Gradient Estimates on Manifolds Using Coupling

M. CRANSTON*

Ray P. Hylan Building, University of Rochester,
Rochester, New York 14627

Communicated by Paul Malliavin
Received August 1990

We use a coupling method to give gradient estimates for solutions to
\[(\frac{1}{2} D + Z)u = 0\]
on a manifold. The size of the gradient depends on a lower bound
on the Ricci curvature of the manifold and bounds on the vector field \(Z\).

\[|\nabla u(x)| < c \sup_{Q'} |u|\]
for \(x \in Q' \subset Q\), where \(Q'\) is an open subset of \(Q\) on which \(u\) is a solution.

One important point is to determine the dependence of \(c\), as in Yau's
estimate, on \(M, Z, Q',\) and \(Q\). This dependence is made explicit in
Theorem 2 below. There it is seen that \(c\) depends on \(M\) only through a
lower bound on Ricci curvature and on the dimension. This dependence
arises in our approach as we need to bound a second variation of arclength
and this variation is controlled by Ricci curvature. The dependence of \(c\) on
\(Z\) arises from a first variation of arclength.

Our approach uses a coupling method on manifolds due to Kendall [4].
Kendall's coupling is a generalization of work of Lindvall and Rogers [6]
on Euclidean space and was preceded by similar work on harmonic maps.

* Research supported by a grant from NSA.
in Kendall [2]. By a coupling, we shall mean that \((X, Y, P^x)_t\) is a diffusion on \(M \times M\) with \(P^x\,^y(X = x, Y = y) = 1\), such that under the measure \(P^x\,^y\) the marginal processes \(X\) and \(Y\) are both diffusions on \(M\) with generator \(L = \frac{1}{2} \mathcal{A} + Z\). Other applications of coupling have appeared in Kendall [5] and Lyons and Sullivan [7]. For gradient estimates, take a solution \(U_t = 0\) on some open set \(Q \subset M\), and define for an open \(Q' \subset Q\), \(\tau_{Q'}(X) = \inf\{t > 0 : X \notin Q'\}\) with a similar definition for \(\tau_{Q'}(Y)\). Finally, set \(T(X, Y) = \inf\{t > 0 : X_t = Y_t\}\) and assume, as we later achieve by construction, that \(Y_t \equiv X_t, t > T(X, Y)\). Then

\[
|u(x) - u(y)| = \left| E^{x, y}[u(X_{\tau_{Q'}}) - u(Y_{\tau_{Q'}}); T(X, Y) > \tau_{Q'}(X) \wedge \tau_{Q'}(Y)] \right| \\
\leq 2 \sup_{Q'} |u| \cdot P^{x, y}(T(X, Y) > \tau_{Q'}(X) \wedge \tau_{Q'}(Y)).
\]  

By modifying Kendall's [4] construction, we are able to create a suitable pair \((X, Y)\) and estimate

\[
P^{x, y}(T(X, Y) > \tau_{Q'}(X) \wedge \tau_{Q'}(Y)).
\]

This is the content of

**Theorem 1.** Suppose \((M, g)\) is a complete \(d\)-dimensional Riemannian manifold with distance \(\rho_M\) and assume

\[
\text{Ric}_M \geq -Kg.
\]

Let \(Z\) be a \(C^1\) vector field on \(M\) for which there is a constant \(m\) such that

\[
|Z(x)| \leq m, \quad x \in M.
\]

Then there is a constant \(c = c(K, d, m)\) such that for all \(x, y \in B(x_0, \delta)\)

\[
P^{x, y}(T(X, Y) > \tau_{B(x_0, 2\delta)}(X) \wedge \tau_{B(x_0, 2\delta)}(Y)) \\
\leq c \left( \frac{1}{\delta} + 1 \right) \rho_M(x, y).
\]  

Furthermore,

\[
P^{x, y}(T(X, Y) = \infty) \leq (2 \sqrt{K(d - 1) + 2m}) \rho_M(x, y) \tag{0.3}
\]

and in case \(Z = 0\),

\[
P^{x, y}(T(X, Y) = \infty) \leq 2 \sqrt{K(d - 1)} \rho_M(x, y). \tag{0.4}
\]

Theorem 1 and the estimate (0.1) imply
THEOREM 2. Suppose \((M, g)\) and \(Z\) are as in Theorem 1. There is a constant \(c = c(K, d, m)\) such that whenever \(\delta > 0\) and \(Lu = 0\) in some \(B(x_0, 2\delta)\) then

\[
|\nabla u(x)| \leq c \left( \frac{1}{\delta} + 1 \right) \sup_{B(x_0, 3\delta/2)} u, \quad x \in B(x_0, \delta). \tag{0.5}
\]

If \(Lu = 0\) on \(M\) and \(u\) is bounded and positive, then

\[
|\nabla u(x)| \leq (2 \sqrt{K(d-1) + m}) \|u\|_\infty. \tag{0.6}
\]

If \(Z = 0\), then

\[
|\nabla u(x)| \leq 2 \sqrt{K(d-1)} \|u\|_\infty. \tag{0.7}
\]

Proof. The first estimate follows easily from (0.1) and (0.2) of Theorem 1 by dividing through by \(\rho_M(x, y)\) and letting \(y \to x\). For the second and third estimates,

\[
|u(x) - u(y)| = |E^{x,y} [u(X_t) - u(Y_t)]|
= |E^{x,y} [u(X_t) - u(Y_t); T(X, Y) > t]| \\
\leq \|u\|_\infty P^{x,y}(T(X, Y) > t).
\]

Letting \(t \to \infty\)

\[
|u(x) - u(y)| \leq \|u\|_\infty P^{x,y}(T(X, Y) = \infty).
\]

Now use (0.3) and (0.4) to get (0.6) and (0.7). \(\square\)

If the \(u\) in Theorem 2 happens to be positive, the bound in (0.5) can be improved to

\[
|\nabla u(x)| \leq \frac{c}{\delta} u(x), \quad x \in B(x_0, \delta). \tag{0.8}
\]

However, the constant \(c\) in (0.8) which comes from the constant in (0.5) and Harnack's inequality now depends on upper bounds on sectional curvature. That is, so far as this author knows, Harnack's constant depends on upper bounds on sectional curvature.

1. THE COUPLING OF KENDALL

The purpose of this section is to review, modify, and extend Kendall's coupling [4] of Brownian motions on \(M\). The key result is formula (1.7)
below which exhibits the behavior of the distance between the two diffusing particles with generator $L = \frac{1}{2} \Delta + Z$. If the development here is too abbreviated, the Kendall [4] or Ikeda and Watanabe [3] reference may prove helpful.

Suppose then that $(M, g)$ is a complete Riemannian manifold. Let $\mathcal{C}(M)$ denote the orthonormal frame bundle and $\pi: \mathcal{C}(M) \to M$ the projection onto $M$. The canonical, torsion-free connection on $M$ gives rise to the horizontal lift

$$H: TM \to T\mathcal{C}(M).$$

Given $\xi \in \mathcal{C}(M)$,

$$H_\xi: T_{\pi\xi}M \to T_\xi\mathcal{C}(M)$$

preserves inner products and $\pi_\ast H_\xi$ is the identity map on $T_{\pi\xi}M$.

Given a semi-martingale $M$, on $\mathbb{R}^d$, $M$, can be printed on $M$ using the above apparatus. Denoting by $\partial$ the Stratonovich differential and by $d$ the Itô differential, begin by creating a moving frame by solving

$$\partial \Xi = H_\Xi \Xi \partial M$$

$$\Xi_0 = \xi$$

and then projecting, define

$$X = \pi \Xi.$$

From the fact that $\pi_\ast H_\xi$ is the identity map, it follows that

$$\partial X = \Xi \partial M,$$

$$X_0 = x = \pi \xi.$$

For our semi-martingale $M$, we shall assume

$$dM_t = dB_t + U_t \, dt,$$

with $B$ a Brownian motion on $\mathbb{R}^d$, $BM(\mathbb{R}^d)$, and $U_t$ a bounded predictable process on $\mathbb{R}^d$. The process $\Xi_t$ (and thus $X_t$) is defined for all $t \geq 0$.

Suppose now that $y \notin C(x)$, the cut locus of $x$. Following Kendall [4] we create another process $Y$ on $M$ started at $y$ with "characteristics" similar to those of $X$.

To this end, take $\gamma$ to be the unique unit speed intervening geodesic between $x$ and $y$ and define the mirror map

$$m_{xy}: T_x M \to T_y M,$$
which acts by parallel transporting along $\gamma$ a vector $v \in T_xM$ to $T_yM$ and then reflecting this in the hyperplane of $T_yM$ which is perpendicular to the incoming geodesic $\gamma$. Note that $m_{xy}$ is an isometry from $T_xM$ onto $T_yM$.

Then, given two starting points $x, y, x \notin C(y)$, one can solve for $t < \sigma_c(Y) \wedge T(X, Y)$, $\sigma_c(Y) = \inf\{t > 0 : Y_t \in C(X_t)\}$, $T(X, Y) = \inf\{t > 0 : X_t = Y_t\}$,

$$
\partial \Psi_t = H \Psi \partial N_t,
$$

$$
\Psi_0 = \eta, \quad \pi \eta = y
$$

$$
Y_t = \pi \Psi_t,
$$

with

$$
dN_t = dA_t + V_t \, dt,
$$

where

$$
dA_t = \Psi_t^{-1} m_{X_t, Y_t} \Xi_t \, dB_t
$$

and $V_t$ is again a bounded predictable process. If $t \geq T(X, Y)$ simply set $Y_t = X_t$.

Note that $A$ is $BM(\mathbb{R}^d)$ since each linear map in the composition (acting on $dB$) is an isometry.

Our goal is to select the processes $U$ and $V$ so that $X$ and $Y$ will each be diffusions on $M$ with generator $\frac{1}{2} A + Z$, where $Z$ is a vector field on $M$, $X$ running for all time, $Y$ up to $\sigma_c(Y)$.

Applying Itô’s formula to $f(X_t)$ will reveal the proper choice. Fix orthonormal bases $\{u^i\}_{i=1}^d$ and $\{v^i\}_{i=1}^d$ for $\mathbb{R}^d$. Then

$$
\partial f(X_t) = f_\ast(X_t) \Xi_t u^i \partial M_i
$$

and setting

$$
g^i(X_t) = \Xi_t u^i f(X_t)
$$

$$
\partial g^i(X_t) = \Xi_t u^i \Xi_t u^j f(X_t) \partial M_j.
$$

Consequently, since $X \partial Y = X \, dY + \frac{1}{2} d\langle X, Y \rangle$,

$$
\begin{align*}
df(X_t) &= \Xi_t u^i f(X_t) \, dB_t + \left[ \Xi_t u^i U_i f(X_t) + \frac{1}{2} \sum_{i=1}^d (\Xi_t u^i)^2 f(X_t) \right] \, dt \\
&= \Xi_t u^i f(X_t) \, dB_t + \left[ \Xi_t u^i U_i f(X_t) + \frac{1}{2} \Delta f(X_t) \right] \, dt.
\end{align*}
$$
Thus if we want $X$ to have generator $\frac{1}{2} \Delta + Z$ we must choose $U$ so that
\[ \mathcal{Z}_t^{-1} Z(X_t) = u_t U_t(t) = U(t). \]
This means we solve the system
\begin{align*}
\partial \mathcal{Z} &= H \mathcal{Z} \partial \mathcal{M} \\
d\mathcal{M} &= dB + \mathcal{Z}_t^{-1} Z(\mathcal{Z}) dt \\
\mathcal{Z}_0 &= \xi \\
\pi \xi &= x.
\end{align*}
(1.1)

Similarly, for $Y$ we want
\[ \mathcal{Z}_t^{-1} Z(Y_t) = v_t V_t(t) = V(t), \]
so we solve the system
\begin{align*}
\partial \mathcal{V} &= H \mathcal{V} \partial \mathcal{N} \\
d\mathcal{N} &= dA + \mathcal{V}_t^{-1} Z(\mathcal{V}) dt \\
dA &= \mathcal{V}_t^{-1} m_{X,Y} dB \\
\mathcal{V}_0 &= \eta \\
\pi \eta &= y.
\end{align*}
(1.2)

We summarize this in

**Proposition 1.** If $X_t = \pi \mathcal{Z}_t$ and $Y_t = \pi \mathcal{V}_t$, where $\mathcal{Z}$ and $\mathcal{V}$ are the solutions to (1.1) and (1.2), respectively, then $X$ and $Y$ are diffusions on $M$ with generator $\frac{1}{2} \Delta + Z$ begun at $x$ and $y$, respectively. Moreover, the pair $(X, Y)$ is a diffusion on $M \times M$ with generator $L$, where
\[ Lf(x, y) = \frac{1}{2} \sum_{i,j=1}^d (\mathcal{Z}u_i + \mathcal{V}v_j)^2 f(x, y) + Z(x) f(x, y) + Z(y) f(x, y). \]
The process $\mathcal{Z}$ is defined for all $t > 0$, the process $Y$ on $[0, \sigma_c(Y))$. 

The process $X$ is defined for all $t \geq 0$, the process $Y$ on $[0, \sigma_c(Y))$. 

580/99/1-9
The proposition allows us to express the Itô expansion for $\rho_M(X_t, Y_t)$, where $\rho_M$ is the distance on $M$ given by the metric $g$. This formula is derived in Kendall [4] in the case $Z \equiv 0$. For nonzero $Z$, one has for $t < \sigma_C(Y) \wedge T(X, Y)$,

$$d\rho_M(X_t, Y_t) = 2 b_t + \frac{1}{2} \sum_{j,i=1}^d (H_{\Xi} \Xi u^j_t + H_{\Psi} \Psi v^j_t)^2 \rho_M(\pi \Xi, \pi \Psi) \, dt$$

$$+ (Z(X_t) \rho_M(X_t, Y_t) + Z(Y_t) \rho_M(X_t, Y_t)) \, dt,$$  \hspace{1cm} (1.3)

where $b$ is $BM(\mathbb{R})$. The second derivative term remains invariant under a change of the bases $\{u^i\}$ and $\{v^i\}$. Thus assume $\Xi u^i = -\gamma'(0)$, $\Psi v^i = \gamma'(\rho_M(x, y))$ and for $i = 2, ..., d$, $v^i = \Psi^{-1} m_{XY} \Xi u^i$. Then the term $2 b_t$ becomes clear as

$$\Xi u^1 \rho_M(X, Y) = 1, \quad \Psi v^1 \rho_M(X, Y) = 1$$

$$\Xi u^i \rho_M(X, Y) = \Psi v^i \rho_M(X, Y) = 0, \quad i = 2, ..., d,$$

and

$$B^1 = \langle u^1, B \rangle = \langle v^1, A \rangle = A^1.$$

This choice of $\{u^i\}$ and $\{v^i\}$ also makes clear

$$(\Xi u^1 \Psi u^1) \rho_M(X, Y) = (\Xi u^1)^2 \rho_M(X, Y) = (\Psi v^1)^2 \rho_M(X, Y) = (\Psi v^1 \Xi u^1) \rho_M(X, Y) = 0.$$

Making use of the second variation of arclength (again the reader is referred to Kendall [4] or Cheeger and Ebin [11] for details) it follows that

$$\frac{1}{2} \sum_{i,j=2}^d (H_{\Xi} \Xi u^j_t + H_{\Psi} \Psi v^j_t)^2 \rho_M(\pi \Xi, \pi \Psi) = \int_{X_t}^Y \sum_{i=2}^d (|\nabla^T W^i|^2 - \langle R(W^i, T), T, W^i \rangle),$$

\hspace{1cm} (1.4)

where the $W^i, i = 2, ..., d,$ are Jacobi fields along $\gamma$ satisfying $W'(0) = \Xi u^i$, $W'(\rho_M(X, Y)) = \Psi v^i$, $T$ is the tangent vector to $\gamma$, and the integral is along the geodesic $\gamma$. Finally, $R$ is the Riemannian curvature tensor.

Using the first variation of arclength (see Cheeger and Ebin [11, p. 4]) or common sense,

$$Z(X_t) \rho_M(X_t, Y_t) + Z(Y_t) \rho_M(X_t, Y_t) = \langle Z(Y_t), T \rangle - \langle Z(X_t), T \rangle.$$  \hspace{1cm} (1.5)
Using (1.4) and (1.5) we may write for \( t < \sigma_c(Y) \),
\[
d\rho_M(X_t, Y_t) = 2 \, dB_t + \left[ \int_{X_t} \sum_{i=2}^{d} \left( |\nabla_T W^i|^2 - \langle R(W^i, T), T, W^i \rangle \right) \right] \, dt \\
+ \left[ \langle Z(Y_t), T \rangle - \langle Z(X_t), T \rangle \right] \, dt.
\] (1.6)

We now wish to prove

**Theorem 3.** The process \( Y \) may be continued past \( \sigma_c(Y) \) in such a way that \( Y \) is a diffusion on \( M \) with generator \( \frac{1}{2} \Delta + Z \). Furthermore, there is an increasing process \( L \) which increases only on the set of times \( \{ t : Y_t \in C(X_t) \} \) such that for all \( t < T(X, Y) \)
\[
d\rho_M(X_t, Y_t) = 2 \, dB_t + \left[ \int_{X_t} \sum_{i=2}^{d} \left( |\nabla_T W^i|^2 - \langle R(W^i, T), T, W^i \rangle \right) \right] \, dt \\
+ \left[ \langle Z(Y_t), T \rangle - \langle Z(X_t), T \rangle \right] \, dt - dL_t,
\] (1.7)

where the coefficients of \( dt \) are set equal to zero on the support of \( L \).

**Proof.** If \((X, Y)\), instead of being coupled, were an independent pair of \( BM(M) \)'s, i.e., \( Z \equiv 0 \), Kendall [4] has shown there is an increasing process \( L \), \( \text{supp} L \subseteq \{ t > 0 : Y_t \in C(X_t) \} \) such that
\[
d\rho_M(X_t, Y_t) = \sqrt{2} \, dw_t + \frac{1}{2} \left[ \Delta_1 \rho_M(X, Y) + \Delta_2 \rho_M(X, Y) \right] \, dt - dL_t,
\]
with \( w \) a \( BM(\mathbb{R}) \) and where \( \Delta_i \) is the Laplace–Beltrami operator applied to the \( i \)th argument, \( i = 1, 2 \). The process \( L \) depends on a nice property of \( A \rho_M(\cdot, x) \) across the cut locus of \( x \) as in Hsu [12] or Yau [10]. It is the local time of \( Y \) on the cut locus of \( X \). This also holds for \( (\frac{1}{2} \Delta + Z) \rho_M(\cdot, x) \), so in the same manner we may conclude that if \((X, Y)\) is an independent pair of diffusions on \( M \times M \), each with generator \( \frac{1}{2} \Delta + Z \), there exists an increasing process \( L \), \( \text{supp} L \subseteq \{ t > 0 : Y_t \in C(X_t) \} \) such that
\[
d\rho_M(X_t, Y_t) = \sqrt{2} \, dw_t + \left[ (\frac{1}{2} \Delta + Z)_1 \rho_M(X, Y) \right. \\
+ \left. (\frac{1}{2} \Delta + Z)_2 \rho_M(X, Y) \right] \, dt - dL_t,
\]
where again the subscripts 1 and 2 indicate the variable on which the operator acts.

We piece together portions of “coupled” and independent diffusions \((X, Y)\) and pass to a limit to obtain our theorem.

Take a process \( X \) satisfying (1.1) running for all \( t \geq 0 \). Select \( \delta > 0 \). Start with \( y \notin C(x) \) and create the coupled pair, denoted \((X, Y^\delta)\). Note that \( \sigma_c(Y^\delta) \) is a stopping time for \( Y^\delta \). At the time \( \sigma_c(Y^\delta) \), uncouple \( X \) and \( Y^\delta \)
and let $Y^\delta$ evolve independently of $X$ as a diffusion with generator $\frac{1}{2} A + Z$. This we allow until

$$\tau_1 = \inf\{ t > \sigma_C(Y^\delta); \rho_M(Y^\delta_{\tau_1}, Y^\delta_{\sigma_C}) = \delta \}.$$ 

Since $\sigma_C$ is a stopping time for $Y^\delta$, the process $\{ Y^\delta_t : 0 \leq t < \tau_1 \}$ is a diffusion with the correct generator. The set $C = \{ (x, y) : y \in C(x) \}$, by Fubini's Theorem, has zero measure with respect to the Riemannian volume element on $M \times M$. As $\tau_1$ is independent of $X$, this implies that a.s. $(X_{\tau_1}, Y^\delta_{\tau_1}) \notin C$ and the Kendall coupling may be resumed. That is, it may be resumed until

$$\sigma_2 = \sigma_c \circ \tau_1.$$ 

After $\sigma_2$, again run $Y^\delta$ independent of $X$ until

$$\tau_2 = \tau_1 \circ \sigma_2.$$ 

In this way we obtain a sequence of stopping times for $Y^\delta$,

$$0 = \tau_0 < \tau_1 \equiv \sigma_c < \tau_1 < \sigma_2 < \tau_2 < \cdots$$

such that

$$(X_t, Y^\delta_t)$$ is the Kendall coupling when $t \in I(\delta) = \bigcup_{j=0}^\infty [\tau_j, \sigma_j)$$

$$(X_t, Y^\delta_t)$$ are independent diffusions with generator $\frac{1}{2} A + Z$ when $t \in I(\delta)^c$,

with the additional stipulation that the construction stops at the time

$$T(X, Y^\delta) = \inf\{ t > 0 : X_t = Y^\delta_t \}$$

and we set $Y^\delta_t = X_t$ for $t > T(X, Y^\delta)$. Now $Y^\delta$ is the solution of

$$\partial Y^\delta_t = 1_{\partial(\delta)}(t) \Psi_t \partial N_t + 1_{\partial Y^\delta}(t) \partial Y^\delta_t.$$ 

Also, by Kendall [4], there is an $\tilde{w}$, $BM(\mathbb{R})$, and an increasing process $L^\delta$ such that

$$d\rho_M(X_t, Y^\delta_t) = 21_{\partial(\delta)}(t) dw_t$$

$$+ \left[ 1_{\partial Y^\delta}(t) \sum_{i=2}^{d} \left( |\nabla_T W^i|^2 - \langle R(W^t, T) T W^t \rangle \right) \right] dt$$

$$+ 1_{\partial(\delta)}(t) \left[ \langle Z(Y^\delta_t), T \rangle - \langle Z(X_t), T \rangle \right] dt$$

$$+ \sqrt{2} 1_{\partial(\delta)}(t) d\tilde{w}_t + 1_{\partial(\delta)}(t) \left[ \left\{ \frac{1}{2} A + Z \right\}_1 \rho_M(X, Y^\delta) + \left\{ \frac{1}{2} A + Z \right\}_2 \rho_M(X, Y^\delta) \right] dt - dL^\delta_t.$$
Here $L_\delta^t$ is the “local time” of $Y_\delta^t$ on $C(X_i)$ which is an increasing functional supported on the times $\{t: Y_\delta^t \in C(X_i)\} \subset I(\delta)$.

Define $B(x, 1) = \{z \in M: \rho_M(x, z) < 1\}$ and $\tau_{B(x, 1)}(X) = \inf\{t > 0: X_t \notin B(x, 1)\}$ with a similar definition for $\tau_{B(y, 1)}(Y_\delta^t)$ and set $\sigma^n = \inf\{t > 0: \rho_M(X_i, Y_\delta^t) < 1/n\}$ and

$$\tau^n = \tau_{B(x, 1)}(X) \land \tau_{B(y, 1)}(Y_\delta^t) \land \sigma^n.$$  

Since $C$ has zero volume,

$$\left| \{t < \tau^n: \text{dist}(Y_\delta^t, C(X_i)) < \delta\} \right| \to 0 \quad \text{as} \quad \delta \to 0.$$  

Also, using local bounds on the geometry of $M$,

$$|I(\delta) \cap [0, \tau^n]| \to 0 \quad \text{a.s. as} \quad \delta \to 0.$$  

In addition, on $B(x, 1) \times B(y, 1)$, using well-known Laplacian bounds (see Greene and Wu [1]) there exist constants $c_1^n$ and $c_2^n$ such that for $t < \tau^n$,

$$c_1^n \leq 1_{H(\delta)}(t) \cdot \left[ (\frac{1}{2} A + Z_1) \rho_M(X_i, Y_\delta^t) + (\frac{1}{2} A + Z_2) \rho_M(X_i, Y_\delta^t) \right] \leq c_2^n.$$  

These bounds imply that in the topology of uniform convergence on $[0, \tau^n]$,

$$Y_\delta^t \to Y \quad \text{a.s. as} \quad \delta \to 0,$$  

where

$$\partial Y_i = \Psi_i \partial N_i,$$  

with $\Psi_i$, taken as in the Kendall coupling unless $Y_i \in C(X_i)$. Thus $\rho_M(X_i, Y_\delta^t) \to \rho_M(X_i, Y_t)$ and consequently $L_\delta \to L$, and for $t < \tau^n$

$$d\rho_M(X_i, Y_t) = 2 db_t + \left[ \int_{X_i} \sum_{i=2}^d \left( \langle \nabla T W^i \rangle^2 - \langle R(W^i, T), T, W^i \rangle \right) dt \right.$$  

$$+ \left[ \langle Z(Y_t), T \rangle - \langle Z(X_i), T \rangle \right] dt - dL_i,$$

where the integrand in the $dt$ term is taken to be 0 when $Y_t \in C(X_i)$. The increasing process $L_i$ being the limit of the process $L_\delta$ increases only on the set $\{t: Y_t \in C(X_i)\}$. This process can be continued past $\tau^n$ indefinitely by a patching-together argument.

\section{Proof of Theorem 1}

The idea is that Brownian particles leave their starting points faster with greater negative Ricci curvature. Similarly, the distance between the
Kendall couple grows faster with greater negative Ricci curvature. Finally, any complication in the topology leads to a decreasing term, \(-L_i\), in the Itô expansion of \(\rho_M(X_i, Y_i)\), which implies that the distance between the couple grows faster on cut locus free manifolds. These observations lead to the conjecture that over all manifolds \(M\) with \(\text{Ric}_M \geq -Kg\), the quantity \(P^{(X,Y)}(T(X, Y) = \infty)\) or \(P^{(x,y)}(T(X, Y) > \tau_{B(x,\delta)}(X) \wedge \tau_{B(y,\delta)}(Y))\) is maximized on the space form (hyperbolic space) with sectional curvature \(-K/(d-1)\). A direct comparison is not made as this would involve a two-dimensional comparison of stochastic processes which in the present case is not valid. However, this conjecture motivates the proof. Take \((N, g_N)\) to be the \(d\)-dimensional hyperbolic space (complete, simply connected) with constant sectional curvature equal to \(-K/(d-1)\). Then \(\text{Ric}_N = -Kg_N\). Let \(W^i\) be the Jacobi fields in formula (1.7). Take \(\{e_M^i(t)\}_{i=1}^d\) to be an orthonormal frame field which is parallel translated along the unique geodesic from \(X_i\) to \(Y_i\) (assume \(Y_i \notin C(X_i)\)) with \(e_M^i(0) = \Xi u^i, e_M^i(\rho) = \Psi v^i\), where \(\rho = \rho_M(X_i, Y_i)\). In \(N\) let \(\tilde{\gamma}\) be a geodesic of length \(\rho\) from some \(v\) to some \(w\) in \(N\). Let \(\{e_N^i(t)\}_{i=1}^d\) be an orthonormal frame field parallel translated along \(\tilde{\gamma}\), the geodesic in \(N\) from \(v\) to \(w\). For \(i = 2, \ldots, d\) set \(W_N^i(s) = f^i(s) e_N^i(s)\), where

\[
f^i(s) = \cosh(s \sqrt{K/(d-1)}) \\
+ \frac{1 - \cosh(\rho \sqrt{K/(d-1)})}{\sinh(\rho \sqrt{K/(d-1)})} \sinh(s \sqrt{K/(d-1)}).
\]

Then the \(W_N^i\) are Jacobi fields along \(\tilde{\gamma}\) in \(N\). By the index lemma applied in \(M\) to the vector fields \(V^i(s) = f^i(s) e_M^i(s)\) it follows that

\[
\int_{X_i}^{Y_i} \sum_{i=2}^{d} (|\nabla_T W^i|^2 - \langle R(W^i, T) T, W^i \rangle)
\]

\[
\leq \int_{X_i}^{Y_i} \sum_{i=2}^{d} (|\nabla_T V^i|^2 - \langle R(V^i, T) T, V^i \rangle)
\]

\[
= \int_{X_i}^{Y_i} \sum_{i=2}^{d} ((f^i(s))^2 - \langle R(V^i, T) T, V^i \rangle)
\]

(and switching to \(N\) by simple comparison of Ricci curvature and integrating along \(\tilde{\gamma}\))

\[
\leq \int_{V} \sum_{i=2}^{d} ((f^i(s))^2 - \langle R_N(W^i_N, T) T, W^i_N \rangle)
\]

(and using integration by parts and the Jacobi equation)

\[
= \sum_{i=2}^{d} (\langle W_N^i(\rho), W_N^i(\rho) \rangle - \langle W_N^i(0), W_N^i(0) \rangle).
\]
But $W''(s) = f''(s) e''(s)$ so this last term is
\[ 2 \sqrt{K(d-1)} \frac{[\cosh(\rho \sqrt{K/(d-1)}) - 1]}{\sinh(\rho \sqrt{K/(d-1)})}. \]

Consequently, if we use the same Brownian motion, $b$, appearing in the formula (1.7) for $\rho_M(X_t, Y_t)$ and define $\rho_t$ by
\[ d\rho_t = 2 db_t + (2 \sqrt{K(d-1)} + 2m) \, dt \]
with initial condition $\rho_0 = \rho_M(x, y)$, then by a comparison theorem (see Malliavin [8] or Ikeda and Watanabe [3]) and noting that
\[ \frac{2 \sqrt{K(d-1)} \frac{[\cosh(\rho \sqrt{K/(d-1)}) - 1]}{\sinh(\rho \sqrt{K/(d-1)})}} \leq 2 \sqrt{K(d-1)}, \]
and
\[ |\langle Z(Y_t), T \rangle - \langle Z(X_t), T \rangle| \leq 2m \]
it follows that
\[ \rho_M(X_t, Y_t) \leq \rho_t \quad \text{for all} \quad t > 0 \ \text{a.s.} \]
(Recall that the integrands in the $dt$ terms are 0 if $Y_t \in C(X_t)$.) Now if
\[ \sigma_a(\rho_M) \equiv \inf\{t > 0 : \rho_M(X_t, Y_t) = a\}, \]
then
\[ T(X, Y) = \sigma_0(\rho_M) \]
so that
\[ T(X, Y) \leq \sigma_0(\rho) \]
and
\[ \sigma_2(\rho_M) \geq \sigma_2(\rho) \quad \text{a.s. by the above comparison.} \]
Thus
\[ P^{(x, y)}(T(X, Y) = \infty) \leq P^{\rho_0}(\sigma_0(\rho) = \infty). \quad (2.1) \]
Also

\[ P^{(x,y)}(T(X, Y) > \tau_{B(x, \delta)}(X) \wedge \tau_{B(y, \delta)}(Y)) \]

\[ = P^{(x,y)}(T(X, Y) > \tau_{B(x, \delta)}(X) \wedge \tau_{B(y, \delta)}(Y) \wedge \sigma_{2\delta}(\rho_M)) \]

\[ \leq P^{(x,y)}(T(X, Y) > \tau_{B(x, \delta)}(X) \wedge \sigma_{2\delta}(\rho_M)) \]

\[ + P^{(x,y)}(T(X, Y) > \tau_{B(y, \delta)}(Y) \wedge \sigma_{2\delta}(\rho_M)) \]

\[ \leq P^{(x,y)}(\sigma_0(\rho) \geq \sigma_{\delta}(\rho_M(X, x)) \wedge \sigma_{2\delta}(\rho)) \]

\[ + P^{(x,y)}(\sigma_0(\rho) \geq \sigma_{\delta}(\rho_M(Y, x)) \wedge \sigma_{2\delta}(\rho)). \]  

(2.2)

Focusing first on inequality (2.1) define

\[ u(\rho_0) = P^{(0,\infty)}(\sigma_0(\rho) = \infty). \]

Then \( u \) satisfies the differential equation

\[ \frac{1}{2} u'' + (\sqrt{K(d-1) + m}) u' = 0 \]

\[ u(0) = 0, \quad u(\infty) = 1. \]

Thus

\[ u(\rho) = 1 - \exp \{-2(\sqrt{K(d-1) + m}) \rho \}, \]

which implies that for \( \rho_0 \) small

\[ u(\rho_0) \leq 2(\sqrt{K(d-1) + m}) \rho_0. \]

This proves (0.3) and (0.4) when \( m = 0. \)

For (0.2), it suffices to handle one term on the extreme right-hand side of (2.2), the other being similar.

According to Kendall [4], there is a local time term \( L^t \), i.e., an increasing process supported on \( C(x) \) such that

\[ dp_{\rho_M}(X_t, x) = dw_t + (\frac{1}{2} A + Z) \rho_M(X_t, x) dt - dL^t, \]

(2.3)

where \( \{w_t : t \geq 0\} \) is \( BM(\mathbb{R}^1) \). Since \( X_t \) never returns to \( x \) and the gradient of \( \rho_M \) has norm one, we have

\[ |Zp_{\rho_M}(X_t, x)| \leq |Z(X_t)| \leq m \quad \text{a.s.} \]

Similarly, if \( X_t \notin C(x) \), by the Laplacian comparison theorem [1]

\[ \frac{1}{2} \Delta \rho_M(X_t, x) \leq \frac{1}{2} \sqrt{K(d-1) \coth(\rho_M(X_t, x) \sqrt{K/(d-1)}).} \]
We take \((\frac{1}{2} \Delta + Z) \rho_M(X_t, x) = 0\) for \(X_t \in C(x)\). Thus, using the same \(BM(\mathbb{R}^1), w\), appearing in (2.3) to define \(\eta_t\) by

\[
d\eta_t = dw_t + (\frac{1}{2} \sqrt{K(d-1)} \coth(\eta_t) \sqrt{K/(d-1)}) + m) dt
\]

\[
\eta_0 = 0
\]

by comparison again, one has

\[\rho_M(X_t, x) \leq \eta_t, \quad \text{for all } t \geq 0 \text{ a.s.}\]

(As before, the \(dt\) coefficients for \(\rho_M(X_t, x)\) are set equal to 0 if \(X_t \in C(x)\).)

Thus

\[
\tau_{B(\infty, \delta)}(X) = \sigma_\delta(\rho_M(X, x)) \geq \sigma_\delta(\eta) \quad \text{a.s.} \quad (2.4)
\]

Using (2.4) in the first term on the far right-hand side of (2.2), it appears that

\[
P^{(x, y)}(\sigma_\delta(\rho) \geq \sigma_\delta(\rho_M(X, x)) \wedge \sigma_\delta(\rho)) \leq P^{(\rho_0, 0)}(\sigma_\delta(\rho) > \sigma_\delta(\eta) \wedge \sigma_\delta(\rho)). \quad (2.5)
\]

The superscript on the probability on the right means \(\rho_0\) is the starting point for \(\rho, 0\) for \(\eta\). Set \(B_1 = (2 \sqrt{K(d-1)} + m), \quad B_2(\eta) = (\frac{1}{2} \sqrt{K(d-1)} \coth(\eta) \sqrt{K/(d-1)}) + m\), from Ito's formula, for a nice function \(h, P^{(\rho_0, 0)} \text{ a.s.}\)

\[
h(\rho_t, \eta_t) = h(\rho_0, 0) + \int_0^\tau \nabla h(\rho_s, \eta_s) d(b_s, \beta_s)
\]

\[
+ \int_0^\tau (2h_{\rho \rho} + h_{\eta \eta} + B_1 h_\rho + B_2(\eta) h_\eta)(\rho_s, \eta_s) ds
\]

\[
+ \int_0^\tau 2h_{\rho \eta}(\rho_s, \eta_s) d\langle b, \beta \rangle.
\]

The right-hand side of (2.5) may be viewed as \(E^{(\rho_0, 0)} 1_A(\rho_t, \eta_t)\), where \(\tau = \inf\{t > 0 : \rho_t \notin (0, 2\delta)\} \text{ or } \eta_t \notin [0, \delta]\} \text{ and } A = \{(2\delta, \eta) : 0 \leq \eta \leq \delta\} \cup \{(\rho, \delta) : 0 < \rho < 2\delta\}. \text{ Thus, } E^{(\rho_0, 0)} 1_A(\rho_t, \eta_t) \text{ may be estimated by selecting a nice } h \geq 0 \text{ for which } h|_A \geq 1_A \text{ and evaluating } E^{(\rho_0, 0)} h(\rho_t, \eta_t).

To this end define

\[
h(\rho, \eta) = \begin{cases} 1 & \text{on } A \\ \rho/2\delta & \text{on } \{(\rho, \eta) : 0 < \rho < 2\delta, 0 < \eta < \delta/2\} \end{cases}
\]
and then extend $h$ so that

$$|h_\eta| \vee |h_\rho| \vee \delta |h_{\rho\rho}| \vee \delta |h_{\eta\eta}| \vee \delta |h_{\rho\delta}| \leq \frac{c}{\delta}. \quad (2.7)$$

Then, using (2.6) and (2.7), since $E^{\rho_0,0}_{\tau} \leq c_\rho_0 \delta$, it follows that

$$P^{\rho_0,0}(\sigma_\delta(\rho) > \sigma_\delta(\eta)) \leq h(\rho_0, 0) + \frac{c}{\delta^2} E^{\rho_0,0}_{\tau}$$

$$\leq \left( \frac{c}{\delta} + c(K, d, m) \right) \rho_0, \quad (2.8)$$

and the proof is complete. \hfill \Box

REFERENCES