial M \neq \emptyset$, together with a smooth map $\phi \colon \partial M \to N$. The problem is to find $u \colon M \to N$ harmonic with $u\big|_{\partial M} = \phi$.

Note that in the flat case $N = \mathbb{R}^n$ solutions are given in terms of Brownian motion by taking expectations, namely as

(i) $u(t, x) = \mathbb{E}[f \circ X_t^x]$ for the heat equation,

(ii) $u(t, x) = \mathbb{E}[f \circ X_{t \wedge \tau^x}^x]$ for the heat equation with a boundary condition,

(iii) $u(x) = \mathbb{E}[\phi \circ X_{\tau^x}^x]$ for the Dirichlet problem,

where X^x is $\mathrm{BM}(M, g)$ started at x, and $\tau^x = \inf\{s \geq 0 : X_s^x \in \partial M\}$ denotes the first hitting time of ∂M. These formulas are meaningless in the general situation; nevertheless there should be a stochastic interpretation involving appropriate expectations (means) of the N-valued random variables $f \circ X_t^x$, $f \circ X_{t \wedge \tau^x}^x$, or $f \circ X_{\tau^x}^x$, respectively. The key observation is given by the following theorem.

Theorem 5.1 *For M compact, let $u \colon [0, T[\times M \to N$ be a smooth solution of the heat equation (5.1) and $(t, x) \in [0, T[\times M$. Then the N-valued process*

$$Y_s = u(t - s, X_s^x), \quad 0 \leq s \leq t, \tag{5.2}$$

is an H^2-martingale on (N, h) with $Y_0 = u(t, x)$ and $Y_t = f \circ X_t^x$. Hence, $u(t, x)$ is an H^2-martingale mean of the N-valued random variable $\xi = f \circ X_t^x$.

Proof. We consider the N-valued semimartingale Y as image of the $\mathbb{R}_+ \times M$-valued semimartingale $\hat{X}_s = (s, X_s^x)$, $0 \leq s \leq t$, under the map $(s, y) \mapsto u(t - s, y)$ and apply Theorem 3.1 to determine its anti-development Z in $T_{f(x)} N$. The fact that u solves the heat equation is equivalent to a vanishing drift component of Z; namely from formula (3.5), resp. (3.6), we get (modulo differentials of local martingales)

$$dZ \overset{\mathrm{m}}{=} //_{s,0} \left(-\partial_t u + \tfrac{1}{2}\tau(u)\right)(t - s, X_s^x)\, ds = 0 \tag{5.3}$$

($//_{s,0}$ denotes parallel transport along Y). Hence, Y is an N-valued martingale, and obviously, $Y_0 = u(t, x)$, resp., $Y_t = u(0, X_t^x) = f \circ X_t^x$. The Riemannian quadratic variation of Y is given as

$$d[Y, Y] \equiv h(dY, dY) = \|du(t - s, \cdot)\|^2(X_s^x)\, ds. \tag{5.4}$$

From (5.4) we conclude immediately $\mathbb{E}([Y, Y]_t) = \mathbb{E}\left(\int_0^t h(dY, dY)\right) < \infty$, verifying that Y is an H^2-martingale. $\square$

The conclusion of Theorem 5.1 also holds for not necessarily compact manifolds M if $\|du\|^2$ is bounded on $[0, t] \times M$. Moreover, Theorem 5.1 can immediately be adapted to the case $\partial M \neq \emptyset$. For instance, if $u \colon [0, T[\times M \to N$ is a solution of the heat equation with a boundary condition, then, for $(t, x) \in [0, T[\times M$,

$$Y_s = u(t - s, X_{s \wedge \tau^x}^x), \quad 0 \leq s \leq t,$$

is an H^2-martingale on (N, h) with $Y_0 = u(t, x)$ and $Y_t = f \circ X_{t \wedge \tau^x}^x$. Likewise, if $u \colon M \to N$ solves the nonlinear Dirichlet problem,

$$Y_s = u(X_{s \wedge \tau^x}^x), \quad s \geq 0,$$

is an H^2-martingale, started at $u(x)$ with limit $u \circ X_{\tau^x}^x$ as $s \to \infty$.

6. Some non-uniqueness results

We now briefly describe some examples indicating that martingale means on manifolds are quite a delicate object, and which illustrate some of the issues arising in a global theory of expectations on manifolds. Here we mainly stick to the lack of uniqueness in the nonlinear Dirichlet problem.

Let $M = B_1(0) \subset \mathbb{R}^m \equiv \{x \in \mathbb{R}^{m+1} : x_{m+1} = 0\}$, the closed unit ball in $\mathbb{R}^m$, and $N = S^m$, the standard m-dimensional sphere. We think of S^m as the unit sphere in $\mathbb{R}^{m+1}$ and view M as unit ball in the equatorial hyperplane $\mathbb{R}^{m+1} \cap \{x_{m+1} = 0\}$. Thus $\partial M = S^{m-1} \overset{\iota}{\hookrightarrow} S^m$ as equator with the inclusion $\iota(x) = x$.

Let X be $\mathrm{BM}(\mathbb{R}^m)$, started at $X_0 = 0$, and τ the first hitting time of $\partial B_1(0)$. The underlying filtration is the standard m-dimensional Brownian filtration. We take the S^m-valued random variable $\xi = X_\tau$ (actually with values in the equator S^{m-1} of S^m), and look for martingales Y on S^m with terminal value ξ, that is $Y_\tau = \xi$ a.s. Recall that this question is related to the Dirichlet problem of finding harmonic maps $u\colon B_1(0) \to S^m$ such that u equals the identity map on $\partial B_1(0) = S^{m-1} \subset S^m$: for each such u the composition $Y = u \circ X$ is a martingale with the desired property.

First, we look for martingales starting from the south (or north) pole and ending up at ξ. For this, let (ϑ, r) be polar coordinates in $B_1(0)$, and write $X = (\Theta, R)$ such that Θ gives the angle of X on S^{m-1} and R the Euclidean distance of X to the origin. We restrict ourselves to "rotationally invariant" martingales Y on S^m of the form

$$Y = (\Theta \cos h \circ R, \sin h \circ R) \tag{6.1}$$

for some suitable C^2-function $h\colon [0,1] \to \mathbb{R}$ such that in addition $Y_0 =$ south pole, i.e., $h(0) = -\pi/2$, and $Y_\tau = \xi$, i.e., $h(1) = 0$. The condition on h to make Y a martingale reads as

$$\frac{d^2 h}{dr^2} + \frac{m-1}{r}\frac{dh}{dr} + \frac{m-1}{2r^2}\sin 2h = 0. \tag{6.2}$$

It is an elementary substitution to transform (6.2) to the equation for a damped pendulum which can be analyzed by standard methods (see also [40], pp. 92–93).

It follows that for $m \geq 7$ any solution h of (6.2) with $h(0) = -\pi/2$, $h'(0) > 0$ lies below the line $h = 0$, and is increasing and asymptotic to $h = 0$. This implies that there is no martingale of the type (6.1) with $Y_0 =$ south pole and $Y_\tau = \xi$.

For $3 \leq m \leq 6$, however, there are infinitely many different martingales starting from the south pole and ending up at ξ. Any solution h of (6.2) with $h(0) = -\pi/2$, $h'(0) > 0$ is now asymptotic to the line $h = 0$ and crosses this line at infinitely many points $0 < r_1 < r_2 < r_3 < \ldots$ with $r_i \to \infty$; hence, as a consequence of the scale invariance of (6.2), for each $i \in \mathbb{N}$

$$h_i(r) = h(r_i\, r), \quad 0 \leq r \leq 1,$$

gives a solution of (6.2) on $[0,1]$ satisfying the boundary conditions $h_i(0) = -\pi/2$ and $h_i(1) = 0$. The corresponding martingales on S^n

$$Y^i = (\Theta \cos h_i \circ R, \sin h_i \circ R)$$

possess Riemannian quadratic variation given by

$$d[Y^i, Y^i] = \left(\tfrac{d}{dr} h_i\right)^2 (R)\, dt + (m-1) R^{-2} \cos^2 h_i(R)\, dt.$$

Exploiting the fact that, by the scaling property of BM, for any $r > 0$ the processes $\big((1/r)\,R_{r^2t} : t \geq 0\big)$ and $(R_t : t \geq 0)$ are equivalent, we get

$$\mathbb{E}\big([Y^i, Y^i]_\tau\big) < \mathbb{E}\big([Y^{i+1}, Y^{i+1}]_\tau\big) \quad \text{for } i = 1, 2 \dots$$

There are several features which should be stressed here. First, only the martingale Y^1 of least H^2-norm

$$\|Y^1\|_{H^2} = \big(\mathbb{E}\,[Y^1, Y^1]_\tau\big)^{1/2}$$

lives completely in the lower hemisphere. A remarkable feature is the failure of estimates on the size of the martingale in terms of the boundary data, as can be seen from

$$\|Y^i\|_{H^2}^2 = \mathbb{E}\big([Y^i, Y^i]_\tau\big) \to \infty \quad \text{as } i \to \infty\,.$$

Moreover, as $i \to \infty$, note that $Y^i_{t \wedge \tau} \to Y^\infty_{t \wedge \tau}$ a.e. for each $t > 0$ where $Y^\infty = u_\infty(X)$ and $u_\infty(x) = \frac{x}{\|x\|}$ for $x \neq 0$ in $B_1(0)$. As explained in Example 4.4, the martingale $Y^\infty \equiv (\Theta, 0)$, defined on the interval $]0, \tau]$, is a martingale without a starting point. It corresponds to $h \equiv 0$ in (6.1) and has the property that $\int_{\varepsilon \wedge \tau}^\tau d[Y^\infty, Y^\infty] < \infty$ for each $\varepsilon > 0$, but $\int_0^\tau d[Y^\infty, Y^\infty] = \infty$. Nevertheless, if we consider the sequence of harmonic maps $u_i \colon B_1(0) \to S^m$,

$$u_i(x) = \big(\tfrac{x}{\|x\|} \cos h_i(\|x\|),\, \sin h_i(\|x\|)\big)\,,$$

it is true that the total energies $E(u_i)$ are uniformly bounded, i.e., $E(u_i) \leq c < \infty$, while $\|du_i\|^2(0) \to \infty$ as $i \to \infty$. This failure of interior estimates on harmonic maps with bounded total energy is again in sharp contrast to the behavior of harmonic functions. We reconsider such phenomena in the subsequent sections.

Finally, if we also take into account martingales Y on S^m with $Y_\tau = \xi$, not necessarily of the rotationally invariant form (6.1), starting points of such martingales may be quite arbitrary. For instance, if $m = 3$, then it is known [37] that for any $x_0 \in B_1(0) \subset \mathbb{R}^3$, there is a smooth harmonic map

$$u \colon B_1(0) \backslash \{x_0\} \to S^2$$

such that $u|_{\partial B_1(0)} = \mathrm{id}_{\partial B_1(0)}$. Each such map $u = u_{x_0}$ for $x_0 \neq 0$ composed with the Brownian motion X on $\mathbb{R}^3$ (started at $X_0 = 0$) provides a martingale $Y = u \circ X$ on $S^2 \subset S^3$ with $Y_0 = u(0)$ and $Y_\tau = \xi$. For any $0 \neq x_0 \in B_1(0)$, we thus get a martingale on S^3 that actually lives on the equator of S^3, ending up at the prescribed value ξ. By studying the construction of the maps $u = u_{x_0}$ [37], it is not very difficult to see that for appropriate choices of x_0 we can achieve that every point of $S^2 \subset S^3$ appears as initial point $u_{x_0}(0)$ of $u_{x_0} \circ X$; moreover $u_{x_0} \circ X$ is even an H^2-martingale. Translated into our terminology, this says that each point on the equator is a martingale mean of ξ.

We emphasize that the above results should not be interpreted as artifacts of an insufficient conceptual framework; they just reflect the topological nature of the problem and the nontriviality of the Dirichlet problem in this context.

7. Long-term behavior of the heat equation

As explained in section 5, associated to the nonlinear heat equation is the following reachability problem for ∇-martingales. Given $u_0 \in C^\infty(M, N)$, a Brownian motion X^x on (M, g) such that $X_0^x = x$, and the random variable $\xi = u_0 \circ X_t^x$ for some $x \in M$, $t > 0$, the problem is to find an N-valued H^2-martingale $Y = (Y_s)_{0 \le s \le t}$ with $Y_t = \xi$ a.e. The filtration is the m-dimensional Brownian filtration ($m = \dim M$) with respect to which X^x is defined. The relevant question is how such a martingale (unambiguously defined as constant martingale for $t = 0$) changes as t increases.

There are results that guarantee existence and uniqueness of martingales with prescribed end states (see [25], [34], [35], [36], [7], [8]). They naturally require strong restrictions on ξ for existence (either on the range of ξ or on the norm of the derivative of ξ considered as a smooth Wiener functional); uniqueness is usually only given within a certain class of martingales with sufficiently small Riemannian quadratic variation. The results of Kendall [25] give existence and uniqueness of martingale means for random variables with values in small domains, like regular geodesic balls; they exclude effects caused by global geometry; see also [26], [27], [28]. Results of Picard ([35], Theorems 2.2.1 and 3.1.1) cover random variables of the type $\xi = u_0 \circ X_t^x$ (u_0 smooth, M and N compact) at least for sufficiently small t and provide existence and uniqueness of martingale means in this case. Darling [7] constructs ∇-martingales on $\mathbb{R}^n$ with prescribed terminal value under local Lipschitz and convexity conditions on the connection ∇.

For small t, existence of martingales ending up at $\xi = u_0 \circ X_t^x$ at time t is guaranteed by Theorem 5.1 and the short-term solvability of the heat equation: There is always a smooth solution $u \colon [0, T[\times M \to N$ of the heat equation with initial condition $u(0, \cdot) = u_0$, provided $T > 0$ is sufficiently small, and then

$$Y_s = u(t - s, X_s^x), \quad 0 \le s \le t,$$

is the martingale on N starting at $u(t, x)$ with $Y_t = u_0 \circ X_t^x$ a.s. Its Riemannian quadratic variation is given by $d[Y, Y] = \|du(t - s, \cdot)\|^2 (X_s^x)\, ds$. We may choose $T = T(u_0) \le \infty$ such that $[0, T[$ is the maximal interval where the solution exists.

Definition 7.1 Let $u_0 \in C^\infty(M, N)$ be a map between Riemannian manifolds (M, g) and (N, h). For $t < T$ and $x \in M$, let $u(t, x) \in N$ be the martingale mean of $u_0 \circ X_t^x$, as defined by the heat flow deformation of u_0. We say that $u(t, \cdot)$ *blows up at time* $t = T$ if there is a point $x_0 \in M$ such that, for any $\varepsilon > 0$,

$$\limsup_{t \nearrow T} \sup_{x \in B_\varepsilon(x_0)} \|d\,u(t, x)\|^2 = \infty.$$

Thus, given the situation of Definition 7.1, we have one of the three alternatives:

(i) Blow-up in finite time, i.e., there exists $T^* = T^*(u_0) > 0$ such that $u(t, \cdot)$ is regular for $0 \le t < T^* < \infty$, but blows up at $t = T^*(u_0)$.

(ii) $u(t, \cdot)$ is regular for all time, but blows up at $T^* = \infty$.

(iii) $u(t, \cdot)$ is regular for all time, and $u(t, \cdot)$ subconverges to a map $u_\infty \equiv u(\infty, \cdot)$ which is necessarily C^∞ and harmonic; moreover u_∞ is homotopic to u_0.

Lemma 7.2 *Let u: $[0, T[\times M \to N$ be a smooth solution of the heat equation*

$$\frac{\partial}{\partial t} u = \frac{1}{2} \tau(u) \,,$$

and let $e(u) = \|du\|^2$ denote the energy density of u. Then it is true that

$$\left(\frac{\partial}{\partial t} - \frac{1}{2}\Delta_M\right)e(u) + \|\nabla du\|^2 \leq K_M e(u) + K_N e(u)^2 \,, \qquad (7.1)$$

where K_M depends on the Ricci curvature of (M, g), and K_N denotes an upper bound for the sectional curvature of (N, h).

Lemma 7.2 is a well-known consequence of the Bochner formula for the energy density $e(u)$ of u (see [9], section (6.8)). In particular, it shows

$$\left(\frac{\partial}{\partial t} - \frac{1}{2}\Delta_M\right)e(u) \leq c\, e(u)\bigl(1 + e(u)\bigr) \qquad (7.2)$$

with a constant $c = c(M, N)$ depending only on the geometry of M and N. Estimate (7.2) is the basic tool to establish small time solvability of the heat equation.

Moreover, estimate (7.1) easily allows one to exploit the role of negative curvature of the target manifold in the heat flow for harmonic maps; for simplicity, assume that both manifolds are compact.

Theorem 7.3 *Let (M, g) and (N, h) be compact Riemannian manifolds, and suppose that the sectional curvature of N is non-positive, i.e. $\mathrm{Riem}^N \leq 0$. Furthermore, let u: $[0, T[\times M \to N$ be a smooth solution of the heat equation. Then, for $0 < t < T$ and any $x \in M$,*

$$\mathbb{E}\bigl[\|du\|^2(t - s, X_s^x)\bigr] \leq \mathbb{E}\bigl[\|du\|^2(0, X_t^x)\bigr] e^{K_M(t-s)} \,, \quad 0 \leq s \leq t \,, \qquad (7.3)$$

where the constant K_M is given by (7.1) and X^x is $\mathrm{BM}(M, g)$ with $X_0^x = x$. In particular, for $s = 0$, we have the a priori estimate

$$\|du\|^2(t, x) \leq \mathbb{E}\bigl[\|du\|^2(0, X_t^x)\bigr] e^{K_M t} \,, \qquad (7.4)$$

which excludes blow-up in finite time.

Proof. Let $0 \leq s \leq s_0 \leq t$. First, note that $K_N = 0$ in (7.1), as a consequence of the curvature assumption $\mathrm{Riem}^N \leq 0$. Thus, by means of Itô's formula,

$$\mathbb{E}\bigl[e(u)(t - s, X_s^x)\bigr] = \mathbb{E}\bigl[e(u)(t - s_0, X_{s_0}^x)\bigr] + \int_s^{s_0} \mathbb{E}\Bigl[\bigl(\tfrac{\partial}{\partial t} - \tfrac{1}{2}\Delta_M\bigr)e(u)(t - \rho, X_\rho^x)\Bigr]\, d\rho$$

$$\leq \mathbb{E}\bigl[e(u)(t - s_0, X_{s_0}^x)\bigr] + K_M \int_s^{s_0} \mathbb{E}\bigl[e(u)(t - \rho, X_\rho^x)\bigr]\, d\rho \,,$$

and hence, by Gronwall's Lemma,

$$\mathbb{E}\bigl[\|du\|^2(t - s, X_s^x)\bigr] \leq \mathbb{E}\bigl[\|du\|^2(t - s_0, X_{s_0}^x)\bigr] e^{K_M(s_0-s)},$$

which completes the proof. $\qquad\qquad\square$

Obviously (7.3) also gives a bound for the H^2-norm of martingales starting at $u(t, x)$ with end state $u_0 \circ X_t^x$ in terms of the initial energy $E(u_0)$.

Returning to the general case, the above discussion provides the following stochastic picture. Given $u_0 \in C^\infty(M, N)$, for $t < T^* = T^*(u_0)$, there is the martingale $Y_s = u(t - s, X^x_s)$, $0 \le s \le t$, starting from $u(t, x)$ and ending at $u_0 \circ X^x_t$. If blow-up occurs at finite time $T^* < \infty$, as $t \nearrow T^*$, the H^2-norm of this martingale will increase to ∞ (for some x_0). Nevertheless, for $t \ge T^*$, there may be new H^2-martingales with different starting point and terminal state $u_0 \circ X^x_t$. Since, for instance, Brownian motion on (M, g) almost surely never hits subsets of codimension at least two, the stochastic description for $t > T^*$ is not affected by singularities of $u(T^*, \cdot)$ in a subset of codimension at least two; for the martingale $u(t - s, X^x_s)$, $0 \le s \le t$, singularities in a polar set are insignificant because X^x will never see them (if not started in such a point). In particular, this will be the case if there is a smooth solution $(t, x) \mapsto u(t, x)$ of the heat equation on a dense open subset of $\mathbb{R}_+ \times M$ whose complement Σ has the property that $\mathrm{BM}(M, g)$ almost surely never hits any of the sections $(\{t\} \times M) \cap \Sigma$ (if not started there). Then, for any $t > 0$, the martingale $u(t - s, X^x_s)$ is well-defined for $0 < s \le t$, up to the possibly non-specified starting point $u(t, x)$ if $(t, x) \in \Sigma$. Of course, at a singularity the homotopy class of the mapping $u(t, \cdot)$ may change.

In analogy to the classical expectation we use the suggestive notation

$$u(t, x) = \mathcal{E}(u_0 \circ X^x_t), \quad t < T^*, \ x \in M \,,$$

for the martingale mean of $u_0 \circ X^x_t$ defined by the unique solution of the heat equation with initial condition $u(0, \cdot) = u_0$; further we write

$$\mathcal{E}^{\mathscr{F}_s}(u_0 \circ X^x_t), \quad 0 \le s \le t \,,$$

for the corresponding martingale $Y_s = u(t - s, X^x_s)$, $0 \le s \le t$. Note that the Markov property and the law of iterated expectations hold in this context. Indeed, if we write $u_t = S_t(u_0)$ for the solution $u(t, \cdot)$ of the heat equation with initial condition $u(0, \cdot) = u_0$, then $S_s\big(S_t(u_0)\big) = S_s(u_t) = S_{s+t}(u_0)$ for $s + t < T^* = T^*(u_0)$ by the unique solvability of the heat equation; therefore in stochastic terms

$$u_t \circ X^x_s = \mathcal{E}(u_0 \circ X^y_t)\big|_{y=X^x_s} = \mathcal{E}^{\mathscr{F}_s}(u_0 \circ X^x_{s+t}), \quad \text{and}$$

$$\mathcal{E}\big[\mathcal{E}^{\mathscr{F}_s}(u_0 \circ X^x_{s+t})\big] = \mathcal{E}(u_0 \circ X^x_{s+t}) \,,$$

as long as $s + t < T^*$. Especially, we get $\mathcal{E}(u_0 \circ X^x_{s+t}) = \mathcal{E}\big[\mathcal{E}(u_0 \circ X^y_t)\big|_{y=X^x_s}\big]$. Taking iterated expectations then gives, as in the usual theory,

$$\mathcal{E}^{\mathscr{F}_r}\big[\mathcal{E}^{\mathscr{F}_s}(u_0 \circ X^x_t)\big] = \mathcal{E}^{\mathscr{F}_r}(u_0 \circ X^x_t), \quad \text{for } 0 \le r \le s \le t < T^* \,.$$

8. Blow-up in finite time – an example

The construction of martingales with prescribed terminal state leads to backward SDEs. It lies outside the intention of this paper to deal in general with the topic of finding such martingales using stochastic methods (see [7] for results in this direction); in the sequel we use martingale theory as a way to comprehend qualitative aspects of the nonlinear heat equation. In this section we discuss an example for finite-time blow-up; only in context of this example we briefly sketch the aspect of backward SDEs.

Let M be an m-dimensional differentiable manifold and r, n positive integers. Given a vector field $A_0 \in \Gamma(TM)$ and a homomorphism $A\colon M \times \mathbb{R}^r \to TM$ of vector bundles over M, we consider the differential operator in Hörmander form

$$L = A_0 + \tfrac{1}{2} \sum_{i=1}^{r} A_i^2$$

with $A_i = A(\,\cdot\,, e_i) \in \Gamma(TM)$; in addition, we fix differentiable maps

$$F\colon M \times \mathbb{R}^n \times \operatorname{Hom}(\mathbb{R}^r, \mathbb{R}^n) \to \mathbb{R}^n \text{ and } f\colon M \to \mathbb{R}^n\,.$$

Problem 1. Find a smooth function $u\colon [0, t] \times M \to \mathbb{R}^n$ solving

$$\begin{cases} \frac{\partial}{\partial s} u = Lu + F(x, u, du \cdot A) \\ u|_{s=0} = f\,. \end{cases} \tag{$*$}$$

Problem 2. Let X be a diffusion on M with generator L, constructed as solution of the Stratonovich SDE on M,

$$\begin{cases} dX = A_0(X)\,dt + A(X) * dW \\ X_0 = x\,, \end{cases}$$

where W is $\operatorname{BM}(\mathbb{R}^r)$ on Wiener space; further let $\xi = f \circ X_t$. Find continuous adapted $\mathbb{R}^n$-, resp. $\operatorname{Hom}(\mathbb{R}^r, \mathbb{R}^n)$-valued processes $(Y_s)_{0 \le s \le t}$ and $(C_s)_{0 \le s \le t}$ such that

$$\begin{cases} dY = C\,dW - F(X, Y, C)\,ds \\ Y_t = \xi\,. \end{cases} \tag{$**$}$$

Example 8.1 Let $F \equiv 0$, and suppose $\xi \in L^2(P; \mathbb{R}^n)$. Then a solution (Y, C) of $(**)$ is given by

$$Y_s = \mathbb{E}^{\mathscr{F}_s}[\xi] = \mathbb{E}[\xi] + \int_0^s C\,dW,$$

where the matrix process C is determined through Y via Itô's representation theorem.

Lemma 8.2 (i) *Any solution of* $(*)$ *gives a solution of* $(**)$ *via* $Y_s = u(t - s, X_s^x)$, $C_s = du(t - s, X_s^x)\,A(X_s^x, \cdot)$; *in particular,* $Y_0 = u(t, x)$.
(ii) *Under a global Lipschitz condition for* F, *and if in addition* $\xi \in L^2(P; \mathbb{R}^n)$, *there exists a unique pair* (Y, C) *of square-integrable continuous adapted processes solving problem* $(**)$.

Proof. Part (i) is checked directly by means of Itô's formula. Part (ii) is covered by the work of Pardoux-Peng (see [32], [33]) on backward SDEs. $\square$

Note that in PDEs connected to the harmonic map problem the function F depends in general quadratically on du, thus violating a global Lipschitz condition for F. Hence, the existence part of Pardoux-Peng cannot be used directly to determine martingale means.

The heat equation as a system of nonlinear parabolic PDEs is hard to deal with explicitly. With enough symmetries, however, the problem can be reduced to a scalar equation in only two variables.

We consider the following example, studied in [5]; see also [17]. Let $M = \mathbb{R}^m$ and $N = S^m$, $m \geq 3$, and let $u_0 \colon \mathbb{R}^m \to S^m$ be defined by

$$u_0(x) = \left(\frac{x}{r} \cos h_0(r),\ \sin h_0(r) \right), \quad r = \|x\|, \tag{8.1}$$

where $h_0 \colon [0, \infty) \to \mathbb{R}$ is differentiable, monotonic increasing, $h_0(0) = -\pi/2$, and $h_0(\infty) = (2k+1)\frac{\pi}{2}$ with $k \in \mathbb{Z}_+$. Let $u(t, \cdot)$ be the solution of the heat equation with $u(0, \cdot) = u_0$, in other words, $u(t, x) = \mathcal{E}(u_0 \circ X_t^x)$, using the terminology of the last section. By symmetry considerations, we have

$$u(t, x) = \left(\frac{x}{r} \cos h(t, r),\ \sin h(t, r) \right) \tag{8.2}$$

with an appropriate function $h(t, r)$ of two variables. Hence, given the S^m-valued random variable $u_0 \circ X_t^x$ for some fixed $t > 0$ (sufficiently small), the martingale Y with initial point $Y_0 = u(t, x)$ and prescribed end state $Y_t = u_0 \circ X_t^x$ is of the form

$$Y_s = u(t - s, X_s^x) = \left(\Theta_s^x \cos h(t - s, R_s^r),\ \sin h(t - s, R_s^r) \right), \quad 0 \leq s \leq t, \tag{8.3}$$

where $X^x = (\Theta^x, R^r)$ denotes the decomposition of $\mathrm{BM}(\mathbb{R}^m)$ into its angle and radial process. It suffices to determine the function $h(t, r)$.

Remark 8.3 A process Y of the form (8.3) is a martingale on S^m with $Y_t = u_0 \circ X_t^x$ if (8.2) solves the heat equation with $u(0, \cdot) = u_0$ which is equivalent to

$$\frac{\partial h}{\partial t} - \frac{1}{2} \frac{\partial^2 h}{\partial r^2} - \frac{m-1}{2r} \frac{\partial h}{\partial r} - \frac{m-1}{4r^2} \sin 2h = 0 \quad \text{where } h(0, \cdot) = h_0(\cdot). \tag{8.4}$$

The Riemannian quadratic variation $[Y, Y]$ of Y is given by

$$d[Y, Y] = \left(\frac{\partial}{\partial r} h(t - s, R) \right)^2 ds + \frac{m-1}{R^2} \cos^2 h(t - s, R)\, ds. \tag{8.5}$$

Now, let u_0 represent a nontrivial homotopy class in $C(\mathbb{R}^m, S^m)$. Suppose that (8.2) is always extendable to a global smooth solution $u \colon \mathbb{R}_+ \times \mathbb{R}^m \to S^m$ of the heat equation, $u(t, x) = \mathcal{E}(u_0 \circ X_t^x)$ for any $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^m$. The energy of u_0 is given by

$$E(u_0) = \mathrm{vol}(S^{m-1}) \int_0^\infty \left[\left(\frac{\partial h}{\partial r} \right)^2 + \frac{m-1}{r^2} \cos^2 h \right] r^{m-1}\, dr. \tag{8.6}$$

Note that, without changing the homotopy type of u_0, the energy of the initial map can be made as small as we like. For instance, in any homotopy class we can find representatives with $h_0(t) \equiv h_0(\infty)$ for $t \geq \varepsilon_0 > 0$ where ε_0 may be chosen arbitrarily small. Taking such a representative means that u_0 is constant outside a small ball of radius ε_0 about the origin. On the other hand, $\mathrm{BM}(\mathbb{R}^m)$ is transient for $m \geq 3$, which implies that the random variable $u_0 \circ X_t^x$ is almost constant for t large, namely equal to north or south pole, depending on k. Thus, we expect

$$\lim_{t \to \infty} u(t, x) = (0, \ldots, 0, 1) \in S^m \ \text{(north pole) if } k \text{ is even, and}$$

$$\lim_{t \to \infty} u(t, x) = (0, \ldots, 0, -1) \in S^m \ \text{(south pole) if } k \text{ is odd.}$$

Of course, u_0 cannot be homotopic to a constant map, if it represents a nontrivial homotopy class. Hence, the deformation $u(t, \cdot)$ of u_0 must develop singularities. The

problem is to verify that this happens actually in finite time. Heuristically, it is then quite obvious that singularities appear at a finite number of points $(t_\nu, 0)$, since by symmetry, as t increases, $u(t, 0)$ cannot move along great circles through the poles: the points $u(t, 0)$ have to "jump" between the poles.

In stochastic terms the problem can be described as follows. Let R denote an m-dimensional Bessel process $\mathrm{Bes}(m)$ (radial part of $\mathrm{BM}(\mathbb{R}^m)$) with generator

$$\frac{1}{2}\Big(\frac{\partial^2}{\partial r^2} + \frac{m-1}{r}\frac{\partial}{\partial r}\Big);$$

assume that R is defined on the one-dimensional Wiener space and adapted to its one-dimensional Brownian filtration $(\mathscr{F}_s)_{s\in\mathbb{R}_+}$. For fixed $t > 0$, set $\xi = h_0 \circ R_t^r$ with $R_0^r = r$. By Lemma 8.2 (i) (for t not too large) there are two adapted real processes $(Y_s, \mathscr{F}_s)_{0\le s\le t}$ and $(C_s, \mathscr{F}_s)_{0\le s\le t}$ such that

$$\begin{cases} dY_s = C_s\, dW_s - F(R_s^r, Y_s)\, ds & \text{on } [0, t] \times \Omega\,, \\ Y_t = \xi\,, \end{cases}$$

where $F(R_s^r, Y_s) = (m-1)\,(2R_s^r)^{-2}\sin(2Y_s)$. Then $h(t, r) = Y_0$. In other words, we get $\xi = Y_s + \int_s^t C_\rho\, dW_\rho - \int_s^t F(R_\rho^r, Y_\rho)\, d\rho$, which implies

$$Y_s = \mathbb{E}^{\mathscr{F}_s}\Big(\xi + \int_s^t \frac{m-1}{(2R_\rho^r)^2}\sin(2Y_\rho)\, d\rho\Big), \tag{8.7}$$

or equivalently,

$$\mathbb{E}^{\mathscr{F}_s}\Big(\xi + \int_0^t \frac{m-1}{(2R_\rho^r)^2}\sin(2Y_\rho)\, d\rho\Big) = Y_0 + \int_0^s C_\rho\, dW_\rho\,.$$

Equation (8.7) represents Y as sum of the "linear conditional expectation" $\mathbb{E}^{\mathscr{F}_s}[\xi]$ plus an additional term taking into account the nonlinearity in the PDE (8.4). For $r = 0$, we get $\sin(2Y_0) = 0$ as a consequence of

$$\int_0^t \frac{m-1}{(2R_\rho^r)^2}\sin(2Y_\rho)\, d\rho < \infty\,.$$

Therefore, $h(t, 0)$ is necessarily an integer multiple of $\pi/2$.

Formula (8.7) could be used for an explicit discussion of the heat flow for maps of the type (8.1). We are not going to pursue this example in further detail now; its main features will be covered by the general theory developed in the next section.

9. Monotonicity properties for manifold-valued martingales

Let (M, g) and (N, h) be Riemannian manifolds where N is compact. By Nash's theorem, we may assume that (N, h) is isometrically embedded into $\mathbb{R}^\ell$ for some $\ell \in \mathbb{N}$. Let $\pi\colon (T\mathbb{R}^\ell)|N \to TN$ denote the orthogonal projection, which gives a linear map $\pi_y\colon \mathbb{R}^\ell \to T_yN$ for each $y \in N$. Writing $\bar{f} = \iota \circ f \in C^\infty(M, \mathbb{R}^\ell)$ for the composition of a function $f \in C^\infty(M, N)$ with the embedding $\iota\colon N \hookrightarrow \mathbb{R}^\ell$, we have

$$\tau(f)(x) = \pi_{f(x)}(\Delta_M \bar{f})(x)\,, \quad x \in M\,. \tag{9.1}$$

Thus under the identification $\tau(f) \equiv \iota_*\tau(f)$ we get $\Delta_M \bar{f} = \tau(f) + \nu(f)$, where the second summand denotes the normal component of $\Delta_M \bar{f}$. Comparison with the composition formula $\Delta_M(\iota \circ f) = \iota_*\tau(f) + \text{trace}(f^*\nabla d\iota)$ shows $\nu(f) = \text{trace}(f^*\nabla d\iota)$. Hence $\nu(f)$ depends quadratically on the differential df. In particular, the last decomposition implies that $f\colon M \to N$ is harmonic if and only if $(\Delta_M \bar{f})(x) - T_{f(x)}N$ for each $x \in M$.

From now on let $M = \mathbb{R}^m$ $(m \geq 3)$. We consider solutions of the heat equation

$$\frac{\partial}{\partial t}u = \tfrac{1}{2}\tau(u), \quad u\big|_{t=0} = u_0, \tag{9.2}$$

where $f = u_0\colon \mathbb{R}^m \to N$ is smooth and of finite energy $E(f) = \int \|df\|^2\,dx < \infty$. We also assume that $e(u_0) = \|df\|^2$ is bounded. In this case (7.2) provides an a priori estimate for $e(u)$ on a small time interval which guarantees the existence of a solution to (9.2) locally, i.e. on a small time interval $[0, t_1[$ for some $t_1 > 0$. The energy inequality

$$E\big(u(t, \cdot)\big) + \int_0^t \int_{\mathbb{R}^m} \Big\|\tfrac{\partial}{\partial s}u\Big\|^2(s, x)\,ds\,dx \leq E(u_0) \tag{9.3}$$

holds in this case for any $t < t_1$. Moreover, $\|du(t, \cdot)\|^2$ is bounded for each $t < t_1$. Via a fixed embedding $\iota\colon N \hookrightarrow \mathbb{R}^\ell$, solutions of the heat equation take values in $\mathbb{R}^\ell$, and (9.2) reads as

$$\frac{\partial}{\partial t}u - \tfrac{1}{2}\Delta u + \tfrac{1}{2}\nu(u) = 0, \quad u\big|_{t=0} = u_0, \tag{9.4}$$

where Δ is the usual Laplacian on $\mathbb{R}^m$.

Now let $u\colon [0, T[\times \mathbb{R}^m \to N$ be a solution of (9.2) such that

$$E(u_0) < \infty, \text{ and } e(u_0) \text{ is bounded on } \mathbb{R}^m. \tag{9.5}$$

The typical situation we have in mind is as in the example of section 8, where $f = u_0$ has small energy $E(f)$, but represents a nontrivial homotopy class in $C^\infty(\mathbb{R}^m, N)$. To describe the evolution of u_0 under the heat flow, we switch to the stochastic picture. As explained in section 5, for each $(t, x) \in [0, T[\times \mathbb{R}^m$, there is an H^2-martingale $Y^{t,x} \equiv (Y_s^{t,x})_{0 \leq s \leq t}$ with starting point $Y_0^{t,x} = u(t, x)$ and terminal value $Y_t^{t,x} = f \circ X_t^x$ a.s. Let

$$[Y^{t,x}, Y^{t,x}] = \int h(dY^{t,x}, dY^{t,x})$$

denote the Riemannian quadratic variation of $Y^{t,x}$. The next theorem shows that the Riemannian quadratic variations of these martingales share a specific monotonicity property which is basically a consequence of Brownian scaling. All subsequent results are essentially a consequence of this property.

Theorem 9.1 (Monotonicity Formula) *Let $u\colon [0, T[\times \mathbb{R}^m \to N$ be a solution of the heat equation (9.2) such that (9.5) is fulfilled. Then, for each $(t, x) \in [0, T[\times \mathbb{R}^m$ and each $\alpha \in\,]0, 1[$,*

$$\Phi\colon r \mapsto \mathbb{E}\int_{\alpha r}^r h(dY^{t,x}, dY^{t,x}) \tag{9.6}$$

defines a non-decreasing function Φ on $]0, t]$.

We start the proof of Theorem 9.1 by giving a first lemma.

Lemma 9.2 *Let* u: $[0, T[\times \mathbb{R}^m \to N$ *be as in Theorem* 9.1 *and* $(t, x) \in [0, T[\times \mathbb{R}^m$. *Then the following two statements are equivalent:*

 (i) *For each* $\alpha \in {}]0, 1[$ *the function* (9.6) *is non-decreasing on* $]0, t]$.

(ii) *The function* $\phi(s) = s\, \mathbb{E}\left[e(u)(t - s, X_s^x)\right]$ *is non-decreasing on* $]0, t]$.

Proof. First note that $h(dY_s^{t,x}, dY_s^{t,x}) = \|du\|^2(t - s, X_s^x)\, ds = e(u)(t - s, X_s^x)\, ds$. Hence, under the assumption that $\phi(s) = s\, \mathbb{E}\left[e(u)(t - s, X_s^x)\right] =: s\, \varphi(s)$ is non-decreasing on $]0, t]$, we get for $0 < r < r' \leq t$, with $\lambda := r'/r > 1$,

$$\int_{\alpha r}^{r} \varphi(s)\, ds = \int_{\alpha r'}^{r'} \varphi(s'/\lambda)\, ds'/\lambda = \int_{\alpha r'}^{r'} (1/s')\, (s'/\lambda)\, \varphi(s'/\lambda)\, ds'$$

$$\leq \int_{\alpha r'}^{r'} (1/s')\, s'\, \varphi(s')\, ds' = \int_{\alpha r'}^{r'} \varphi(s')\, ds',$$

which verifies condition (i). On the other hand, supposing that (i) is true, consider $\Phi(r) = \mathbb{E} \int_{\alpha r}^{r} h(dY^{t,x}, dY^{t,x}) = \int_{\alpha r}^{r} \varphi(s)\, ds$ for fixed $\alpha \in {}]0, 1[$. Then by assumption $\Phi'(r) = \varphi(r) - \alpha\, \varphi(\alpha r) \geq 0$; hence $\alpha r\, \varphi(\alpha r) \leq r\, \varphi(r)$ for any $r \in {}]0, t]$. Because $\alpha \in {}]0, 1[$ is arbitrary this shows (ii). □

Proof (of Theorem 9.1) By Lemma 9.2 we have to show that

$$s \mapsto s\, \varphi(s) = s\, \mathbb{E}\left[e(u)(t - s, X_s^x)\right] \tag{9.7}$$

is non-decreasing on $]0, t]$. We write $s\, \varphi(s) = s\, \mathbb{E}\left[e(\tilde{u})(s, X_s^0)\right]$ with

$$\tilde{u}(r, y) := u(t - r, x + y)$$

solving $\partial_r \tilde{u} + \frac{1}{2}\tau(\tilde{u}) = 0$ on $[0, t] \times \mathbb{R}^m$, where ∂_r denotes differentiation with respect to the time variable. The claim is a consequence of the following lemma. □

Lemma 9.3 *Over an interval* $]a, b[\subset \mathbb{R}_+$ *let* $\tilde{u}$: $]a, b[\times \mathbb{R}^m \to N$ *be a solution of*

$$\left(\partial_r + \tfrac{1}{2}\tau\right) \tilde{u} = 0 \tag{9.8}$$

such that $e(\tilde{u}) = \|d\tilde{u}\|^2$ *and* $\|\partial_r \tilde{u}\|$ *are bounded on* $]a, b[\times \mathbb{R}^m$. *Then, for a Brownian motion* X^x *on* $\mathbb{R}^m$ *with starting point* x, *the function*

$$\phi:\ s \mapsto s\, \varphi(s) = s\, \mathbb{E}\left[e(\tilde{u})(s, X_s^x)\right] \tag{9.9}$$

is non-decreasing on $]a, b[$.

Proof. We follow a similar calculation of Struwe [41], and show that

$$s \mapsto \phi(s^2) = s^2\, \varphi(s^2)$$

is non-decreasing on $]\sqrt{a}, \sqrt{b}\,]$. Of course, we may assume that $x = 0$. First note that, for fixed $s > 0$, the rescaled function

$$\tilde{u}_s(r, y) := \tilde{u}(s^2 r, sy), \quad (r, y) \in {}\left]a/s^2, b/s^2\right[\times \mathbb{R}^m$$

also solves (9.8).

Since, by the scaling property of $\mathrm{BM}(\mathbb{R}^m, \mathrm{can})$,

$$\mathbb{E}\big[\|d\tilde{u}_s\|^2(1, X_1^0)\big] = s^2\,\mathbb{E}\big[\|d\tilde{u}\|^2(s^2, sX_1^0)\big] = s^2\,\mathbb{E}\big[\|d\tilde{u}\|^2(s^2, X_{s^2}^0)\big]\,,$$

it suffices to check $\frac{d}{ds}\mathbb{E}\big[\|d\tilde{u}_s\|^2(1, X_1^0)\big] \geq 0$. Moreover, by the same reasoning as above, it is enough to consider $s = 1$, i.e., to establish

$$\frac{d}{ds}\Big|_{s=1}\mathbb{E}\big[\|d\tilde{u}_s\|^2(1, X_1^0)\big] \geq 0\,. \tag{9.10}$$

From the definition of $\tilde{u}_s$ follows $\frac{d}{ds}\big|_{s=1}\tilde{u}_s(r, y) = 2r\,\partial_r\tilde{u}(r, y) + d\tilde{u}(r, y)\cdot y$. Notice that $\tilde{u}(r, \cdot)$ is a function from $\mathbb{R}^m$ to $\mathbb{R}^\ell$, and its differential $d\tilde{u}(r, \cdot)$ is hence a function from $\mathbb{R}^m$ to $\mathbb{R}^m \otimes \mathbb{R}^\ell$. Moreover, by the isometry of the embedding $N \hookrightarrow \mathbb{R}^\ell$, we have $\|d\tilde{u}\|^2 = \langle d\tilde{u}, d\tilde{u}\rangle$, where $\langle A, B\rangle = \sum_{i=1}^m \langle Ae_i, Be_i\rangle_{\mathbb{R}^\ell}$ for an orthonormal basis $(e_1, \ldots, e_m)$ of $\mathbb{R}^m$. Since $P\{X_1^0 \in dy\} = g(y)\,dy$, where g denotes standard normal distribution on $\mathbb{R}^m$, we get:

$$\begin{aligned}
\frac{d}{ds}\Big|_{s=1}\mathbb{E}\big[\|d\tilde{u}_s\|^2(1, X_1^0)\big] &= \frac{d}{ds}\Big|_{s=1}\mathbb{E}\big[\langle d\tilde{u}_s, d\tilde{u}_s\rangle(1, X_1^0)\big] \\
&= 2\,\mathbb{E}\big[\langle d\tilde{u}(1, y), d\big(2\,\partial_r\tilde{u}(1, y) + d\tilde{u}(1, y)\cdot y\big)\rangle\big|_{y=X_1^0}\big] \\
&= 2\int_{\mathbb{R}^m}\big\langle d\tilde{u}(1, y), d\big(2\,\partial_r\tilde{u}(1, y) + d\tilde{u}(1, y)\cdot y\big)\big\rangle\,g(y)\,dy \\
&= 2\,\Big[\int_{\mathbb{R}^m}\big\langle d\tilde{u}(1, y), d\big[\big(2\,\partial_r\tilde{u}(1, y) + d\tilde{u}(1, y)\cdot y\big)g(y)\big]\big\rangle\,dy \\
&\quad - \int_{\mathbb{R}^m}\big\langle d\tilde{u}(1, y), (\mathrm{grad}\,g)(y) \otimes \big(2\,\partial_r\tilde{u}(1, y) + d\tilde{u}(1, y)\cdot y\big)\big\rangle\,.
\end{aligned}$$

We substitute $(\mathrm{grad}\,g)(y) = -y\,g(y)$ into the second integral, and apply Green's formula to the first integral, exploiting $\Delta\tilde{u} = -2\,\partial_r\tilde{u} + \nu(\tilde{u})$ and the fact that

$$2\,\partial_r\tilde{u}(1, y) + d\tilde{u}(1, y)\cdot y \in T_{\tilde{u}(1,y)}N\,, \quad y \in \mathbb{R}^m\,.$$

Thus, we finally have

$$\begin{aligned}
\frac{d}{ds}\Big|_{s=1}\mathbb{E}\big[\|d\tilde{u}_s\|^2(1, X_1^0)\big] &= \frac{d}{ds}\Big|_{s=1}\mathbb{E}\big[\langle d\tilde{u}_s, d\tilde{u}_s\rangle(1, X_1^0)\big] \\
&= 2\,\Big[\int_{\mathbb{R}^m}\big\langle 2\,\partial_r\tilde{u}(1, y), 2\,\partial_r\tilde{u}(1, y) + d\tilde{u}(1, y)\cdot y\big\rangle_{\mathbb{R}^\ell}\,g(y)\,dy \\
&\quad + \int_{\mathbb{R}^m}\big\langle d\tilde{u}(1, y)\cdot y, 2\,\partial_r\tilde{u}(1, y) + d\tilde{u}(1, y)\cdot y\big\rangle_{\mathbb{R}^\ell}\,g(y)\,dy\Big] \\
&= 2\int_{\mathbb{R}^m}\big\|2\,\partial_r\tilde{u}(1, y) + d\tilde{u}(1, y)\cdot y\big\|^2 g(y)\,dy \\
&= 2\,\mathbb{E}\big\|2\,\partial_r\tilde{u}(1, X_1^0) + d\tilde{u}(1, X_1^0)\cdot X_1^0\big\|^2 \geq 0\,.
\end{aligned}$$

This completes the proof of Theorem 9.1. $\qquad\qquad\qquad\square$

The proof of Lemma 9.3 allows us to extract explicit information about the growth of function (9.7) which in turn gives a quantitative version of the Monotonicity Formula.

Corollary 9.4 *Let* $]a, b[\subset \mathbb{R}_+$ *and let* $\tilde{u}\colon]a, b[\times \mathbb{R}^m \to N$ *be a solution of*

$$\left(\partial_r + \tfrac{1}{2}\tau\right)\tilde{u} = 0 \tag{9.11}$$

such that $e(\tilde{u}) = \|d\tilde{u}\|^2$ *and* $\|\partial_r\tilde{u}\|$ *are bounded functions on* $]a, b[\times \mathbb{R}^m$. *Further let* $\phi\colon]a, b[\to \mathbb{R}$ *be defined by*

$$\phi(s) = s\,\mathbb{E}\big[e(\tilde{u})(s, X_s^x)\big]. \tag{9.12}$$

Then for any $s < t$ *in* $]a, b[$ *and* $x \in \mathbb{R}^m$ *the following equality holds:*

$$\phi(s) + \int_s^t \mathbb{E}\left\{\frac{\|2\rho\,\partial_r\tilde{u}(\rho, X_\rho^x) + d\tilde{u}(\rho, X_\rho^x)\cdot X_\rho^x\|^2}{\rho}\right\}d\rho = \phi(t)\,.$$

Proof. Again it suffices to consider the case $x = 0$. Note that for solutions $\tilde{u}$ of (9.11), where $\phi(s) \equiv \phi(s; \tilde{u})$ is given by (9.12), we get from the proof of Lemma 9.3

$$\frac{d}{ds}\Big|_{s=1}\phi(s^2; \tilde{u}) = 2\,\mathbb{E}\big\|2\,\partial_r\tilde{u}(1, X_1^0) + d\tilde{u}(1, X_1^0)\cdot X_1^0\big\|^2.$$

On the other hand, we have

$$2\rho^2\,\phi'(\rho^2; \tilde{u}) = \frac{d}{ds}\Big|_{s=1}\phi(\rho^2 s^2; \tilde{u}) = \frac{d}{ds}\Big|_{s=1}s^2\,\mathbb{E}\big[\|d\tilde{u}_\rho\|^2(s^2, X_{s^2}^0)\big] = \frac{d}{ds}\Big|_{s=1}\phi(s^2; \tilde{u}_\rho)\,,$$

where $\tilde{u}_\rho(s, y) = \tilde{u}(\rho^2 s, \rho y)$ is the corresponding rescaled solution of (9.11). The two equations combined give

$$\phi'(\rho; \tilde{u}) = \mathbb{E}\big\|2\rho\,\partial_r\tilde{u}(\rho, X_\rho^0) + d\tilde{u}(\rho, X_\rho^0)\cdot X_\rho^0\big\|^2/\rho$$

which implies the assertion. $\qquad\square$

Theorem 9.1 should be seen as a parabolic version of the monotonicity inequality developed by Schoen and Uhlenbeck (see [39], [38]) for energy minimizing harmonic maps: If $f \in L_1^2(\Omega, N)$ is energy minimizing on some domain $\Omega \subset \mathbb{R}^m$, i.e., if each point x of Ω has a neighborhood U such that $E(f) \le E(f')$ for every $f' \in L_1^2(\Omega, N)$ satisfying $f = f'$ on $\Omega\backslash U$, then

$$r^{2-m}E_r(f) \le \text{const}\,s^{2-m}E_s(f) \tag{9.13}$$

for $0 < r < s < \text{dist}(x, \partial\Omega)$ and $x \in \Omega$, where $E_r(f) = \int_{B_r(x)}\|df\|^2$ is the energy of f on the r-ball $B_r(x)$ about x. Note that in comparison to (9.13) the parabolic monotonicity formula 9.1 (expressed by the monotonicity of the function (9.7)) uses the heat kernel instead of Lebesgue measure on $\mathbb{R}^m$ which seems to be a natural way to take into account the inhomogenity of space-time. We continue with a variation of the Monotonicity Formula.

Corollary 9.5 *Let* $u\colon [0, T[\times \mathbb{R}^m \to N$ *be a smooth solution of the heat equation* (9.2) *such that* (9.5) *is fulfilled. For fixed* $(t, x) \in [0, T[\times \mathbb{R}^m$ *and* $\rho \ge 0$, *consider the martingale* $Y^{t,x;\rho}$ *on* N,

$$Y_s^{t,x;\rho} = u(\rho + t - s, X_s^x), \quad \rho \le s \le \rho + t\,. \tag{9.14}$$

Then the function

$$s \mapsto s\,\mathbb{E}\big[e(u)(\rho + t - s, X_s^x)\big] \tag{9.15}$$

is non-decreasing on $[\rho, \rho + t]$, and for each $\alpha \in \,]0, 1[$

$$r \mapsto \mathbb{E} \int_{\alpha r}^{r} h(dY^{t,x;\rho}, dY^{t,x;\rho}) = \mathbb{E} \int_{\alpha r}^{r} \left[e(u)(\rho + t - s, X_s^x) \right] ds \qquad (9.16)$$

is non-decreasing on $[\rho/\alpha, \rho + t]$.

Proof. Since $\tilde{u}(r, y) := u(\rho + t - r, y)$ solves $\left(\partial_r + \frac{1}{2} \tau \right) \tilde{u} = 0$ on $[\rho, \rho + t]$, the statement about (9.15) is an immediate consequence of Lemma 9.3. Again the statement about (9.16) follows by the same argument as in the proof of Lemma 9.3. $\qquad \square$

Remark 9.6 Let $u \colon [0, T[\, \times \mathbb{R}^m \to N$ be a smooth solution of the heat equation (9.2) such that (9.5) is fulfilled. The same argument as in 9.5 can be applied to

$$Y_s^{T,x;\rho} = u(\rho + T - s, X_s^x), \quad \rho < s \leq \rho + T,$$

showing that

$$s \mapsto s \, \mathbb{E} \left[e(u)(\rho + T - s, X_s^x) \right]$$

is non-decreasing on $]\rho, T + \rho]$. Especially,

$$\lim_{s \searrow 0} \mathbb{E} \left[e(u)(T - s, X_{\rho+s}^x) \right]$$

exists for any $\rho > 0$, even if u blows up at x at time T.

10. Blow-ups in finite time and small energy

Crucial for this section is the observation that Theorem 9.1 allows one to use almost verbatim the machinery developed by Schoen and Uhlenbeck (cf. [38], Theorem 2.2) to establish a priori C^1-bounds for solutions of the heat equation with small energy (see also [41]).

Theorem 10.1 *There exists a constant $\varepsilon_0 = \varepsilon_0(m, N) > 0$ depending only on m and N such that for any solution $u \colon [0, T[\, \times \mathbb{R}^m \to N$ of the heat equation satisfying condition (9.5) the following is true: If $\Phi(r) = \mathbb{E} \int_{r/2}^{r} h(dY^{t,x}, dY^{t,x}) < \varepsilon_0$ for some $(t, x) \in [0, T[\, \times \mathbb{R}^m$ and some r such that $0 < r \leq t < T$, then $\|du\|^2(t, x) \leq C$ with a constant $C = C\big(r, m, N, E(u_0)\big)$.*

Proof. Let $u \colon [0, T[\, \times \mathbb{R}^m \to N$ be an arbitrary solution of the heat equation. Fix $(t_0, x_0) \in [0, T[\, \times \mathbb{R}^m$, and consider $\Phi(r_0) = \mathbb{E} \int_{r_0/2}^{r_0} h(dY^{t_0,x_0}, dY^{t_0,x_0})$ for some $0 < r_0 \leq t_0$. We show that there are (universal) constants $\varepsilon_0 = \varepsilon_0(m, N) > 0$ and $C = C\big(r_0, m, N, E(u_0)\big)$ such that if $\Phi(r_0) < \varepsilon_0$ then $\|du\|^2(t_0, x_0) \leq C$.

Denote by $B_\sigma(x) = \{ y \in \mathbb{R}^m : \|y - x\| < \sigma \}$ the (open) Euclidean ball about $x \in \mathbb{R}^m$ of radius σ. Further, for $\sigma \geq 0$ and $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^m$, consider the "parabolic ball"

$$P(\sigma; t, x) := [t - \sigma^2, t] \times \bar{B}_\sigma(x)$$

at (t, x), where $\bar{B}_\sigma(x) = \{ y \in \mathbb{R}^m : \|y - x\| \leq \sigma \}$ is the closed ball about x of radius σ. Let $0 < r_1 \leq 1$ such that $r_1^2 = r_0/\delta$ with $\delta \geq 4$. First, we claim that for $0 < \sigma < r_1$ and any point $(t, x) \in P(r_1; t_0, x_0)$ the following estimate holds:

$$\frac{1}{\sigma^m} \int_{\sigma^2}^{2\sigma^2} \int_{B_\sigma(x)} e(u)(t + \sigma^2 - s, y) \, ds \, dy \leq c \, \Phi(r_0) + \frac{c^*}{\delta} \, E(u_0) \,, \qquad (10.1)$$

with constants $c = c(m)$ and $c^* = c^*(r_0, m)$. Indeed, note that for $(s, y) \in [\sigma^2, 2\sigma^2] \times B_\sigma(x)$ the Gaussian transition kernel may be estimated from below as

$$p_s(x, y) = \left(\frac{1}{2\pi s} \right)^{m/2} \exp\left(-\frac{\|y - x\|^2}{2s} \right) \geq \frac{1}{c_1} \, \sigma^{-m} \,,$$

where $c_1 = c_1(m)$. In combination with Corollary 9.5 this estimate gives

$$\frac{1}{\sigma^m} \int_{\sigma^2}^{2\sigma^2} \int_{B_\sigma(x)} e(u)(t + \sigma^2 - s, y) \, ds \, dy \leq c_1 \int_{\sigma^2}^{2\sigma^2} \mathbb{E} \, e(u)(t + \sigma^2 - s, X_s^x) \, ds$$

$$\leq c_1 \int_{r_1^2}^{2r_1^2} \mathbb{E} \, e(u)(t + \sigma^2 - s, X_s^x) \, ds \,.$$

Further, if we set $\tau = t_0 - t$, then $0 \leq \tau \leq r_1^2$, and hence for $s \in [r_1^2, 2r_1^2]$

$$\frac{p_s(x, y)}{p_{s+\tau}(x, y)} \leq \left(\frac{s + \tau}{s} \right)^{m/2} \leq c_2 = c_2(m) \,,$$

uniformly in y. Thus, by using Theorem 9.1 once again, we get

$$\mathbb{E} \int_{r_1^2}^{2r_1^2} e(u)(t + \sigma^2 - s, X_s^x) \, ds \leq c_2 \, \mathbb{E} \int_{r_1^2}^{2r_1^2} e(u)(t + \sigma^2 - s, X_{s+\tau}^x) \, ds$$

$$= c_2 \, \mathbb{E} \int_{r_1^2+\tau}^{2r_1^2+\tau} e(u)(t_0 + \sigma^2 - s, X_s^x) \, ds$$

$$\leq c_2 \, \mathbb{E} \left\{ \int_{r_1^2}^{2r_1^2} \dots \, ds + \int_{2r_1^2}^{4r_1^2} \dots \, ds \right\}$$

$$\leq 2c_2 \, \mathbb{E} \int_{r_0/2}^{r_0} e(u)(t_0 + \sigma^2 - s, X_s^x) \, ds$$

$$= 2c_2 \, \mathbb{E} \int_{r_0/2-\sigma^2}^{r_0-\sigma^2} e(u)(t_0 - s, X_{s+\sigma^2}^x) \, ds \,.$$

Finally, for $r_0/2 - \sigma^2 \leq s \leq r_0 - \sigma^2$, we can use the obvious estimate

$$\left| p_{s+\sigma^2}(x, y) - p_s(x_0, y) \right| \leq$$

$$\leq \left(\frac{1}{2\pi} \right)^{m/2} \left| \left(\frac{1}{s + \sigma^2} \right)^{m/2} \exp\left(-\frac{\|y - x\|^2}{2(s + \sigma^2)} \right) - \left(\frac{1}{s} \right)^{m/2} \exp\left(-\frac{\|y - x_0\|^2}{2s} \right) \right|$$

$$\leq c_3 \left| \left(\frac{1}{s + \sigma^2} \right)^{m/2} - \left(\frac{1}{s} \right)^{m/2} \right| + c_4 \left| \exp\left(-\frac{\|y - x\|^2}{2(s + \sigma^2)} \right) - \exp\left(-\frac{\|y - x_0\|^2}{2s} \right) \right|$$

$$\leq c_5 \, r_1^2 = c_6/\delta \,,$$

uniformly in y, with constants c_i depending on m and r_0. From there, together with

$$\mathbb{E}\int_{r_0/2-\sigma^2}^{r_0-\sigma^2} e(u)(t_0-s, X_s^{x_0}) \le \mathbb{E}\left\{\int_{r_0/4}^{r_0/2}\ldots\,ds + \int_{r_0/2}^{r_0}\ldots\,ds\right\}\le 2\,\Phi(r_0)\,,$$

we get, as claimed,

$$\frac{1}{\sigma^m}\int_{\sigma^2}^{2\sigma^2}\int_{B_\sigma(x)} e(u)(t+\sigma^2-s, y)\,ds\,dy \le 4c_2\,\Phi(r_0) + c_2\,c_6\,r_0\,E(u_0)/\delta$$

$$= c(m)\,\Phi(r_0) + c^*(m, r_0)\,E(u_0)/\delta\,.$$

We will exploit (10.1) for σ small. First observe that there exists $\sigma_0 \in [0, r_1[$ such that

$$(r_1-\sigma_0)^2\sup_{P(\sigma_0;t_0,x_0)} e(u) = \max_{0\le\sigma\le r_1}(r_1-\sigma)^2\sup_{P(\sigma;t_0,x_0)} e(u)\,. \tag{10.2}$$

Moreover, there is a point $(t^*, x^*) \in P(\sigma_0; t_0, x_0)$ such that

$$e_0 := \sup_{P(\sigma_0;t_0,x_0)} e(u) = e(u)(t^*, x^*)\,.$$

We may assume that $e_0 > 0$. If we set $\rho_0 = (r_1 - \sigma_0)/2$, then $0 < \sigma_0 + \rho_0 < r_1$, and hence from the choice of σ_0 and (t^*, x^*)

$$\sup_{P(\rho_0;t^*,x^*)} e(u) \le \sup_{P(\sigma_0+\rho_0;t_0,x_0)} e(u) \le 4\,e_0\,. \tag{10.3}$$

Introduce $\rho_1 := \sqrt{e_0}\,\rho_0$, and rescale u as

$$v(t, x) := u\left(\frac{t-t_0}{e_0} + t^*, \frac{x-x_0}{\sqrt{e_0}} + x^*\right)\,, \quad (t, x) \in P(\rho_1; t_0, x_0)\,,$$

then v provides a solution of the nonlinear heat equation on $P(\rho_1; t_0, x_0)$ chosen such that $e(v)(t_0, x_0) = 1$. Moreover, by (10.3) we have

$$\sup_{P(\rho_1;t_0,x_0)} e(v) = \frac{1}{e_0}\sup_{P(\rho_0;t^*,x^*)} e(u) \le 4\,.$$

We want to show that $\rho_1 \le 1$. First note that Lemma 7.2 implies

$$\left(\tfrac{\partial}{\partial t} - \tfrac{1}{2}\Delta\right)e(v) \le C_1\,e(v)$$

on $P(\rho_1; t_0, x_0)$, where the constant C_1 depends only on N. Hence, if instead of $e(v)$ we consider

$$\tilde{e}(t, x) := \exp\big(C_1(t_0 - t)\big)\,e(v)(t, x)$$

on $P(\rho_1; t_0, x_0)$, and if $\rho_1 \ge 1$, we get from Moser's weak Harnack inequality for subsolutions of parabolic equations ([31], Theorem 3, p. 113) the estimate

$$1 = e(v)(t_0, x_0) \le C_2\int_{P(1;t_0,x_0)} e(v)(s, y)\,ds\,dy\,, \tag{10.4}$$

where the constant C_2 depends on C_1. Scaling back gives

$$\int_{P(1;t_0,x_0)} e(v)(s, y)\,ds\,dy = e_0^{m/2}\int_{P(1/\sqrt{e_0};t^*,x^*)} e(u)(s, y)\,ds\,dy\,,$$

so that (10.4) leads to

$$1 \le C_2 e_0^{\,m/2} \int_{t^* - 1/e_0}^{t^*} \int_{B_{1/\sqrt{e_0}}(x^*)} e(u)(s, y)\, ds\, dy$$

$$= C_2 e_0^{\,m/2} \int_{1/e_0}^{2/e_0} \int_{B_{1/\sqrt{e_0}}(x^*)} e(u)(t^* + 1/e_0 - s, y)\, ds\, dy$$

$$\le C_3\, \Phi(r_0) + C_4\, E(u_0)/\delta = C_3\, \varepsilon_0 + C_4\, E(u_0)/\delta\,,$$

with uniform constants $C_3 = C_3(m, N)$ and $C_4 = C_4(r_0, m, N)$; note that the inequality in the last line follows from (10.1), since $1/\sqrt{e_0} \le \rho_0 < r_1$ as a consequence of the assumption $\rho_1 = \sqrt{e_0}\,\rho_0 \ge 1$. If we choose $\varepsilon_0 < 1/(2 C_3)$ and $\delta > 2\, C_4\, E(u_0)$, we obtain a contradiction. Thus only $\rho_1 \le 1$ is possible, and hence, by the choice of σ_0,

$$\max_{0 \le \sigma \le r_1} (r_1 - \sigma)^2 \sup_{P(\sigma; t_0, x_0)} e(u) = (r_1 - \sigma_0)^2 \sup_{P(\sigma_0; t_0, x_0)} e(u) \le 4\rho_0^2 e_0 = 4\rho_1^2 \le 4\,.$$

Finally, choosing $\sigma = r_1/2$, hence $(r_1 - \sigma)^2 = r_0/4\delta$, we get for δ sufficiently large

$$\sup_{P\left((r_0/4\delta)^{1/2}; t_0, x_0\right)} e(u) \le 16\,\delta/r_0\,. \tag{10.5}$$

The proof is complete. $\square$

From (10.5) we read off more precise quantitative information how smallness of Riemannian quadratic variation translates into gradient bounds for solutions of the heat equation.

Corollary 10.2 *Let* $u\colon [0, T[\,\times \mathbb{R}^m \to N$ *be a solution of the heat equation satisfying condition* (9.5). *Suppose that*

$$\Phi(r_0) = \mathbb{E} \int_{r_0/2}^{r_0} h(dY^{t_0, x_0}, dY^{t_0, x_0}) < \varepsilon_0$$

for some $(t_0, x_0) \in [0, T[\,\times \mathbb{R}^m$ *such that* $0 < r_0 \le t_0 < T$, *where* ε_0 *is determined by Theorem* 10.1. *Then, there exists* $\delta = \delta\big(r_0, m, N, E(u_0)\big)$ *such that*

$$\sup\big\{ e(u)(t_0 - s, x)\colon\ 0 \le s \le r_0/\delta\,,\ \|x - x_0\|^2 \le r_0/\delta \big\} \le 4\delta/r_0\,. \tag{10.6}$$

Moreover, δ *can be chosen independent of* r_0 *as long as* r_0 *exceeds some positive lower bound.*

Theorem 10.1 can easily be adapted to obtain a global existence result for solutions of the heat equation.

Theorem 10.3 *For any* $T > 0$ *there exists a constant* $\varepsilon_1 = \varepsilon_1(m, N, T)$ *depending only on* m, N *such that any solution* $u\colon [0, T[\,\times \mathbb{R}^m \to N$ *of the heat equation with* $e(u_0)$ *bounded on* $\mathbb{R}^m$ *and* $E(u_0) < \varepsilon_1$ *can be extended to a global (smooth) solution* $u\colon [0, \infty[\,\times \mathbb{R}^m \to N$ *which converges to a constant harmonic map* u_∞ *as* $t \to \infty$.

Proof. Given $(t, x) \in [0, T[\,\times \mathbb{R}^m$ and $0 < r \le t$, we get by means of the Monotonicity Formula (see Corollary 9.5 with $\rho = 0$)

$$\mathbb{E} \int_{r/2}^{r} h(dY^{t,x}, dY^{t,x}) = \int_{r/2}^{r} \mathbb{E}\, e(u)(t-s, X_s^x)\, ds = \int_{r/2}^{r} \frac{1}{s}\, s\, \mathbb{E}\big[e(u)(t-s, X_s^x)\big]\, ds$$

$$\leq t\, \mathbb{E}\big[e(u_0)(X_t^x)\big] \int_{r/2}^{r} \frac{1}{s}\, ds = t\, \mathbb{E}\big[e(u_0)(X_t^x)\big] \ln 2$$

$$\leq c(t,m)\, E(u_0) < \varepsilon_0\,,$$

for $\varepsilon_1 < \varepsilon_0/c(t,m)$ where $\varepsilon_0 = \varepsilon_0(m,N)$ is the constant determined by Theorem 10.1. Note that the explicit form of the constant $c(t,m)$ is $t \ln 2\,(2\pi t)^{-m/2}$ and that $m \geq 3$ by assumption; hence the same constant ε_1 that works for t also applies for $t' > t$; especially ε_1 may be chosen independent of t as long as t exceeds some positive value, say $t \geq t_0$. However, as a consequence of Corollary 10.2, if $E(u_0) \leq \varepsilon_1 = \varepsilon_1(m,N,t_0)$, then there is a positive $\alpha = \alpha\big(t_0, m, N, E(u_0)\big)$ and a uniform a priori bound

$$\|du\|^2(s, \cdot) \leq C/T \tag{10.7}$$

for $T - \alpha < s < T$ with a universal constant C depending only on t_0, m, N, $E(u_0)$. Now, given a smooth solution of the heat equation $u\colon [0, T[\, \times \mathbb{R}^m \to N$ with $e(u_0)$ bounded and $E(u_0) < \varepsilon_1 = \varepsilon_1(m, N, t_0)$, let $u\colon [0, T'[\, \times \mathbb{R}^m \to N$ be its maximal extension in the sense that u can not be extended beyond the time interval $[0, T'[$ as a smooth solution of the heat equation. Hence, if $T' < \infty$, then $u(t, \cdot)$ will blow up at $t = T'$. But by the above reasoning

$$\|du\|^2(s, \cdot) \leq C/T'$$

for $T' - \alpha < s < T'$, which shows that blow up at time T' is impossible. As a consequence, we must have $T' = \infty$, which means that u is extendable to a global solution $u\colon [0, \infty[\, \times \mathbb{R}^m \to N$. The fact that $u(t, \cdot) \to u_\infty \equiv \text{const}$ as $t \to \infty$ can easily be derived from the uniform global decay

$$\|du\|^2(t, \cdot) \leq C/t$$

for $t \geq T$. $\qquad\qquad\square$

Note that in Theorem 10.3 the initial map u_0 is automatically homotopic to a constant map. On the other hand, in case of homotopically nontrivial maps u_0, Theorem 10.3 immediately leads to blow-up results in finite time.

Corollary 10.4 *Let $T > 0$, and let $\varepsilon_1 = \varepsilon_1(m, N, T)$ be determined by Theorem 10.3. Then, for homotopically nontrivial smooth initial data $u_0\colon \mathbb{R}^m \to N$ such that $e(u_0) \in L^\infty$ and $E(u_0) < \varepsilon_1$, solutions of the heat equation*

$$\tfrac{\partial}{\partial t} u = \tfrac{1}{2}\tau(u), \quad u\big|_{t=0} = u_0 \tag{10.8}$$

blow up before time T. In fact, blow-up time $T^* = T^*(u_0)$ approaches 0 as $E(u_0)$ decreases to 0.

Proof. Otherwise, by Theorem 10.3, there would exist a global solution to (10.8) inducing a homotopy $u_0 \simeq u_\infty \equiv \text{const}$, in contradiction to the assumption that u_0 is homotopically nontrivial. $\qquad\square$

11. Conclusion

Deformation by heat flow reduces the total energy of the initial map $u_0 \colon M \to N$, while the average energy along paths of Brownian motion (measured by the H^2-norm of the martingales with prescribed terminal state $u_0 \circ X_t^x$) may blow up. We discussed the case $M = \mathbb{R}^m$ with the flat Brownian motion; analogous results can be given in the general situation by exploiting the fact that locally about each point any Riemannian metric is approximately Euclidean (see [42]).

Specifically, if $u \colon [0, t_0[\times M \to N$ is a smooth solution of the heat equation, we have the N-valued martingale $Y_s = u(t_0 - s, X_s^x)$, $0 < s \le t_0$, and the function $\Phi(r) = \mathbb{E} \int_{r/2}^{r} h(dY, dY)$ is non-decreasing on $]0, t_0]$. Hence, for $x \in M$, there are two possibilities:

(i) $\lim_{r \to 0} \Phi(r) > 0$, which means that blow-up occurs at (t_0, x); in this case the process $(Y_s)_{0 < s \le t_0}$ is a martingale without starting point.

(ii) $\lim_{r \to 0} \Phi(r) = 0$, which implies that the heat flow u can be extended around (t_0, x); moreover $Y_0 = u(t_0, x)$.

There are many important related questions. Of course, the above discussion stresses the desirability of constructing appropriate martingales with purely probabilistic methods and without relying on solutions the heat equation. For instance, let $f \colon M \to N$ be a smooth map between compact Riemannian manifolds (M, g) and (N, h). Is there always an H^2-martingale Y_s (defined at least for $0 < s \le t$) with terminal state $f \circ X_t^x$, even for $t > T^*$, if T^* is the first blow-up time in the heat flow with initial map f? What is a straightforward stochastic way to construct such martingales in general situations? (We deal with these questions in [41], [42]).

Note that the heat equation is no longer well-defined as classical PDE beyond the first singularity; one has to switch to the framework of distributional solutions. Nevertheless, questions about martingale means of random variables of the type $\xi = f \circ X_t^x$ make sense for arbitrary values of t.

Finding starting points of martingales with a prescribed terminal state puts the heat equation in a quite general setting and provides "solutions" to the heat equation in a weak but canonical sense.

References

1. Arnaudon, M.: Espérances conditionnelles et $\mathcal{C}$-martingales dans les variétés. In: Azéma, J. et al. (eds.) Séminaire de Probabilités XXVIII. Lect. Notes in Math. **1583**, 300–311 (1994)

2. Arnaudon, M.: Barycentres convexes et approximations des martingales continues dans les variétés. In: Azéma, J. et al. (eds.) Séminaire de Probabilités XXIX. Lect. Notes in Math. **1613**, 70–85 (1995)

3. Chang, K.-C., Ding, W.-Y., Ye, R.: Finite-time blow-up of the heat flow of harmonic maps from surfaces. J. Differ. Geom. **36**, 507–515 (1992)

4. Chen, Y., Ding, W.-Y.: Blow-up and global existence for heat flows of harmonic maps. Invent. math. **99**, 567–578 (1990)

5. Coron, J.M., Ghidaglia, J.M.: Explosion en temps fini pour le flot des applications harmoniques. C. R. Acad. Sci. Paris, Sér. I. **308**, 339–344 (1989)

6. Darling, R. W. R.: Martingales in manifolds – definition, examples, and behaviour under maps. In: Azéma, J., Yor, M. (eds.) Séminaire de Probabilités XVI, 1980/81. Supplement: Géométrie Différentielle Stochastique. Lect. Notes in Math. **921**, 217–236 (1982)

7. Darling, R. W. R.: Constructing gamma-martingales with prescribed limit, using backwards SDE. Ann. Probab. **23**, 1234–1261 (1995)

8. Darling, R. W. R.: Martingales on noncompact manifolds: maximal inequalities and prescribed limits. Ann. Inst. Henri Poincaré, Probab. Stat. **32**, 1–24 (1996)

9. Eells, J., Lemaire, L.: A report on harmonic maps. Bull. London Math. Soc. **10**, 1–68 (1978)

10. Eells, J., Lemaire, L.: Selected topics in harmonic maps. American Mathematical Society: Reg. Conf. Ser. Math. **50** (1983)

11. Eells, J., Lemaire, L.: Another report on harmonic maps. Bull. London Math. Soc. **20**, 385–524 (1988)

12. Eells, J., Sampson, J. H.: Harmonic mappings of Riemannian manifolds. Amer. J. Math. **86**, 109–160 (1964)

13. Elworthy, K. D.: Stochastic Differential Equations on Manifolds. London Math. Soc. Lecture Notes Series **70**. Cambridge: Cambridge University Press, 1982

14. Elworthy, K. D.: Geometric aspects of diffusions on manifolds. In: Hennequin, P. L. (ed.) École d'Été de Probabilités de Saint–Flour XV–XVII. Lect. Notes in Math. **1362**, 277–425 (1988)

15. Emery, M.: Stochastic Calculus in Manifolds (with an appendix by P. A. Meyer). Berlin Heidelberg New York: Springer, 1989

16. Emery, M., Mokobodzki, G.: Sur le barycentre d'une probabilité dans une variété. In: Azéma, J. et al. (eds.) Séminaire de Probabilités XXV. Lect. Notes in Math. **1485**, 277–425 (1991)

17. Grotowski, J. F.: Harmonic map heat flow for axially symmetric data. Manuscr. Math. **73**, 207–228 (1991)

18. Hamilton, R. S.: Harmonic maps of manifolds with boundary. Lect. Notes in Math. **471** (1975)

19. Hackenbroch, W., Thalmaier, A.: Stochastische Analysis. Eine Einführung in die Theorie der stetigen Semimartingale. Stuttgart: Teubner, 1994

20. Ikeda, N., Watanabe, S.: Stochastic Differential Equations and Diffusion Processes. 2nd Edition. Amsterdam: North Holland Publ. Comp. 1989

21. Jost, J.: Ein Existenzbeweis für harmonische Abbildungen, die ein Dirichletproblem lösen, mittels der Methode des Wärmeflusses. Manuscr. Math. **34**, 17–25 (1981)

22. Jost, J.: Nonlinear Methods in Riemannian and Kählerian Geometry. DMV Seminar, Bd. 10. Basel: Birkhäuser, 1988

23. Karcher, H.: Riemannian center of mass and mollifier smoothing. Comm. Pure Appl. Math. **30**, 509–541 (1977)

24. Kendall, W. S.: Martingales on manifolds and harmonic maps. In: Durrett, R., Pinsky, M. A. (eds.) Geometry of Random Motion. American Mathematical Society. Contemporary Mathematics. Vol. 73, 121-157 (1988). Also: Huang, H., Kendall, W. S. Correction note to "Martingales on manifolds and harmonic maps". Stochastics **37**, 253–257 (1991)

25. Kendall, W. S.: Probability, convexity, and harmonic maps with small image I: Uniqueness and fine existence. Proc. London Math. Soc. **61**, 371–406 (1990)

26. Kendall, W. S.: Convexity and the hemisphere. J. Lond. Math. Soc. (2) **43**, 567–576 (1991)

27. Kendall, W. S.: Convex geometry and nonconfluent Γ-martingales I: tightness and strict convexity. In: Barlow, M. T., Bingham, N. H. (eds.) Stochastic Analysis. Proceedings of the Durham Symposium on Stochastic Analysis, 1990. London Mathematical Society Lecture Note Series. Vol. 167, 163–178 (1991)

28. Kendall, W. S.: Convex geometry and nonconfluent Γ-martingales II: well-posedness and Γ-martingale convergence. Stochastics **38**, 135–147 (1992)

29. Kendall, W. S.: Probability, convexity, and harmonic maps II. Smoothness via probabilistic gradient inequalities. J. Funct. Anal. **126**, 228–257 (1994)

30. Kobayashi, S., Nomizu, K.: Foundations of Differential Geometry. Vol. II. New York: Interscience Publishers, 1969

31. Moser, J.: A Harnack inequality for parabolic differential equations. Comm. Pure Appl. Math. **17**, 101–134 (1964)

32. Pardoux, E., Peng, S.: Adapted solution of a backward stochastic differential equation. Systems and Control Letters **14**, 55–61 (1990)

33. Peng, S.: Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. Stochastics **37**, 61–74 (1991)

34. Picard, J.: Martingales sur le cercle. In: Azéma, J. et al. (eds.) Séminaire de Probabilités XXV. Lect. Notes in Math. **1372**, 147–160 (1989)
35. Picard, J.: Martingales on Riemannian manifolds with prescribed limit. J. Funct. Anal. **99**, 223–261 (1991)
36. Picard, J.: Barycentres et martingales sur une variété. Ann. Inst. Henri Poincaré, Prob. Stat. **30**, 647–702 (1994)
37. Poon, C.-C.: Some new harmonic maps from B^3 to S^2. J. Differ. Geom. **34**, 165–168 (1991)
38. Schoen, R.M.: Analytic aspects of the harmonic map problem. In: Chern, S.S. (ed.) Seminar on Nonlinear Partial Differential Equations. Mathematical Sciences Research Institute Publications. New York Berlin Heidelberg Tokyo: Springer-Verlag, 1984
39. Schoen, R.M., Uhlenbeck, K.: A regularity theory for harmonic maps. J. Differ. Geom. **17**, 307–335 (1982)
40. Schoen, R.M., Uhlenbeck, K.: Regularity of minimizing harmonic maps into the sphere. Invent. math. **78**, 89–100 (1984)
41. Struwe, M.: On the evolution of harmonic maps in higher dimensions. J. Differ. Geom. **28**, 485–502 (1988)
42. Thalmaier, A.: Martingales on Riemannian manifolds and the nonlinear heat equation. In: Davies, I.M., Truman, A., Elworthy, K.D. (eds.) Stochastic Analysis and Applications, Gregynog, 1995. Proc. of the Fifth Gregynog Symposium. Singapore: World Scientific Press, 1996, 429–440
43. Thalmaier, A.: Martingale solutions to the nonlinear heat equation: existence and partial regularity (in preparation)
44. White, B.: Infima of energy functionals in homotopy classes of mappings. J. Differ. Geom. **23**, 127–142 (1986)