Z. Wahrscheinlichkeitstheorie verw. Gebiete 32, 1–64 (1975)

The Central Limit Problem for Geodesic Random Walks

Erik Jørgensen

0. Introduction

The purpose of the present work is to consider the problem of defining the concept of a random walk in a general Riemannian manifold $\mathscr{M}$, and to investigate the behavior in the limit of a sequence of such random walks. It will be shown that such a sequence, under reasonable assumptions, converges to a diffusion process in $\mathscr{M}$, and in particular Brownian motion processes will be obtained as limits of sequences of random walks with identically distributed steps.

The results which we arrive at in this paper are general versions of well-known classical results concerning the transition from random walks to diffusion processes, for instance: the central limit theorem and Donsker's theorem (as formulated in: [1] p. 68ff.), and they specialize to these results when we take $\mathscr{M}=E^N=N$-dimensional Euclidean space. Also, what they essentially show is that the combination of a large number of mutually independent, individually negligible, and identically distributed random causes leads to a "normally distributed" random effect, even though the rule according to which these causes are combined is nonlinear, noncommutative, and altogether not in any way derived from a natural group structure.

It might at this point be appropriate to mention, that the article is not concerned with the study of diffusion processes as such. If one wishes to undertake such a study, one does not have to worry about the convergence problem that is investigated here. Instead, one may proceed to define such a process directly in terms of its infinitesimal characteristics, for instance, by letting it be the solution of a stochastic differential equation, or by prescribing its infinitesimal generator. This approach has, for example, been taken in [6, 10–12] and [16]. In this connection, it should also be noted that Gangolli ([7]) by generalizing an idea contained in [15] has shown how one may construct a diffusion process in a differentiable manifold by injecting, via the exponential map, the differentials of corresponding diffusion processes defined on the tangent spaces of the manifold. Although Gangolli's work apparently has much in common with ours, the similarity is only superficial as one will easily discover by comparing the results obtained in the two papers in the case where the underlying manifold is E^N.

Our approach to the problem in question is based upon the use of semigroup methods similar to the ones used in the papers: [9, 21] and [23]. It may briefly be outlined thusly: Starting with a family of subprobability measures defined on the tangent spaces to the manifold $\mathscr{M}$ we first construct a random walk and next an associated Markov process in $\mathscr{M}$. Then considering a sequence of such families we set up conditions on basis of which we prove that the corresponding sequence of Markov processes converges weakly to a diffusion process in $\mathscr{M}$. All of this is formulated in the language of semigroup theory and thus the basic convergence problem becomes the problem of deciding whether a given sequence of semigroups is convergent. In the present context this problem is rather complicated, and its solution takes up a major part of the article.

The article itself is divided into three main sections and an appendix. Section 1 starts out with a brief survey of our notation, which we have tried to keep as close as possible to that of our main reference, namely, the books: [6] vol. I and II. After that follows a theorem by Stone concerning the weak convergence of a sequence of Markov processes. This theorem is translated into a corresponding theorem about the convergence of a sequence of semigroups and is from then on used as the basic test for weak convergence. Finally, we state a lemma which links the convergence of a sequence of semigroups to that of the corresponding sequence of infinitesimal generators. The proof of this lemma is based upon ideas contained in the paper: [9].

In Section 2 we begin with a description of the concept we have named "a geodesic random walk." Following that is the proof of our main theorem, which states under which conditions a sequence of such random walks converges to a diffusion process. As mentioned earlier, the proof of this theorem is rather long, so in order to simplify it we have broken it into several smaller parts, each of which has been stated as a lemma. The section closes with a few remarks about the properties of the resulting limit process.

Section 3 is devoted to the study of a special class of processes in $\mathscr{M}$, namely, those we have chosen to name Brownian motions. Here we define a Brownian motion to be a diffusion process that may be obtained as the limit of a sequence of geodesic random walks with identically distributed steps. Such a process may be characterized by the condition that the coefficients occurring in the expression for its differential generator are constant, where the word "constant" is taken to mean "invariant under parallel translations." In particular, we define the Wiener process in $\mathscr{M}$ to be the Brownian motion that occurs as the limit of a sequence of random walks where the individual steps have a symmetric distribution. As one would expect, the differential generator of the Wiener process turns out to be $\frac{1}{2}\nabla^2$, where ∇^2 is the Laplace operator on $\mathscr{M}$.

Now, in case $\mathscr{M}$ is a homogeneous space one usually defines the concept of a Brownian motion in a different fashion. That is, instead of demanding that it be invariant under parallel translations one defines (see [24–26]) a Brownian motion to be a diffusion process that is invariant under isometries of $\mathscr{M}$. The two definitions do not in general lead to the same class of processes, but, nevertheless, the Wiener process is a Brownian motion according to both of them. A brief discussion of this situation has been included. The section finishes with a classification of the different types of Brownian motions that may exist on a given manifold. In particular, it is shown that the Wiener process is the only Brownian motion there is on a two-dimensional complete Riemannian manifold with nonzero curvature. This generalizes a corresponding result obtained by Yosida ([24]) for the case $\mathscr{M}=S^2$.

Finally, the appendix contains a listing of those concepts from differential geometry which are used in the three main sections, together with a few non-standard notational conventions. Also our two basic axioms, referred to as $\mathscr{A}$ and $\mathscr{B}$ in the main text, have been listed here.

1. Basic Concepts

1.1. Throughout this paper $\mathscr{M}$ will denote a fixed N-dimensional complete Riemannian manifold[1] with metric tensor g and corresponding distance function $d(\cdot,\cdot)$. In case $\mathscr{M}$ is not compact let $\mathscr{M}_\Delta = \mathscr{M} \cup \{\Delta\}$ be the one-point compactification of $\mathscr{M}$, and define a metric $\delta(\cdot,\cdot)$ on $\mathscr{M}_\Delta$ by setting

$$\delta(p,q) = \sum_{n=1}^{\infty} 2^{-n} |f_n(p) - f_n(q)|, \qquad p, q \in \mathscr{M}_\Delta,$$

where

$$f_n(q) = \begin{cases} 1 & \text{if } q = \Delta \\ \dfrac{d(p_n, q)}{1 + d(p_n, q)} & \text{if } q \in \mathscr{M}, \end{cases}$$

and the sequence $\{p_n\}_{n=1}^{\infty}$ is dense in $\mathscr{M}$. The topology on $\mathscr{M}$ generated by $\delta(\cdot,\cdot)$ is then the same as the topology generated by $d(\cdot,\cdot)$ (see, for instance, [13] p. 144ff.). In case $\mathscr{M}$ is already compact, we simply adjoin Δ as an isolated point and define the metric $\delta(\cdot,\cdot)$ on $\mathscr{M}_\Delta = \mathscr{M} \cup \{\Delta\}$ by letting $d = \delta$ on $\mathscr{M} \times \mathscr{M}$ and

$$\delta(p, \Delta) = 1 \qquad \text{for } p \in \mathscr{M}.$$

In both cases we then have

Lemma 1.1. *The restriction of $\delta(\cdot,\cdot)$ to $\mathscr{M} \times \mathscr{M}$ uniformly continuous with respect to $d(\cdot,\cdot)$.*

Define the following classes of real-valued functions on $\mathscr{M}$

$$C = C(\mathscr{M}) = \{f : f \text{ is bounded and continuous}\},$$

$$\tilde{C} = \tilde{C}(\mathscr{M}) = \{f : f \in C \text{ and } \lim_{p \to \infty} f(p) \text{ exists}\},$$

$$C_0 = C_0(\mathscr{M}) = \{f : f \in \tilde{C} \text{ and } \lim_{p \to \infty} f(p) = 0\},$$

$$C_K = C_K(\mathscr{M}) = \{f : f \in C \text{ and } \operatorname{supp}(f) \text{ is compact}\},$$

where $\operatorname{supp}(f)$ denotes the support of f, while

$$\lim_{p \to \infty} f(p) = a$$

means: For every $\varepsilon > 0$ there exists a compact $K \subset \mathscr{M}$ such that

$$|f(p) - a| < \varepsilon \qquad \text{for } p \in \mathscr{M} \setminus K = K^c.$$

Clearly $C = \tilde{C} = C_0 = C_K$ if $\mathscr{M}$ is compact.

[1] See the appendix for notation and references.

For $f \in C$ set

$$\|f\| = \sup_{p \in \mathcal{M}} |f(p)|.$$

This defines a norm relative to which each of the spaces C, $\tilde{C}$, C_0 is a (real) Banach space. Note that C_K is dense in C_0, and that $\tilde{C}$ consists of exactly those functions in C which may be continuously extended to all of $\mathcal{M}_\Delta$. This we make use of occasionally by identifying $\tilde{C}(\mathcal{M})$ and $C(\mathcal{M}_\Delta)$, in an obvious fashion, when $\mathcal{M}$ is not compact.

Finally, whenever the term "measurable" is used it will refer to the basic σ-algebra $\mathcal{B}(\mathcal{B}_\Delta)$ of Borel sets in $\mathcal{M}(\mathcal{M}_\Delta)$. The class of bounded Borel measurable functions will be denoted $B(B_\Delta)$.

1.2. Let $(X_t, t \geqq 0)$ be a stationary Markov process[2] with state space $(\mathcal{M}, \mathcal{B})$ and transition function

$$P(t, p, \Gamma) = \operatorname{Prob}\{X_{t+s} \in \Gamma \mid X_s = p\},$$

where $0 \leqq t < \infty$, $p \in \mathcal{M}$, and $\Gamma \in \mathcal{B}$. The semigroup $(T_t, t \geqq 0)$ associated with $(X_t, t \geqq 0)$ is then defined by

$$T_t f(p) = \int_{\mathcal{M}} f(q)\, P(t, p, dq), \tag{1.1}$$

for $0 \leqq t$ and $f \in B$. $(T_t, t \geqq 0)$ is a positive contraction ("sup" norm!) semigroup on B. If also

$$f \in C_0 \Rightarrow T_t f \in C_0 \quad \text{for } t \geqq 0, \text{ and } \lim_{t \downarrow 0} \|T_t f - f\| = 0, \tag{1.2}$$

then $(T_t, t \geqq 0)$ will be called a C_0-semigroup. We shall be working almost exclusively with such semigroups, and unless otherwise specified we always take the domain to be C_0 rather than B. Similarly, we take for the domain of A, the infinitesimal generator, the dense subset $\mathcal{D}$ of C_0 on which it is defined, i.e., if "s-lim" denotes "limit relative to the norm $\|\cdot\|$" then

$$\mathcal{D} = \left\{ f \colon f \in C_0 \text{ and } Af = s\text{-}\lim_{t \downarrow 0} \frac{1}{t} [T_t f - f] \text{ exists} \right\}.$$

Recall, that the operator A is linear, closed, and in general unbounded, and that moreover

$$f \in \mathcal{D}, \quad t \geqq 0 \Rightarrow T_t f \in \mathcal{D} \quad \text{and} \quad A T_t f = T_t A f = \frac{d}{dt} T_t f \tag{1.3}$$

$$T_t f - f = \int_0^t A T_s f\, ds,$$

the differentiation and integration being performed in the Banach space C_0. Because the weak and the strong generators coincide for a C_0-semigroup ([6] p. 77, Lemma 2.11), we can also assert that a function $f \in C_0$ is contained in $\mathcal{D}$ if only it satisfies

$$\begin{aligned} &\|T_t f - f\| \leqq t \cdot \text{constant} < \infty, \text{ for } t \geqq 0, \text{ and} \\ &\lim_{t \downarrow 0} \frac{1}{t} (T_t f(p) - f(p)) = g(p) \text{ exists for each } p \in \mathcal{M}, \\ &\text{and } g \text{ is contained in } C_0. \end{aligned} \tag{1.4}$$

[2] For precise definitions see for instance [2] or [6].

Finally, for $\lambda>0$, the map $(\lambda-A)\colon \mathscr{D}\to C_0$ is a bijection from $\mathscr{D}$ to C_0, whose inverse R_λ, the resolvent of A (or of $(T_t, t\geqq 0)$), is given by the formula

$$R_\lambda g=\int_0^\infty e^{-\lambda s}\, T_s g\, ds, \qquad g\in C_0. \tag{1.5}$$

R_λ is linear and bounded with norm $\|R_\lambda\|\leqq\lambda^{-1}$, moreover

$$s\text{-}\lim_{\lambda\to\infty} R_\lambda g=g, \qquad \text{for } g\in C_0.$$

On the other hand, if we are originally given a C_0-semigroup $(T_t, t\geqq 0)$ then we may assume that it comes from some "nice" Markov process $(X_t, t\geqq 0)$. Namely, from $(T_t, t\geqq 0)$ we can first construct a transition function $P(t,p,\Gamma)$ on $(\mathscr{M},\mathscr{B})$, and then extend this function to $(\mathscr{M}_\Delta,\mathscr{B}_\Delta)$ by setting

$$\tilde P(t,p,\Gamma)=\begin{cases}1 & \text{if } p=\Delta \text{ and } \Delta\in\Gamma\\ 0 & \text{if } p=\Delta \text{ and } \Delta\notin\Gamma\\ P(t,p,\Gamma\setminus\{\Delta\})+1-P(t,p,\mathscr{M}) & \text{if } p\neq\Delta \text{ and } \Delta\in\Gamma\\ P(t,p,\Gamma) & \text{if } p\neq\Delta \text{ and } \Delta\notin\Gamma.\end{cases}$$

Then, if we let Ω be the set of functions $\omega(t)$ on $[0,\infty)$ with values in $\mathscr{M}_\Delta$ which are right continuous, have left hand limits, and which moreover satisfy

$$\omega(t_0)=\Delta \Rightarrow \omega(s)=\Delta \qquad \text{for } s\geqq t_0,$$

we may assert[3] that there exists a standard process $(X_t, t\geqq 0)$ with transition function $\tilde P(t,p,\Gamma)$ and sample functions belonging to Ω. The semigroup $(\tilde T_t, t\geqq 0)$ associated with this process is determined by

$$\tilde T_t \tilde f(p)=\begin{cases}\tilde f(\Delta)+T_t f_0(p) & \text{for } p\neq\Delta\\ \tilde f(\Delta) & \text{for } p=\Delta\end{cases} \tag{1.6}$$

when $\tilde f\in\tilde C(\mathscr{M}_\Delta)$ and where $f_0\in C_0$ is the restriction of $\tilde f-\tilde f(\Delta)$ to $\mathscr{M}$.

In case we only want to study the process $(X_t, t\geqq 0)$ up till the time it leaves $\mathscr{M}$ we introduce the lifetime

$$\zeta(\omega)=\begin{cases}\infty \text{ if } X_t(\omega)\in\mathscr{M} \text{ for all } t\geqq 0\\ \inf\{t\colon X_t(\omega)=\Delta\} \text{ otherwise}.\end{cases}$$

Here

$$\operatorname{Prob}\{\zeta>s\,|\,X_0=q\}=P(s,q,\mathscr{M}),$$

and in particular if $P(s,q,\mathscr{M})\equiv 1$, that is if $(T_t, t\geqq 0)$ is conservative, we may assume[4] that $\zeta\equiv\infty$. Also, if the condition ([6] p. 91)

$$\lim_{t\downarrow 0}\frac{1}{t}\sup_{q\in K} P\big(t,q,B^c(\varepsilon,q)\big)=0,$$

[3] [2] p. 46ff., or [6] Chapter III.

[4] [6] p. 87.

where $B(\varepsilon, q)=\{p: p\in\mathcal{M}, d(p,q)<\varepsilon\}$, is satisfied for every $\varepsilon>0$ and every compact subset K of $\mathcal{M}$, then the sample functions of $(X_t, t\geqq 0)$ may be assumed to be continuous on the interval $[0, \zeta)$.

From now on we shall mainly be working with semigroups, but at the same time we shall interpret our results as being results about the corresponding Markov processes. In this context, one of the first problems we must deal with is the following: Given a sequence $\{(\overset{n}{T}_t, t\geqq 0)\}_{n=1}^{\infty}$ of C_0-semigroups and a corresponding sequence of processes $\{(X_t^{(n)}, t\geqq 0)\}_{n=1}^{\infty}$ (on $\mathcal{M}_\Delta$). Which conditions must be imposed upon the sequence $\{(\overset{n}{T}_t, t\geqq 0)\}$ in order to ensure the convergence[5] of the sequence $\{(X_t^{(n)}, t\geqq 0)\}$? To find the answer to this question we take as our point of departure the following theorem by Stone

Proposition 1.1. *The sequence* $\{(X_t^{(n)}, t\geqq 0)\}_{n=1}^{\infty}$ *converges weakly to* $(X_t^{(0)}, t\geqq 0)$ *if*

(i) *the finite dimensional distributions of* $(X_t^{(n)}, t\geqq 0)$ *converge weakly to the finite dimensional distributions of* $(X_t^{(0)}, t\geqq 0)$; *and*

(ii) *for every* $\varepsilon>0$ *and* $L>0$

$$\lim_{\substack{n\to\infty \\ c\to 0}} \operatorname{Prob}\Big\{\sup_{\substack{t-c<t_1<t<t_2<t+c \\ 0\leqq t_1<t<t_2\leqq L}} \big[\min\big[\delta(X_{t_1}^{(n)}, X_t^{(n)});\, \delta(X_t^{(n)}, X_{t_2}^{(n)})\big]\big]>\varepsilon\Big\}=0.$$

Let us assume that there is a $C_0(\mathcal{M})$-semigroup $(\overset{0}{T}_t, t\geqq 0)$ such that

$$s\text{-}\lim_{n\to\infty} \overset{n}{T}_t f=\overset{0}{T}_t f, \qquad \text{for all } f\in C_0(\mathcal{M}) \text{ and } t\geqq 0, \tag{1.7}$$

and let $(X_t^{(0)}, t\geqq 0)$ be associated with $(\overset{0}{T}_t, t\geqq 0)$. Then we have first

Lemma 1.2. *Let the initial distributions*

$$\mu_n(\Gamma)=\operatorname{Prob}\{X_0^{(n)}\in\Gamma\}, \qquad \Gamma\in\mathcal{B}(\mathcal{M}_\Delta)$$

be given and satisfy $\mu_n\to\mu_0$ *weakly as* $n\to\infty$; *then* (1.7) *implies that the finite dimensional distributions of the processes* $(X_t^{(n)}, t\geqq 0)$ *converge weakly to the finite dimensional distributions of the process* $(X_t^{(0)}, t\geqq 0)$.

To prove this we need the simple

Lemma 1.3. *Let* $\{f_n\}_{n=1}^{\infty}$ *be a sequence of functions in* $C(\mathcal{M}_\Delta)$ *converging uniformly to* $f_0\in C(\mathcal{M}_\Delta)$, *and let* $\{\mu_n\}_{n=1}^{\infty}$ *be a sequence of probability measures on* $\mathcal{M}_\Delta$ *converging weakly to the measure* μ_0; *then*

$$\lim_{n\to\infty}\int_{\mathcal{M}_\Delta} f_n(p)\,\mu_n(dp)=\int_{\mathcal{M}_\Delta} f_0(p)\,\mu_0(dp).$$

Proof of Lemma 1.2. Define as in (1.6) $C(\mathcal{M}_\Delta)$-semigroups $(\overset{\tilde n}{T}_t, t\geqq 0)$ by

$$\overset{\tilde n}{T}_t \tilde f(p)=\begin{cases}\tilde f(\Delta)+\overset{n}{T}_t f_0(p) & \text{if } p\in\mathcal{M}\\ \tilde f(\Delta) & \text{if } p=\Delta\end{cases}$$

---

[5] Weak convergence as introduced in [18] and [20]. However, because we are working on a half-open parameter interval, $0\leqq t<\infty$, we find it convenient to use a slightly modified version of the standard theory. This modification has been introduced by Stone, and we refer the reader to his paper [22] for details. Here we shall contend ourselves with listing his convergence criterion as our Proposition 1.1.

where as before $\tilde{f}\in C(\mathcal{M}_\Delta)$ and f_0 is the restriction of $\tilde{f}-\tilde{f}(\Delta)$ to $\mathcal{M}$, then (1.7) implies

$$s\text{-}\lim_{n\to\infty} \overset{\tilde{n}}{T}_t\tilde{f}=\overset{\tilde{0}}{T}_t\tilde{f}, \quad \text{for } \tilde{f}\in C(\mathcal{M}_\Delta) \text{ and all } t\geqq 0. \tag{1.8}$$

We shall prove that for every k-tuple, $0\leqq t_1<t_2<\cdots<t_k$, and every function $g(p_1,\ldots,p_k)\in C(\mathcal{M}_\Delta^k)=C(\mathcal{M}_\Delta\times\cdots\times\mathcal{M}_\Delta)$ is

$$\mathscr{E}^q\{g(X_{t_1}^{(n)},\ldots,X_{t_k}^{(n)})\}=F_n(q)\in C(\mathcal{M}_\Delta)$$

and

$$s\text{-}\lim_{n\to\infty} F_n(q)=F_0(q),$$

where $\mathscr{E}^q\{\cdot\}$ denotes expectation relative to the measure[6] $\mu(\Gamma)=\chi_\Gamma(q)$. This in conjunction with Lemma 1.3 will imply the validity of Lemma 1.2. The proof is carried out by induction. For $k=1$ the statement is just (1.8), so assume the statement is true for $k=k_0\geqq 1$, and let

$$0\leqq s<t_1<\cdots<t_{k_0}, \qquad g(q,p_1,\ldots,p_{k_0})\in C(\mathcal{M}_\Delta^{k_0+1});$$

then

$$\mathscr{E}^p\{g(X_s^{(n)},X_{t_1}^{(n)},\ldots,X_{t_{k_0}}^{(n)})\}=\overset{\tilde{n}}{T}_s G_n(p)$$

with

$$G_n(q)=\mathscr{E}^q\{g(q,X_{t_1-s}^{(n)},\ldots,X_{t_{k_0}-s}^{(n)})\}$$

provided $G_n(\cdot)\in C(\mathcal{M}_\Delta)$. To prove the latter let $\varepsilon>0$ be given and choose $q^1,\ldots,q^r\in\mathcal{M}_\Delta$ with neighborhoods $V_1,\ldots,V_r$ such that

$$\bigcup_{j=1}^r V_j=\mathcal{M}_\Delta,$$

and

$$\sup_{\substack{q\in V_i\\(p_1,\ldots,p_{k_0})\in\mathcal{M}_\Delta}} |g(q,p_1,\ldots,p_{k_0})-g(q^i,p_1,\ldots,p_{k_0})|<\varepsilon.$$

Also let $\{\varphi_j\}_{j=1}^r$ be a partition of unity[7] subordinate to $V_1,\ldots,V_r$ then

$$\begin{aligned}G_n(q)&=\sum_{i=1}^r \mathscr{E}^q\{\varphi_i(q)\,g(q^i,X_{t_1-s}^{(n)},\ldots,X_{t_{k_0}-s}^{(n)})\}\\&\quad+\mathscr{E}^q\left\{\sum_{i=1}^r\varphi_i(q)[g(q,X_{t_1-s}^{(n)},\ldots,X_{t_{k_0}-s}^{(n)})-g(q^i,X_{t_1-s}^{(n)},\ldots,X_{t_{k_0}-s}^{(n)})]\right\}\\&=S_1^{(n)}(q)+S_2^{(n)}(q),\end{aligned}$$

where $S_1^{(n)}(\cdot)\in C(\mathcal{M}_\Delta)$ according to the induction hypothesis, while

$$|S_2^{(n)}(q)|<\mathscr{E}^q\left\{\varepsilon\sum_{i=1}^r\varphi_i(q)\right\}=\varepsilon.$$

Thus $G_n(\cdot)$ can be approximated uniformly by functions in $C(\mathcal{M}_\Delta)$ which implies

$$G_n(\cdot)\in C(\mathcal{M}_\Delta).$$

---

[6] χ_Γ is the characteristic function of the set Γ.

[7] That is: $\varphi_i\in C(\mathcal{M}_\Delta)$, $\operatorname{supp}(\varphi_i)\subset V_i$, $0\leqq\varphi_i\leqq 1$, and $\sum_{i=1}^r\varphi_i\equiv 1$.

Applying the second part of the induction hypothesis on the $S_1^{(n)}(\cdot)$ next, we get

$$\text{s-}\lim_{n\to\infty} S_1^{(n)}(p) = S_1^{(0)}(p),$$

hence

$$\overline{\lim_{n\to\infty}} \sup_{q\in\mathcal{M}_\Delta} |G_n(q) - G_0(q)| \leqq 2\varepsilon,$$

and as the left hand side is independent of ε it must be zero, that is

$$\text{s-}\lim_{n\to\infty} G_n(q) = G_0(q),$$

and then also [8]

$$\text{s-}\lim_{n\to\infty} \overset{\bar{n}}{T}_t G_n(p) = \overset{\bar{0}}{T}_t G_0(p)$$

which completes the induction. Note that what we essentially have managed to do is to show that condition (i) in Proposition 1.1 may be replaced by (1.8), which is easier to verify in a given situation. There is a corresponding simplification of condition (ii) available:

Lemma 1.4. *Assume the hypotheses in Lemma* 1.2 *are satisfied, and that furthermore*[9].

For every $\varepsilon > 0$ *there exists an* $\alpha > 0$ *so*

$$\sup_{p\in\mathcal{M}_\Delta,\ 0<t} \frac{1}{t} \tilde{P}_n(t, p, B_\Delta^c(\varepsilon, p)) \leqq \alpha, \tag{1.9}$$

for n *sufficiently large; then the sequence* $\{(X_t^{(n)}, t \geqq 0)\}_{n=1}^{\infty}$ *converges weakly to* $(X_t^{(0)}, t \geqq 0)$.

Here $B_\Delta(\varepsilon, p) = \{q \colon q \in \mathcal{M}_\Delta, \delta(p, q) < \varepsilon\}$ and B^c is the complement of B.

In view of Proposition 1.1 we only need to show that (1.9) implies (ii). Now the quantity

$$\tilde{P}_n^q\{\sup_{\substack{0\leqq t_1<t_2<t_3\leqq L\\ t_2-t_1<\beta,\ t_3-t_2<\beta}} [\min(\delta(X_{t_1}^{(n)}, X_{t_2}^{(n)}); \delta(X_{t_2}^{(n)}, X_{t_3}^{(n)}))] > \varepsilon\}$$

is majorized by

$$\left(\frac{4L}{\beta}\right)\left[\sup_{\substack{0\leqq t\leqq 2\beta\\ p\in\mathcal{M}_\Delta}} \tilde{P}_n\left(t, p, B_\Delta^c\left(\frac{\varepsilon}{4}, p\right)\right)\right]^2$$

as one may prove by combining the proofs of Lemma 6.4 and Lemma 6.6 in [5] (p. 129 and p. 136 resp.). Also, for given L, β, ε we may choose α so (1.9) holds with $\varepsilon/4$ substituted for ε. Then for n large enough

$$\left(\frac{4L}{\beta}\right)\left[\sup_{\substack{0\leqq t\leqq 2\beta\\ p\in\mathcal{M}_\Delta}} \tilde{P}_n\left(t, p, B_\Delta^c\left(\frac{\varepsilon}{4}, p\right)\right)\right]^2 \leqq \left(\frac{4L}{\beta}\right)[2\beta\alpha]^2 = 16L\alpha^2\beta \to 0, \quad \text{for } \beta\to 0,\ n\to\infty,$$

and this proves the lemma.

[8] $\|\overset{\bar{n}}{T}_t G_n - \overset{\bar{n}}{T}_t G_0\| \leqq \|\overset{\bar{n}}{T}_t(G_n - G_0)\| + \|\overset{\bar{n}}{T}_t G_0 - \overset{\bar{0}}{T}_t G_0\| \leqq \|G_n - G_0\| + \|\overset{\bar{n}}{T}_t G_0 - \overset{\bar{0}}{T}_t G_0\|$.

[9] $\tilde{P}_n(t, p, \Gamma)$ is the transition function associated with $(\overset{\bar{n}}{T}_t, t \geqq 0)$.

1.3. The essential idea behind the following lemma is contained in the paper [9] (see [9] Lemma 4.1).

Lemma 1.5. *Let $\{(\overset{n}{T}_t, t \geqq 0)\}_{n=1}^{\infty}$ be a sequence of $C_0(\mathscr{M})$-semigroups satisfying the conditions:*

(a) *For every $t \geqq 0, f \in C_0$, and $\varepsilon > 0$ there exists a compact subset K_ε of $\mathscr{M}$ such that*

$$|\overset{n}{T}_t f(p)| < \varepsilon \qquad \text{for all } n \text{ if } p \notin K_\varepsilon.$$

(b) *For every $t \geqq 0$ and $f \in C_0$ the family $\{\overset{n}{T}_t f\}_{n=1}^{\infty}$ is equicontinuous.*

(c) *There exists a dense subset $\mathscr{D}$ of C_0 such that*

$$\mathscr{D} \subset \bigcap_{n=1}^{\infty} \mathscr{D}_n \quad \text{and} \quad \underset{n\to\infty}{s\text{-lim}}\, A_n f = Af \qquad \text{exists for } f \in \mathscr{D}.$$

Here A_n with domain $\mathscr{D}_n$ is the infinitesimal generator for $(\overset{n}{T}_t, t \geqq 0)$. Then there exists a subsequence $\{n_k\}_{k=1}^{\infty}$ such that

$$\underset{k\to\infty}{s\text{-lim}}\, \overset{n_k}{T}_t f = T_t f$$

exists for all $t \geqq 0$ and $f \in C_0$. The family of operators $(T_t, t \geqq 0)$ forms a C_0-semigroup whose infinitesimal generator $(\dot{A}, \dot{\mathscr{D}})$ is an extension of $(A, \mathscr{D})$.

Proof. Let $\{f_n\}_{n=1}^{\infty}$ be dense in C_0 and let $\{t_l\}_{l=1}^{\infty}$ be an enumeration of the positive rationals. For t_l, f_j fixed, the sequence $\{\overset{n}{T}_{t_l} f_j\}_{n=1}^{\infty}$ satisfies $\|\overset{n}{T}_{t_l} f_j\| \leqq \|f_j\|$ so because of (b) there exists according to the Ascoli-Arzela theorem a subsequence $\{\overset{n_k}{T}_{t_l} f_j\}_{k=1}^{\infty}$ that converges, uniformly on compact subsets of $\mathscr{M}$, to a continuous function $g_{l,j}$ for which $\|g_{l,j}\| \leqq \|f_j\|$.

Next, (a) implies that the convergence is uniform on all of $\mathscr{M}$; namely, to a given $\varepsilon > 0$ we may choose a compact subset K of $\mathscr{M}$ so

$$|\overset{n}{T}_{t_l} f_j(p)| < \frac{\varepsilon}{2} \qquad \text{for all } n \text{ and } p \notin K;$$

hence, for k sufficiently large

$$|\overset{n_k}{T}_{t_l} f_j(p) - g_{l,j}(p)| < \varepsilon \qquad \text{for } p \in K,$$

and

$$|\overset{n_k}{T}_{t_l} f_j(p) - g_{l,j}(p)| \leqq |\overset{n_k}{T}_{t_l} f_j(p)| + |g_{l,j}(p)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \qquad \text{for } p \notin K,$$

which demonstrates the uniform convergence and also shows $g_{l,j} \in C_0$. It may be assumed that the sequence $\{n_k\}$ is independent of (l, j) (use the diagonal method). Also in order to make the remainder of the proof easier to read we shall assume $\{n_k\} = \{n\}$. Thus, for every (l, j) the limit

$$T_{t_l} f_j = \underset{n\to\infty}{s\text{-lim}}\, \overset{n}{T}_{t_l} f_j$$

exists, and

$$T_{t_l} f_j \in C_0, \qquad \|T_{t_l} f_j\| \leqq \|f_j\|.$$

Now, let $g \in C_0$ and choose $\{f'_k\}_{k=1}^\infty \subset \{f_j\}_{j=1}^\infty$ with $s\text{-}\lim_{k\to\infty} f'_k = g$. We have

$$\|\overset{n}{T}_{t_l} g - \overset{m}{T}_{t_l} g\| \leqq \|\overset{n}{T}_{t_l} g - \overset{n}{T}_{t_l} f'_k\| + \|\overset{n}{T}_{t_l} f'_k - \overset{m}{T}_{t_l} f'_k\| + \|\overset{m}{T}_{t_l} f'_k - \overset{m}{T}_{t_l} g\|$$
$$\leqq 2\|g - f'_k\| + \|\overset{n}{T}_{t_l} f'_k - \overset{m}{T}_{t_l} f'_k\|.$$

From which it follows that $\{\overset{n}{T}_{t_l} g\}_{n=1}^\infty$ is a Cauchy sequence in the Banach space C_0, and thus convergent. We set

$$T_{t_l} g = s\text{-}\lim_{n\to\infty} \overset{n}{T}_{t_l} g, \qquad g \in C_0.$$

Next, if $f \in \mathscr{D}$ and $0 < t_l < t$

$$\|\overset{n}{T}_t f - \overset{m}{T}_t f\| \leqq \|\overset{n}{T}_t f - \overset{n}{T}_{t_l} f\| + \|\overset{n}{T}_{t_l} f - \overset{m}{T}_{t_l} f\| + \|\overset{m}{T}_{t_l} f - \overset{m}{T}_t f\|$$
$$\leqq \|\overset{n}{T}_{t_l} f - \overset{m}{T}_{t_l} f\| + \|\overset{n}{T}_{t-t_l} f - f\| + \|\overset{m}{T}_{t-t_l} f - f\|.$$

Moreover, for $s \geqq 0$

$$\|\overset{k}{T}_s f - f\| = \left\|\int_0^s \frac{d}{dt}(\overset{k}{T}_t f)\, dt\right\| = \left\|\int_0^s \overset{k}{T}_t A_k f\, dt\right\| \leqq \int_0^s \|A_k f\|\, dt \leqq s \cdot \sup_j \|A_j f\| = s \cdot c$$

because $\{A_k f\}_{k=1}^\infty$ converges. Consequently

$$\|\overset{n}{T}_t f - \overset{m}{T}_t f\| \leqq \|\overset{n}{T}_{t_l} f - \overset{m}{T}_{t_l} f\| + 2c \cdot (t - t_l),$$

and it may be concluded that the sequence $\{\overset{n}{T}_t f\}_{n=1}^\infty$ converges in C_0. As before the limit is denoted $T_t f$. Finally, for an arbitrary $g \in C_0$ we can find a sequence $\{f_k\} \subset \mathscr{D}$ converging to g, and so it follows that $T_t g = s\text{-}\lim_{n\to\infty} \overset{n}{T}_t g$ exists.

So far we have obtained a family $(T_t, t \geqq 0)$ of bounded linear maps of C_0 into itself. It must next be shown that they constitute a C_0-semigroup. First, each $(\overset{n}{T}_t, t \geqq 0)$ satisfies the conditions:

$$\overset{n}{T}_0 f = f, \quad \|\overset{n}{T}_t\| \leqq 1, \quad f \geqq 0 \Rightarrow \overset{n}{T}_t f \geqq 0$$

so clearly $(T_t, t \geqq 0)$ also satisfies them. Moreover, we already know

$$f \in C_0 \Rightarrow T_t f \in C_0;$$

so it only remains to verify

$$\text{(a)} \quad T_{t+s} = T_t \circ T_s, \qquad \text{(b)} \quad s\text{-}\lim_{t\downarrow 0} T_t f = f, \quad \text{for } f \in C_0.$$

To prove (b) it suffices to show that the class

$$\mathscr{L} = \{f : f \in C_0 \text{ and } s\text{-}\lim_{t\downarrow 0} T_t f = f\}$$

is closed in C_0 and contains $\mathscr{D}$. So assume

$$\{f_j\} \subset \mathscr{L} \quad \text{and} \quad s\text{-}\lim_{j\to\infty} f_j = f$$

then

$$\|T_t f - f\| \leqq \|T_t(f-f_j)\| + \|T_t f_j - f_j\| + \|f_j - f\| \leqq 2\|f-f_j\| + \|T_t f_j - f_j\|,$$

so

$$\overline{\lim_{t\downarrow 0}} \|T_t f - f\| \leqq 2\|f - f_j\| \to 0 \qquad \text{for } j \to \infty$$

which shows $\mathscr{L}$ is closed. Next, if $g \in \mathscr{D}$

$$\|T_t g - g\| = \lim_{n\to\infty} \|\overset{n}{T}_t g - g\| \leqq t \cdot \sup_n \|A_n g\| = t \cdot c,$$

hence $\|T_t g - g\| \to 0$ for $t \to 0$ so $g \in \mathscr{L}$. This takes care of (b). To prove (a), let $f \in C_0$ and $t, s \geqq 0$ then

$$\begin{aligned}\|T_t(T_s f) - T_{t+s} f\| &\leqq \|T_t T_s f - \overset{n}{T}_t \overset{n}{T}_s f\| + \|\overset{n}{T}_{t+s} f - T_{t+s} f\| \\ &\leqq \|T_t T_s f - \overset{n}{T}_t T_s f\| + \|\overset{n}{T}_t(T_s f - \overset{n}{T}_s f)\| + \|\overset{n}{T}_{n+s} f - T_{t+s} f\| \\ &\leqq \|T_t(T_s f) - \overset{n}{T}_t(T_s f)\| + \|T_s f - \overset{n}{T}_s f\| + \|\overset{n}{T}_{t+s} f - T_{t+s} f\| \xrightarrow[n\to\infty]{} 0.\end{aligned}$$

As the left hand side is independent of n it must be zero and this proves (a).

Let for $\lambda > 0$ R_λ and $\overset{n}{R}_\lambda$ be the resolvents for $(T_t, t \geqq 0)$ and $(\overset{n}{T}_t, t \geqq 0)$ respectively. For $f \in C_0$

$$\|R_\lambda f - \overset{n}{R}_\lambda f\| = \left\| \int_0^\infty e^{-\lambda t}(T_t f - \overset{n}{T}_t f)\, dt \right\| \leqq \int_0^\infty e^{-\lambda t} \varphi_n(t)\, dt$$

where

$$0 \leqq \varphi_n(t) = \|T_t f - \overset{n}{T}_t f\| \leqq 2\|f\|, \qquad \lim_{n\to\infty} \varphi_n(t) = 0,$$

so the "bounded convergence theorem" yields

$$\lim_{n\to\infty} \|R_\lambda f - \overset{n}{R}_\lambda f\| = 0. \tag{1.10}$$

Let $\dot{A}$ with domain $\dot{\mathscr{D}}$ be the infinitesimal generator of $(T_t, t \geqq 0)$. Choose $f \in \mathscr{D}$ and set

$$g_n = f - A_n f$$

then

$$g = s\text{-}\lim_{n\to\infty} g_n = f - Af$$

exists. Also

$$\|f - R_1 g\| = \|\overset{n}{R}_1 g_n - R_1 g\| \leqq \|\overset{n}{R}_1(g_n - g)\| + \|\overset{n}{R}_1 g - R_1 g\| \to 0, \qquad \text{for } n \to \infty,$$

because $\|\overset{n}{R}_1\| \leqq 1$. Thus $f = R_1 g$ which in turn implies

$$f \in \dot{\mathscr{D}} \quad \text{and} \quad f - \dot{A} f = g.$$

So altogether

$$f \in \mathscr{D} \;\Rightarrow\; f \in \dot{\mathscr{D}} \quad \text{and} \quad \dot{A} f = A f,$$

and this completes the proof of Lemma 1.5.

Remark. In order to prove that the sequence $\{(\overset{n}{T}_t, t \geqq 0)\}_{n=1}^\infty$ itself converges it is enough to prove that any two convergent subsequences have the same limit. Now, if $(S_t, t \geqq 0)$ and $(T_t, t \geqq 0)$ are two such limits introduce the corresponding resolvents

$$R_\lambda f = \int_0^\infty e^{-\lambda t} T_t f\, dt, \qquad \overset{0}{R}_\lambda f = \int_0^\infty e^{-\lambda t} S_t f\, dt$$

then the uniqueness theorem for Laplace transforms yields

$$R_\lambda f(p) = \overset{0}{R}_\lambda f(p) \quad \text{for all } \lambda > 0 \Leftrightarrow T_t f(p) = S_t f(p) \text{ for all } t \geqq 0,$$

and so it suffices to show $R_\lambda f = \overset{0}{R}_\lambda f$ for all $\lambda > 0$ and $f \in C_0$ in order to be able to conclude that the semigroups are the same. Define

$$\mathscr{R}^\lambda = \{g : g = (\lambda - A) f, f \in \mathscr{D}\}$$

then an argument similar to the one given in the last part of the proof of Lemma 1.5 yields

$$R_\lambda g = \overset{0}{R}_\lambda g \qquad \text{for all } g \in \mathscr{R}^\lambda.$$

In particular if $\mathscr{R}^\lambda$ is dense in C_0

$$R_\lambda g = \overset{0}{R}_\lambda g \qquad \text{for all } f \in C_0. \tag{1.11}$$

Now, if (1.11) holds for some $\lambda_0 > 0$ then it holds for all $\lambda > 0$. To see this, consider first λ satisfying $0 < \lambda < 2\lambda_0$. According to the resolvent equation we have for $g \in C_0$

$$R_\lambda g - R_{\lambda_0} g = (\lambda_0 - \lambda) R_{\lambda_0} R_\lambda g,$$
$$\overset{0}{R}_\lambda g - \overset{0}{R}_{\lambda_0} g = (\lambda_0 - \lambda) \overset{0}{R}_{\lambda_0} \overset{0}{R}_\lambda g = (\lambda_0 - \lambda) R_{\lambda_0} \overset{0}{R}_\lambda g.$$

So by subtracting the second equation from the first:

$$R_\lambda g - \overset{0}{R}_\lambda g = (\lambda_0 - \lambda) R_{\lambda_0} [R_\lambda g - \overset{0}{R}_\lambda g].$$

Take norms on both sides and use that $\|\lambda_0 R_{\lambda_0}\| \leqq 1$ then

$$\|R_\lambda g - \overset{0}{R}_\lambda g\| \leqq \frac{|\lambda_0 - \lambda|}{\lambda_0} \|R_\lambda g - \overset{0}{R}_\lambda g\|, \qquad |\lambda_0 - \lambda| < \lambda_0$$

which can hold only if $R_\lambda g = \overset{0}{R}_\lambda g$. Thus we have shown that (1.11) holds for $0 < \lambda < 2\lambda_0$ if it holds for $\lambda = \lambda_0$, but then an induction argument shows that it must hold for all $\lambda > 0$. So altogether we may state, that if the assumption: "For some $\lambda_0 > 0$ the set $\mathscr{R}^{\lambda_0}$ is dense in C_0" is added to the assumption in Lemma 1.5 then the sequence $\{(\overset{n}{T}_t, t \geqq 0)\}_{n=1}^\infty$ itself converges.

When we do have convergence, this convergence is in a certain sense uniform on compact t-intervals. Namely, we have

Lemma 1.6. *Let $\{(\overset{n}{T}_t, t \geqq 0)\}_{n=0}^\infty$ be a sequence of C_0-semigroups with generators $\{(A_n, \mathscr{D}_n)\}_{n=0}^\infty$ and which satisfies*

$$s\text{-}\lim_{n\to\infty} \overset{n}{T}_t f = \overset{0}{T}_t f \qquad \textit{for } t \geqq 0 \textit{ and } f \in C_0,$$

$$s\text{-}\lim_{n\to\infty} A_n f = A_0 f \qquad \textit{for } f \in \mathscr{D} \subset \bigcap_{n=0}^\infty \mathscr{D}_n$$

with $\mathscr{D}$ dense in C_0. Then for $t_0>0, f\in C_0$

$$\lim_{n\to\infty} \sup_{0\leqq t\leqq t_0} \|\overset{n}{T}_t f-\overset{0}{T}_t f\|=0.$$

Proof. Let $f\in C_0$, $t_0>0$, and $\varepsilon>0$ be given.

Choose $f_\varepsilon\in\mathscr{D}$ so

$$\|f-f_\varepsilon\|<\frac{\varepsilon}{3},$$

and note that

$$\sup_n \|A_n f_\varepsilon\|=L<\infty$$

because the sequence $\{A_n f_\varepsilon\}_{n=1}^\infty$ converges to $A_0 f_\varepsilon$. Next, select points

$$0=s_0<s_1<\cdots<s_k=t_0$$

such that

$$\max_{1\leqq i\leqq k}(s_i-s_{i-1})<\frac{\varepsilon}{12L}$$

and then choose n_ε so

$$\max_{0\leqq i\leqq k}\|\overset{0}{T}_{s_i} f_\varepsilon-\overset{n}{T}_{s_i} f_\varepsilon\|<\frac{\varepsilon}{6}, \qquad \text{for } n\geqq n_\varepsilon.$$

Now, if $t\in[0,t_0]$ then $s_i\leqq t\leqq s_{i+1}$ for some i, thus

$$\begin{aligned}
\|\overset{0}{T}_t f-\overset{n}{T}_t f\| &\leqq \|\overset{0}{T}_t f-\overset{0}{T}_t f_\varepsilon\|+\|\overset{0}{T}_t f_\varepsilon-\overset{n}{T}_t f_\varepsilon\|+\|\overset{n}{T}_t f_\varepsilon-\overset{n}{T}_t f\|\\
&\leqq 2\|f-f_\varepsilon\|+\|\overset{0}{T}_t f_\varepsilon-\overset{n}{T}_t f_\varepsilon\|<\frac{2\varepsilon}{3}+\|\overset{0}{T}_t f_\varepsilon-\overset{n}{T}_t f_\varepsilon\|\\
&\leqq \frac{2\varepsilon}{3}+\|\overset{0}{T}_t f_\varepsilon-\overset{0}{T}_{s_i} f_\varepsilon\|+\|\overset{0}{T}_{s_i} f_\varepsilon-\overset{n}{T}_{s_i} f_\varepsilon\|+\|\overset{n}{T}_{s_i} f_\varepsilon-\overset{n}{T}_t f_\varepsilon\|\\
&\leqq \frac{2\varepsilon}{3}+(t-s_i)\|A_0 f_\varepsilon\|+\frac{\varepsilon}{6}+(t-s_i)\|A_n f_\varepsilon\|<\varepsilon
\end{aligned}$$

for $n\geqq n_\varepsilon$. This proves the lemma.

2. The Main Results

2.1. By now we have finished all the preliminary work and are in a position where we can formulate and prove the central result of the present paper. This result is a version of the central limit theorem where instead of considering sums of vector-valued random variables we consider "sums" of random geodesic segments in a Riemannian manifold. Such a "sum" could naturally be called a geodesic random walk, and what we shall do is to consider a sequence of such random walks where within a given time-period the number of steps taken increases indefinitely while at the same time the length of the individual step decreases to zero. In the limit we then expect to obtain a diffusion process in $\mathscr{M}$, and in particular we expect to obtain the Brownian motion if we let the individual steps in the random walks be identically distributed.

However, before we give the final rigorous formulation of our problem let us do a little pseudo-mathematics in order to motivate this formulation.

Assume there is defined on $\mathscr{M}$ a family[10] of subprobability measures $\{\nu_p\}_{p\in\mathscr{M}}$, where each ν_p is a measure on $\mathscr{M}_p$. Let $\tau>0$ and $p_0\in\mathscr{M}$ be given and define for $0\leqq t\leqq\tau$ the random path $\hat{\xi}(t)$ starting at p_0 by[11]

$$\hat{\xi}(t)=\exp_{p_0}\left\{\frac{t}{\tau}\left[\tau\mu_{p_0}+\sqrt{\tau}(X_{p_0}-\mu_{p_0})\right]\right\}$$

where X_{p_0} is a random variable with values in $\mathscr{M}_{p_0}$ and distribution ν_{p_0} and

$$\mu_{p_0}=\int_{\mathscr{M}_{p_0}} X\,\nu_{p_0}(dX)=\text{mean value of } X_{p_0}.$$

Also, if $\nu_{p_0}(\mathscr{M}_{p_0})<1$, set

$$\operatorname{Prob}\{\hat{\xi}(t)=\Delta \text{ for } t>0\}=1-\nu_{p_0}(\mathscr{M}_{p_0}).$$

In general "define" $\hat{\xi}(t)$ by induction

$$p_n=\hat{\xi}(n\tau),\qquad n=0,1,\ldots,$$

$$\hat{\xi}(t)=\exp_{p_n}\left\{\frac{t-n\tau}{\tau}\left[\tau\mu_{p_n}+\sqrt{\tau}(X_{p_n}-\mu_{p_n})\right]\right\}$$

for $n\tau\leqq t\leqq(n+1)\tau$. If $p_n=\Delta$ set $\hat{\xi}(t)=\Delta$ for $n\tau\leqq t$, while if $p_n\in\mathscr{M}$ set

$$\operatorname{Prob}\{\hat{\xi}(t)=\Delta \text{ for } n\tau<t\}=1-\nu_{p_n}(\mathscr{M}_{p_n}).$$

Now, if there exists a stochastic process $\hat{\xi}(t)$ satisfying these "definitions" it may serve as a reasonable mathematical model for diffusion in $\mathscr{M}$ (in the limit for $\tau\to 0$).

The sequence $\{\hat{\xi}(n\tau)\}_{n=0}^{\infty}$ is a Markov process with stationary transition probabilities

$$P\{\hat{\xi}((n+1)\tau)=\Delta\mid\hat{\xi}(n\tau)=\Delta\}=1$$

$$P\{\hat{\xi}((n+1)\tau)\in A\mid\hat{\xi}(n\tau)=p\}$$
$$=(1-\nu_p(\mathscr{M}_p))\,\chi_A(\Delta)+\nu_p\{\mu_p+\tau^{-\frac{1}{2}}[\exp_p^{-1}(A\setminus\{\Delta\})-\tau\mu_p]\},\qquad p\neq\Delta;$$

while for $n\tau<t<(n+1)\tau$ $\hat{\xi}(t)$ is just a geodesic segment from $\hat{\xi}(n\tau)$ to $\hat{\xi}((n+1)\tau)$.

Because the process $\hat{\xi}(t)$ is a little awkward to handle we shall introduce another process $\xi(t)$ which in a certain sense is just as good as $\hat{\xi}(t)$ and which also is much easier to work with. Let $(n(t), 0\leqq t<\infty)$ be a Poisson process independent of $\hat{\xi}(t)$, with parameter τ^{-1} and right-continuous sample paths. Set

$$\xi_\tau(t)=\hat{\xi}(n(t)\cdot\tau),$$

then $\xi_\tau(t)$ is a Markov process with right-continuous paths and transition function

$$P_\tau(t,p,\Gamma)=e^{-t/\tau}\sum_{k=0}^{\infty}\frac{(t/\tau)^k}{k!}P_\tau^{(k)}(p,\Gamma)$$

[10] See Section 4.4.

[11] "$\exp_p$" denotes the exponential map in $\mathscr{M}$.

where $P_\tau^{(k)}(\cdot,\cdot)$ is the k-th iterate of the transition function

$$P(p,\Gamma)=\mathrm{Prob}\{\hat{\xi}(\tau)\in\Gamma\,|\,\hat{\xi}(0)=p\}.$$

For t fixed the mean value of $n(t)$ is t/τ, so loosely spoken

$$\xi_\tau(t)\sim\hat{\xi}\left(\frac{t}{\tau}\cdot\tau\right)=\hat{\xi}(t);$$

this is what we mean by saying $\xi_\tau(t)$ is just as good as $\hat{\xi}(t)$. Also $\xi_\tau(t)$ has the advantage of being a Markov process, which makes it possible for us to define it by specifying its associated semigroup, thus circumventing the problem of translating the above statements into mathematics. We shall show that $\xi_\tau(\cdot)$ converges weakly to a diffusion process $\eta(\cdot)$ as $\tau\to 0$, and by estimating the difference between $\hat{\xi}(t)=\hat{\xi}_\tau(t)$ and $\xi_\tau(t)$ properly, we could also show that $\hat{\xi}_\tau(\cdot)\to\eta(\cdot)$ as $\tau\to 0$; but this seems hardly worthwhile, and we shall not worry about doing it.

2.2. From now on assume that our manifold $\mathscr{M}$ is of class $\mathscr{A}\mathscr{B}$[12]. Let $\{\nu_p^{(n)}\}_{p\in\mathscr{M}}$, $n=1,2,\ldots$, be a sequence of families of subprobability measures on the tangent spaces $\mathscr{M}_p$, satisfying the following conditions for every $\delta>0$[13]

$\mathscr{C}0$: For each $n\geqq 1$ is

$$s\text{-}\lim_{r\to\infty}\nu_p^{(n)}\{Y\colon Y\in\mathscr{M}_p,\ \|Y\|\geqq r\}=0.$$

$\mathscr{C}1$: $$s\text{-}\lim_{n\to\infty} n\{1-\nu_p^{(n)}(\mathscr{M}_p)\}=k(p).$$

$\mathscr{C}2$: $$s\text{-}\lim_{n\to\infty} n\,\nu_p^{(n)}\{Y\colon Y\in\mathscr{M}_p,\ \|Y\|\geqq\delta\sqrt{n}\}=0.$$

$\mathscr{C}3$: $$s\text{-}\lim_{n\to\infty}\int_{\{\|Y\|<\delta\sqrt{n}\}} Y\nu_p^{(n)}(dY)=\mu_p\in\mathscr{M}_p.$$

$\mathscr{C}4$: $$s\text{-}\lim_{n\to\infty}\int_{\{\|Y\|<\delta\sqrt{n}\}} Y\otimes Y\nu_p^{(n)}(dY)=\mu_p^{(2)}\in\mathscr{M}_p\otimes\mathscr{M}_p.$$

The functions $k(\cdot)$, $\mu_\cdot$, and $\mu_\cdot^{(2)}$ defined by $\mathscr{C}1$, 3, 4 are assumed to satisfy certain conditions. We require

$\mathscr{C}5$: For some α, $0<\alpha\leqq 1$, the functions $k(\cdot)$, $\mu_\cdot$, and $\mu_\cdot^{(2)}$ are Hölder continuous of order α. For each $p\in\mathscr{M}$, $\mu_p^{(2)}-\mu_p\otimes\mu_p=\sigma_p^2$ is strictly positive definite[14]. Finally

$$\max\big[\sup_p k(p),\ \sup_p\|\mu_p\|,\ \sup_p\|\mu_p^{(2)}\|\big]=M_0<\infty. \tag{2.1}$$

It should be noted that $\mathscr{C}2$ implies the validity of $\mathscr{C}3$ and $\mathscr{C}4$ for all $\delta>0$ once they have been proven to hold for a particular $\delta>0$, and also that the limits μ_p and $\mu_p^{(2)}$ are independent of the choice of δ. One may even show (see [4] p. 188,

[12] That is, satisfies both axioms $\mathscr{A}$ and $\mathscr{B}$ in 4.3.

[13] Recall that "s-lim" means "uniformly on $\mathscr{M}$". We shall also use the notation $f_n\xrightarrow[n\to\infty]{u}f$, instead of: $s\text{-}\lim_{n\to\infty} f_n=f$.

[14] That is $\langle\xi\otimes\xi,\sigma^2\rangle=0$ must imply $\xi=0$. This condition is used only to prove the uniqueness part of our limit theorem.

Lemma 1) that there exists a sequence $\{\delta_n\}_{n=1}^{\infty}$ such that

$$0<\cdots<\delta_n<\delta_{n-1}<\cdots<\delta_1\leqq 1;\quad \sqrt{n}\,\delta_n\uparrow\infty \quad\text{and}\quad \delta_n\downarrow 0 \quad\text{for } n\to\infty, \tag{2.2}$$
$$s\text{-}\lim_{n\to\infty} n\,\nu_p^{(n)}\{Y\colon Y\in\mathcal{M}_p,\ \|Y\|\geqq\delta_n\sqrt{n}\}=0,$$

and it follows easily that δ may be replaced with δ_n in the statements $\mathscr{C}3$, $\mathscr{C}4$.

Introduce the notation

$$\begin{aligned}
&H^{\delta}(p,n)=\{Y\colon Y\in\mathcal{M}_p,\ \|Y\|<\delta\sqrt{n}\},\\
&G(p,n)=H^{\delta_n}(p,n),\\
&k_n(p)=n\{1-\nu_p^{(n)}(\mathcal{M}_p)\},\\
&\mu_{p,n}=\int_{H^1(p,n)} Y\nu_p^{(n)}(dY),\\
&\mu_{p,n}^{(2)}=\int_{H^1(p,n)} Y\otimes Y\nu_p^{(n)}(dY).
\end{aligned}$$

Then $\mu_{p,n}\in\mathcal{M}_p$, $\mu_{p,n}^{(2)}\in\mathcal{M}_p\otimes\mathcal{M}_p$, and from the remarks above it follows that

$$\mu_p=s\text{-}\lim_{n\to\infty}\mu_{p,n}=s\text{-}\lim_{n\to\infty}\int_{G(p,n)} Y\nu_p^{(n)}(dY),$$
$$\mu_p^{(2)}=s\text{-}\lim_{n\to\infty}\mu_{p,n}^{(2)}=s\text{-}\lim_{n\to\infty}\int_{G(p,n)} Y\otimes Y\nu_p^{(n)}(dY).$$

We may also assume that

$$\max\big[\sup_{p,n}k_n(p),\ \sup_{p,n}\|\mu_{p,n}\|,\ \sup_{p,n}\|\mu_{p,n}^{(2)}\|\big]=M<\infty$$

because of (2.1) and the uniform convergence.

Define on $B=B(\mathcal{M})$ the operators

$$\begin{aligned}
\overset{n}{T}f(p)&=\int_{\mathcal{M}_p} f\left(\exp_p\left\{\frac{1}{n}\mu_{p,n}+\frac{1}{\sqrt{n}}(Y-\mu_{p,n})\right\}\right)\nu_p^{(n)}(dY)\\
&=\int_{\mathcal{M}_p} f(\exp_p Z)\,\tilde{\nu}_p^{(n)}(dZ),\\
\overset{n}{S}f(p)&=\overset{n}{T}f(p)+\frac{1}{n}k_n(p)f(p)=\int_{\mathcal{M}_p} f(\exp_p Z)\,\theta_p^{(n)}(dZ),
\end{aligned}$$

where $\tilde{\nu}_p^{(n)}$ is the distribution of

$$Z=\frac{1}{n}\mu_{p,n}+\frac{1}{\sqrt{n}}(Y-\mu_{p,n})$$

and

$$\theta_p^{(n)}(\Gamma)=\tilde{\nu}_p^{(n)}(\Gamma)+\chi_\Gamma(0)\big(1-\tilde{\nu}_p^{(n)}(\mathcal{M}_p)\big).$$

In particular, if $f=\chi_\Gamma$ with $\Gamma\in\mathscr{B}$ then

$$\overset{n}{T}f(p)=P_n(p,\Gamma),\qquad \overset{n}{S}f(p)=\overset{0}{P}_n(p,\Gamma)$$

where $P_n(p, \cdot)$ and $\overset{0}{P}_n(p, \cdot)$ are subprobability measures on $\mathscr{M}$. Thus we may write

$$\overset{n}{T} f(p) = \int_{\mathscr{M}} f(q)\, P_n(p, dq); \qquad \overset{n}{S} f(p) = \int_{\mathscr{M}} f(q)\, \overset{0}{P}_n(p, dq),$$

and as the notation suggests we want P_n and $\overset{0}{P}_n$ to be transition functions. If we compare with our model from Section 2.1 $P_n(p, \Gamma)$ of course represents

$$\operatorname{Prob}\left\{\hat{\xi}\left(\frac{i}{n}\right) \in \Gamma \,\middle|\, \hat{\xi}\left(\frac{i-1}{n}\right) = p\right\},$$

while the auxiliary quantities $\overset{0}{P}_n$ are introduced in order to facilitate the proofs of some estimates to be given later on. They correspond to the situation[15] where $v_p^{(n)}(\mathscr{M}_p) \equiv 1$ and thus they may be expected to behave better than the P_n.

Now, in order to get the functions P_n and $\overset{0}{P}_n$ to be transition functions we must impose some kind of restriction upon the variation of the $v_p^{(n)}$ with p. Our final condition takes care of this[16]

$\mathscr{C}6$: For each n $\lim\limits_{p \to \infty} k_n(p)$ exists. Also, there shall exist constants $K > 0$, $\delta_0 > 0$, and α, $0 < \alpha \leqq 1$, such that for $0 < \delta \leqq \delta_0$ and all $n \geqq 1$

$$\sup_{d(p,q) \leqq \delta} n\, |\tilde{v}_p^{(n)} - \tilde{v}_q^{(n)}| \leqq K \delta^{\alpha}.$$

We can now prove

Lemma 2.1. *For each n $k_n(\cdot) \in \tilde{C}$. Moreover, the operators $\overset{n}{T}$ and $\overset{n}{S}$ map $\tilde{C}$ into $\tilde{C}$ and C_0 into C_0.*

Proof. First, we have the estimate

$$\begin{aligned} |k_n(p) - k_n(q)| &= n\, |v_p^{(n)}(\mathscr{M}_p) - v_q^{(n)}(\mathscr{M}_q)| \\ &= n\, |\tilde{v}_p^{(n)}(\mathscr{M}_p) - \tilde{v}_q^{(n)}(\mathscr{M}_q)| \leqq n\, |\tilde{v}_p^{(n)} - \tilde{v}_q^{(n)}| \leqq K [d(p,q)]^{\alpha} \end{aligned}$$

whenever $d(p, q) \leqq \delta_0$. Thus $k_n(\cdot)$ is Hölder continuous of order α. Moreover, $\lim\limits_{p \to \infty} k_n(p)$ exists per assumption, hence $k_n(\cdot) \in \tilde{C}$. Note that this argument in combination with $\mathscr{C}1$ actually shows that $k(\cdot)$ is Hölder continuous of order α and is contained in $\tilde{C}$. Next, if $f \in \tilde{C}$ then we may write

$$f(p) = f_0(p) + b$$

with $f_0 \in C_0$ and b constant. Thus to finish the proof of the lemma it suffices to show that $\overset{n}{T}$ maps C_0 into itself and that $\overset{n}{T} 1 \in \tilde{C}$, and the second statement is easily proved:

$$\overset{n}{T} 1(p) = 1 - \frac{1}{n} k_n(p), \quad \text{so} \quad \overset{n}{T} 1 \in \tilde{C}.$$

[15] If $v_p^{(n)}(\mathscr{M}_p) \equiv 1$ then $\overset{n}{S} f = \overset{n}{T} f$.

[16] The notation is explained in the appendix. Note that we may use the same α in $\mathscr{C}5$ and in $\mathscr{C}6$.

To prove the first, let $f \in C_0$ be given and let p_0, p be a pair of points in $\mathscr{M}$. Denote parallel translation along a short geodesic $\widehat{p_0 p}$ by τ_p, then

$$\begin{aligned}
\overset{n}{T}f(p)-\overset{n}{T}f(p_0) &= \int_{\mathscr{M}_p} f(\exp_p X)\,\tilde{\nu}_p^{(n)}(dX) - \int_{\mathscr{M}_{p_0}} f(\exp_{p_0} Z)\,\tilde{\nu}_{p_0}^{(n)}(dZ) \\
&= \int_{\mathscr{M}_{p_0}} f(\exp_p[\tau_p Z])(\tau_p^{-1}\tilde{\nu}_p^{(n)})(dZ) - \int_{\mathscr{M}_{p_0}} f(\exp_{p_0} Z)\,\tilde{\nu}_{p_0}^{(n)}(dZ) \\
&= \int_{\mathscr{M}_{p_0}} f(\exp_p[\tau_p Z])(\tau_p^{-1}\tilde{\nu}_p^{(n)}-\tilde{\nu}_{p_0}^{(n)})(dZ) \\
&+ \int_{\mathscr{M}_{p_0}} \{f(\exp_p[\tau_p Z])-f(\exp_{p_0} Z)\}\,\tilde{\nu}_{p_0}^{(n)}(dZ).
\end{aligned} \tag{2.3}$$

The first integral is estimated thusly

$$\begin{aligned}
\Big|\int_{\mathscr{M}_{p_0}} f(\exp_p[\tau_p Z])(\tau_p^{-1}\tilde{\nu}_p^{(n)}-\tilde{\nu}_{p_0}^{(n)})(dZ)\Big| &\leqq \|f\|\;|\tilde{\nu}_p^{(n)}-\tilde{\nu}_{p_0}^{(n)}| \\
&\leqq \frac{\|f\|}{n} K[d(p,p_0)]^{\alpha} \to 0 \qquad \text{for } p \to p_0,
\end{aligned}$$

while the second integral is seen to tend to zero when $p \to p_0$ because of the bounded convergence theorem and the continuity of the map

$$p \to \exp_p(\tau_p Z)$$

at p_0. Thus $\overset{n}{T}f$ is continuous. Next, let $\varepsilon > 0$ be given and choose a compact set $\Gamma \subset \mathscr{M}$ such that

$$|f(p)| < \tfrac{1}{2}\varepsilon \qquad \text{for } p \notin \Gamma,$$

then

$$|\overset{n}{T}f(p)| \leqq \tfrac{1}{2}\varepsilon + \Big|\int_\Gamma f(q)\,P_n(p,dq)\Big| \leqq \tfrac{1}{2}\varepsilon + \|f\|\,P_n(p,\Gamma)$$

where by definition

$$P_n(p,\Gamma) = \tilde{\nu}_p^{(n)}(\exp_p \Gamma) = \nu_p^{(n)}\left\{Y: \exp_p\left[\frac{1}{n}\mu_{p,n} + \frac{1}{\sqrt{n}}(Y-\mu_{p,n})\right] \in \Gamma\right\}.$$

If $r > 4M_0$ then

$$\begin{aligned}
\nu_p^{(n)}\left\{Y: \left\|\frac{1}{n}\mu_{p,n} + \frac{1}{\sqrt{n}}(Y-\mu_{p,n})\right\| > r\right\} &\\
\leqq \nu_p^{(n)}\{Y: \|Y-\mu_{p,n}\| > \tfrac{3}{4}r\sqrt{n}\} &\\
\leqq \nu_p^{(n)}\{Y: \|Y\| > \tfrac{1}{2}r\sqrt{n}\} < \frac{\varepsilon}{2\|f\|} &
\end{aligned}$$

for r sufficiently large according to $\mathscr{C}0$.

So, if we set

$$\Gamma_1 = \{q: q \in \mathscr{M} \text{ and } d(q,\Gamma) \leqq r\}$$

then Γ_1 is compact[17] and for $p \in \Gamma_1^c$

$$P_n(p, \Gamma) \leqq \tilde{\nu}_p^{(n)}\{Y\colon \|Y\| > r\} < \frac{\varepsilon}{2\|f\|}$$

that is

$$|\overset{n}{T} f(p)| \leqq \frac{\varepsilon}{2} + \|f\| \, P_n(p, \Gamma) < \varepsilon \quad \text{for } p \notin \Gamma_1,$$

and we have thus shown $\overset{n}{T} f \in C_0$. This finishes the proof of the lemma.

Note. It is easy to show that the operators $\overset{n}{T}$ and $\overset{n}{S}$ map B into itself also, and that in fact they are positive contraction operators on B. In the following we shall, however, unless otherwise indicated think of them as being operators on C_0.

Define for $p \in \mathcal{M}$, $\Gamma \in \mathcal{B}(\mathcal{M}) = \mathcal{B}$

$$P_n^{(0)}(p, \Gamma) = \overset{0}{P}{}_n^{(0)}(p, \Gamma) = \chi_\Gamma(p)$$

and for $t \geqq 0, f \in C_0$

$$\overset{n}{T}_t f(p) = e^{-nt} \sum_{k=0}^{\infty} \frac{(nt)^k}{k!} \int_{\mathcal{M}} f(q) \, P_n^{(k)}(p, dq)$$

$$\overset{n}{S}_t f(p) = e^{-nt} \sum_{k=0}^{\infty} \frac{(nt)^k}{k!} \int_{\mathcal{M}} f(q) \, \overset{0}{P}{}_n^{(k)}(p, dq)$$

where $\overset{(0)}{P}{}_n^{(k)}$ is the k-th iterate of $\overset{(0)}{P}_n$. If we set

$$\overset{(0)}{P}_n(t, p, \Gamma) = e^{-nt} \sum_{k=0}^{\infty} \frac{(nt)^k}{k!} \overset{(0)}{P}{}_n^{(k)}(p, \Gamma)$$

then

$$\overset{n}{T}_t f(p) = \int_{\mathcal{M}} f(q) \, P_n(t, p, dq); \quad \overset{n}{S}_t f(p) = \int_{\mathcal{M}} f(q) \, \overset{0}{P}_n(t, p, dq),$$

and we have

Proposition 2.1. *The families* $(\overset{n}{T}_t, t \geqq 0)$ *and* $(\overset{n}{S}_t, t \geqq 0)$ *are* C_0*-semigroups. Their infinitesimal generators are respectively*

$$A_n f = n\{\overset{n}{T} f - f\}, \quad \text{and} \quad \overset{0}{A}_n f = n\{\overset{n}{S} f - f\}$$

with the common domain $\mathcal{D}_n = C_0$. *If* $f \geqq 0$

(a) $$0 \leqq \overset{n}{T}_t f \leqq \overset{n}{S}_t f$$

and in general for $\Gamma \in \mathcal{B}$

(b) $$P_n(t, p, \Gamma) \leqq \overset{0}{P}_n(t, p, \Gamma).$$

Proof. For the first part it suffices to work with the $\overset{n}{T}_t$. Let $f \in C_0$, then according to Lemma 2.1 $\overset{n}{T} f \in C_0$ and more generally

$$\int_{\mathcal{M}} f(q) \, P_n^{(k)}(p, dq) = \overset{n}{T}{}^k f(p) \in C_0, \quad \text{for } k \geqq 0.$$

[17] Because $\mathcal{M}$ is complete.

Also $\|\overset{n}{T}{}^k f\| \leqq \|f\|$ so the series defining $\overset{n}{T}_t f$ converges uniformly on $\mathscr{M}$, and this implies at once that $\overset{n}{T}_t f \in C_0$. It is also easy to see that

$$\|\overset{n}{T}_t f\| \leqq \|f\|,$$

that $\overset{n}{T}_t$ is a positive operator, and that $(\overset{n}{T}_t, t \geqq 0)$ is a semigroup. Finally

$$\|\overset{n}{T}_t f - f\| \leqq (1-e^{-nt})\,\|f\| + e^{-nt} \sum_{k=1}^{\infty} \frac{(nt)^k}{k!}\,\|f\| = 2\,\|f\|\,(1-e^{-nt})$$

so

$$s\text{-}\lim_{t \downarrow 0} \overset{n}{T}_t f = f,$$

thus $(\overset{n}{T}_t, t \geqq 0)$ is a C_0-semigroup. Next, for $f \in C_0$

$$\begin{aligned}
&\left\| \frac{1}{t}[\overset{n}{T}_t f - f] - A_n f \right\| \\
&= \left\| \frac{1}{t}[\overset{n}{T}_t f - f] - n[\overset{n}{T} f - f] \right\| = \left\| \frac{1}{t} e^{-nt} \sum_{k=0}^{\infty} \frac{(nt)^k}{k!}[\overset{n}{T}{}^k f - f] - n[\overset{n}{T} f - f] \right\| \\
&= \left\| e^{-nt} \sum_{k=2}^{\infty} \frac{(nt)^{k-1}}{k!} \left\{ n \sum_{j=0}^{k-1} \overset{n}{T}{}^j (\overset{n}{T} f - f) \right\} - (1-e^{-nt})\, n[\overset{n}{T} f - f] \right\| \\
&\leqq e^{-nt} \sum_{k=2}^{\infty} \frac{(nt)^{k-1}}{(k-1)!} \|n(\overset{n}{T} f - f)\| + (1-e^{-nt})\,\|n(\overset{n}{T} f - f)\| \\
&= 2(1-e^{-nt})\,\|n(\overset{n}{T} f - f)\| \to 0 \qquad \text{for } t \downarrow 0
\end{aligned}$$

which shows A_n with domain $\mathscr{D}_n = C_0$ is the infinitesimal generator for $(\overset{n}{T}_t, t \geqq 0)$.

To prove the second part of the proposition, note that if $f \in B, f \geqq 0$ then

$$0 \leqq \overset{n}{T} f \leqq \overset{n}{T} f + \frac{1}{n} k_n f = \overset{n}{S} f,$$

and in general for $l > 1$

$$\overset{n}{T}{}^l f = \overset{n}{T}{}^{l-1}(\overset{n}{T} f) \leqq \overset{n}{T}{}^{l-1}(\overset{n}{S} f) \leqq \cdots \leqq \overset{n}{S}{}^l f,$$

which immediately yields

$$\overset{n}{T}_t f \leqq \overset{n}{S}_t f, \qquad \text{for } f \geqq 0$$

and this proves (a). As (b) is a special case of (a) this completes the proof of the proposition. In the passing, one might note that in order to infer the validity of (b) from that of (a) one only has to know that (a) is true for $f \in C_0$.

Our goal is now to demonstrate that the semigroups $(\overset{n}{T}_t, t \geqq 0)$ and $(\overset{n}{S}_t, t \geqq 0)$ converge as $n \to \infty$ to C_0-semigroups $(T_t, t \geqq 0)$ and $(S_t, t \geqq 0)$. To carry out this demonstration, we first make use of Lemma 1.5 to show the existence of convergent subsequences, and then we show that the limits hereby obtained are independent of the particular subsequences that have been chosen.

First, let us verify that condition (b) from Lemma 1.5 is satisfied. This is a consequence of

Proposition 2.2. *If $f \in \tilde{C}$ and $t \geqq 0$ then the families $\{\overset{n}{T}_t f\}_{n=1}^{\infty}$ and $\{\overset{n}{S}_t f\}_{n=1}^{\infty}$ are equicontinuous.*

Proof. We start with a few estimates. Define

$$\alpha_n = 2\sup_{p\in\mathcal{M}} \tilde{v}_p^{(n)}\{Z\colon Z\in\mathcal{M}_p,\ \|Z\|>1\},$$

then according to $\mathscr{C}2$

(a) $$\lim_{n\to\infty} n\alpha_n = 0.$$

Next, for some $\eta>0$ and all p, n

$$\{Y\colon Y\in\mathcal{M}_p,\ \|\mu_{p,n}+\sqrt{n}(Y-\mu_{p,n})\|\leqq n\}\subset\{Y\colon Y\in\mathcal{M}_p,\ Y<\eta\sqrt{n}\},$$

so $\mathscr{C}4$ implies

$$\begin{aligned}\int_{\{\|Z\|\leqq 1\}} \|Z\|^2\,\tilde{v}_p^{(n)}(dZ) &= \int_{\{\|\mu_{p,n}+\sqrt{n}(Y-\mu_{p,n})\|\leqq n\}} \left\|\frac{1}{n}(\mu_{p,n}+\sqrt{n}(Y-\mu_{p,n}))\right\|^2 v_p^{(n)}(dY)\\ &\leqq \frac{1}{n}\int_{\{Y<\eta\sqrt{n}\}}\left[\frac{1}{\sqrt{n}}M+M+\|Y\|^2\right] v_p^{(n)}(dY)\\ &\leqq \frac{1}{n}\left[2M+\sqrt{N}\left\|\int_{\{Y<\eta\sqrt{n}\}} Y\otimes Y v_p^{(n)}(dY)\right\|\right]\leqq\frac{c_0}{n}\end{aligned}$$

for some constant c_0, that is independent of p and n. Thus

(b) $$\int_{\{\|Z\|\leqq 1\}} \|Z\|^2\,\tilde{v}_p^{(n)}(dZ)\leqq\frac{c_0}{n}.$$

Let α be a constant satisfying $0<\alpha\leqq 1$, and let p, q be points in $\mathcal{M}$ with $d(p,q)=\delta\leqq\min(1,\delta_0)$[18], then from (b) and axiom $\mathscr{B}$[19]

(c) $$\begin{aligned}\int_{\{\|Z\|\leqq 1\}} [d(\exp_p Z, \exp_q \tau_{pq} Z)]^\alpha\,\tilde{v}_p^{(n)}(dZ) &\leqq \delta^\alpha \int_{\{\|Z\|\leqq 1\}} [1+c\|Z\|^2]\,\tilde{v}_p^{(n)}(dZ)\\ &\leqq \delta^\alpha\left[1+\frac{c\,c_0}{n}\right]=\delta^\alpha\beta_n.\end{aligned}$$

Now, if $f\in\tilde{C}$ and satisfies the Lipschitz condition

$$|f(q_1)-f(q_2)|\leqq a\,d(q_1,q_2);\qquad q_1,q_2\in\mathcal{M}, \tag{2.4}$$

then, using the expression (2.3), we obtain

$$\begin{aligned}|\overset{n}{T}f(p)-\overset{n}{T}f(q)| &\leqq \|f\|\cdot|\tilde{v}_p^{(n)}-\tilde{v}_q^{(n)}|+\left|\int_{\mathcal{M}_q}[f(\exp_p\tau_{qp}Z)-f(\exp_q Z)]\,\tilde{v}_q^{(n)}(dZ)\right|\\ &\leqq \|f\|\cdot[|\tilde{v}_p^{(n)}-\tilde{v}_q^{(n)}|+2\tilde{v}_q^{(n)}\{\|Z\|>1\}]\\ &\quad+\int_{\{\|Z\|\leqq 1\}} a\,d(\exp_p\tau_{qp}Z,\exp_q Z)\,\tilde{v}_q^{(n)}(dZ)\\ &\leqq \|f\|\cdot\left[\frac{K}{n}\delta^\alpha+\alpha_n\right]+a\,\delta\,\beta_n,\end{aligned}$$

according to $\mathscr{C}6$, (c), and the definition of α_n.

---

[18] δ_0 comes from $\mathscr{C}6$.

[19] As usual τ_{pq} denotes parallel translation along a short geodesic segment $\widehat{pq}$.

Next, estimate $\overset{n}{T}{}^2 f$

$$
\begin{aligned}
&|\overset{n}{T}{}^2 f(p)-\overset{n}{T}{}^2 f(q)| \\
&\quad \leqq \|\overset{n}{T} f\| \cdot \left[\frac{K}{n}\delta^\alpha+\alpha_n\right]+\int_{\{\|Z\|\leqq 1\}} |\overset{n}{T} f(\exp_p \tau_{qp} Z)-\overset{n}{T} f(\exp_q Z)|\, \tilde{v}_q^{(n)}(dZ) \\
&\quad \leqq \|f\| \cdot \left[\frac{K}{n}\delta^\alpha+2\alpha_n+\int_{\{\|Z\|\leqq 1\}} \frac{K}{n}[\delta(1+c\|Z\|^2)]^\alpha\, \tilde{v}_q^{(n)}(dZ)\right]+a\,\delta\,\beta_n^2 \\
&\quad \leqq \|f\| \cdot \left[2\alpha_n+\frac{K}{n}\delta^\alpha(1+\beta_n)\right]+a\,\delta\,\beta_n^2
\end{aligned}
$$

provided $\delta\,\beta_n \leqq 1$.

Iterating this procedure we get

$$|\overset{n}{T}{}^l f(p)-\overset{n}{T}{}^l f(q)| \leqq \|f\| \cdot \left[l\alpha_n+\frac{K}{n}\delta^\alpha\sum_{j=0}^{l-1}\beta_n^j\right]+a\,\delta\,\beta_n^l$$

for $l=1,2,\ldots,nL$ provided $\delta\,\beta_n^{nL}\leqq 1$. But

$$\beta_n^{nL}=\left(1+\frac{c\,c_0}{n}\right)^{nL}\leqq e^{Lcc_0}$$

so it suffices to demand

$$\delta=d(p,q)\leqq e^{-Lcc_0}.$$

Also

$$\sum_{j=0}^{l-1}\beta_n^j=\frac{\beta_n^{nL}-1}{\beta_n-1}\leqq n\,\frac{e^{Lcc_0}}{c\,c_0}$$

and consequently

(d) $$|\overset{n}{T}{}^l f(p)-\overset{n}{T}{}^l f(q)|\leqq \|f\| \cdot \left[n\,L\,\alpha_n+\frac{K\,e^{Lcc_0}}{c\,c_0}\delta^\alpha\right]+a\,\delta\,e^{Lcc_0}$$

if

$$\delta=d(p,q)<\min(1,\delta_0,e^{-Lcc_0}), \quad 0\leqq l\leqq nL.$$

The operators $\overset{n}{S}$ are estimated in a similar fashion.

First

$$
\begin{aligned}
|\overset{n}{S} f(p)-\overset{n}{S} f(q)| &\leqq |\overset{n}{T} f(p)-\overset{n}{T} f(q)|+\frac{1}{n}|k_n(p)-k_n(q)|\,|f(p)|+\frac{1}{n}|k_n(q)|\,|f(p)-f(q)| \\
&\leqq |\overset{n}{T} f(p)-\overset{n}{T} f(q)|+\frac{M}{n}|f(p)-f(q)|+\|f\|\,\frac{K}{n}[d(p,q)]^\alpha,
\end{aligned}
$$

for $d(p,q)\leqq\delta_0$. So proceeding as before we get

$$|\overset{n}{S}{}^l f(p)-\overset{n}{S}{}^l f(q)|\leqq \|f\| \cdot \left[\alpha_n\sum_{j=0}^{l-1}\left(1+\frac{M}{n}\right)^j+2\frac{K}{n}\delta^\alpha\sum_{j=0}^{l-1}\left(\beta_n+\frac{M}{n}\right)^j\right]+a\,\delta\left(\beta_n+\frac{M}{n}\right)^l$$

for $0 \leqq l \leqq nL$ provided $\delta\left(\beta_n + \frac{M}{n}\right)^{nL} \leqq 1$. Moreover

$$\sum_{j=0}^{nL-1}\left(1+\frac{M}{n}\right)^j \leqq \frac{n}{M}\left(1+\frac{M}{n}\right)^{nL} \leqq n\frac{e^{ML}}{M},$$

and

$$\sum_{j=0}^{nL-1}\left(\beta_n+\frac{M}{n}\right)^j = \sum_{j=0}^{nL-1}\left(1+\frac{c c_0+M}{n}\right)^j \leqq n\frac{e^{(cc_0+M)L}}{c c_0+M}$$

so altogether

$$\text{(e)} \qquad |\overset{n}{S}{}^l f(p) - \overset{n}{S}{}^l f(q)| \leqq \frac{e^{(cc_0+M)L}}{M}\left[\|f\|(n\alpha_n + 2K\delta^\alpha) + a\delta M\right]$$

if

$$\delta = d(p,q) \leqq e^{-(cc_0+L)}, \qquad 0 \leqq l \leqq nL.$$

Comparing (d) and (e) we see that they are essentially equivalent. It will be demonstrated below that (d) implies the equicontinuity of the family $\{\overset{n}{T}_t f\}_{n=1}^\infty$, and it is then clear that (e) in a similar fashion implies the equicontinuity of the family $\{\overset{n}{S}_t f\}_{n=1}^\infty$.

Let $t_0 > 0$ and $\varepsilon_0 > 0$ be given. From Chebyshev's inequality applied to the Poisson variable ξ with parameter $n t_0$ we get for $L = 2 + [t_0]$

$$e^{-nt_0}\sum_{k=nL+1}^{\infty}\frac{(n t_0)^k}{k!} = \mathrm{Prob}\{\xi \geqq nL+1\} \leqq \mathrm{Prob}\{|\xi - n t_0| \geqq nL+1-n t_0\}$$
$$\leqq \frac{n t_0}{(nL+1-n t_0)^2} < \frac{t_0}{n}.$$

Choose n_0 such that for $n \geqq n_0$

$$\frac{6 t_0 \|f\|}{n} < \varepsilon_0 \quad \text{and} \quad 3\|f\|\{2+t_0\}\, n\alpha_n < \varepsilon_0$$

then (d) yields for $n \geqq n_0$ and $\delta = d(p,q) < e^{-(2+t_0)cc_0}$

$$|\overset{n}{T}_{t_0} f(p) - \overset{n}{T}_{t_0} f(q)| = e^{-nt_0}\left|\sum_{k=0}^{\infty}\frac{(n t_0)^k}{k!}(\overset{n}{T}{}^k f(p) - \overset{n}{T}{}^k f(q))\right|$$
$$\leqq e^{-nt_0}\left[\sum_{l=0}^{nL}\frac{(n t_0)^l}{l!}|\overset{n}{T}{}^l f(p) - \overset{n}{T}{}^l f(q)| + 2\|f\| \cdot \sum_{l=nL+1}^{\infty}\frac{(n t_0)^l}{l!}\right]$$
$$\leqq \frac{\varepsilon_0}{3} + \|f\| \cdot \left[L n\alpha_n + \frac{K e^{Lcc_0}}{c c_0}\delta^\alpha\right] + a\delta\, e^{Lcc_0}$$
$$< \frac{2\varepsilon_0}{3} + \tilde{K}[d(p,q)]^\alpha < \varepsilon_0$$

if $d(p,q) < (\tilde{K}^{-1}\varepsilon_0)^\alpha$, where the constant $\tilde{K}$ is independent of p, q and n.

Now, each of the functions $\overset{n}{T}_{t_0} f$ is contained in $\tilde{C}$; hence there exists a $\gamma > 0$ so for $0 < n \leqq n_0$

$$d(p,q) < \gamma \Rightarrow |\overset{n}{T}_{t_0} f(p) - \overset{n}{T}_{t_0} f(q)| < \varepsilon_0,$$

and this in conjunction with the result above gives

For every $\varepsilon>0$ there exists a $\delta>0$ such that

(f) $$|\overset{n}{T}_{t_0} f(p)-\overset{n}{T}_{t_0} f(q)|<\varepsilon \quad \text{if } d(p,q)<\delta.$$

The class $\mathscr{L}$ of functions in $\tilde{C}$ which satisfy condition (2.4) is dense in $\tilde{C}$, so if $g\in\tilde{C}$ and $\varepsilon>0$ are given we can first find $f\in\mathscr{L}$ so

$$\|f-g\|<\frac{\varepsilon}{3},$$

and then choose $\delta>0$ so

$$d(p,q)<\delta \Rightarrow |\overset{n}{T}_{t_0} f(p)-\overset{n}{T}_{t_0} f(q)|<\frac{\varepsilon}{3}$$

for all n. Then we obtain for $d(p,q)<\delta$

$$\begin{aligned}|\overset{n}{T}_{t_0} g(p)-\overset{n}{T}_{t_0} g(q)| &\leqq |\overset{n}{T}_{t_0} g(p)-\overset{n}{T}_{t_0} f(p)|+|\overset{n}{T}_{t_0} f(p)-\overset{n}{T}_{t_0} f(q)|+|\overset{n}{T}_{t_0} f(q)-\overset{n}{T}_{t_0} g(q)|\\ &\leqq 2\|g-f\|+|\overset{n}{T}_{t_0} f(p)-\overset{n}{T}_{t_0} f(q)|<\varepsilon\end{aligned}$$

for all n, which shows that the family $\{\overset{n}{T}_{t_0} g\}_{n=1}^{\infty}$ is equicontinuous and thus finishes the proof of Proposition 2.2.

Next, we turn our attention to condition (c) of Lemma 1.5. Here we have[20]

Proposition 2.3. *Let $f\in C^{(2,\alpha)}\cap C_K$ then the following relations hold*

(a) $$s\text{-}\lim_{n\to\infty} A_n f=\mathscr{A}f=\tfrac{1}{2}\langle D^2 f,\sigma^2\rangle+\langle Df,\mu\rangle-k\cdot f$$

(b) $$s\text{-}\lim_{n\to\infty} \overset{0}{A}_n f=\overset{0}{\mathscr{A}}f=\tfrac{1}{2}\langle D^2 f,\sigma^2\rangle+\langle Df,\mu\rangle.$$

Proposition 2.3 clearly implies condition (c) because the class of C^∞ functions with compact support is dense in C_0 and is contained in $C^{(2,\alpha)}$. To prove the proposition note first that (b) is a special case of (a) so it suffices to show (a). Next

$$\begin{aligned}A_n f(p)&=n\{\overset{n}{T}f(p)-f(p)\}\\ &=n\int_{\mathscr{M}_p}[f(\exp_p Y)-f(p)]\,\tilde{\nu}_p^{(n)}(dY)-n[1-\tilde{\nu}_p^{(n)}(\mathscr{M}_p)]\,f(p)\end{aligned} \tag{2.5}$$

and

$$n[1-\tilde{\nu}_p^{(n)}(\mathscr{M}_p)]=k_n(p)\xrightarrow[n\to\infty]{u}k(p), \tag{2.5'}$$

according to $\mathscr{C}$1, thus we only have to worry about the integral. Here we follow the well-known recipe: First truncate appropriately, then substitute for the difference $f(\exp_p Y)-f(p)$ the first two terms of its Taylor expansion and take limits. So assume f is a given function in $C^{(2,\alpha)}\cap C_K$, then according to Lemma 4.3, for $p\in\mathscr{M}$ and $Y_0\in\mathscr{M}_p$

$$f(\exp_p Y_0)-f(p)=\langle Df, Y_0\rangle_p+\tfrac{1}{2}\langle D^2 f, Y_0\otimes Y_0\rangle_p+R(p,Y_0)$$

where

$$|R(p,Y_0)|\leqq K\|Y_0\|^{2+\alpha}, \qquad K \text{ constant}.$$

[20] $C^{(2,\alpha)}$ is defined in Section 4.2.

Insert

$$Y_0=\frac{1}{n}\mu_{p,n}+\frac{1}{\sqrt{n}}(Y-\mu_{p,n})$$

then

$$\begin{aligned}
&n\left[f\left(\exp_p\left\{\frac{1}{n}\mu_{p,n}+\frac{1}{\sqrt{n}}(Y-\mu_{p,n})\right\}\right)-f(p)\right]\\
&\quad=\langle Df,\mu_{p,n}\rangle+\sqrt{n}\langle Df,Y-\mu_{p,n}\rangle\\
&\quad+\tfrac{1}{2}\langle D^2f,(Y-\mu_{p,n})\otimes(Y-\mu_{p,n})\rangle+\frac{1}{\sqrt{n}}\langle D^2f,\mu_{p,n}\otimes(Y-\mu_{p,n})\rangle\\
&\quad+\frac{1}{2n}\langle D^2f,\mu_{p,n}\otimes\mu_{p,n}\rangle+nR\left(p,\frac{1}{n}\mu_{p,n}+\frac{1}{\sqrt{n}}(Y-\mu_{p,n})\right).
\end{aligned}\tag{2.6}$$

This quantity has to be integrated over $\mathcal{M}_p$. Write[21] $\mathcal{M}_p=G(p,n)\cup G(p,n)^c$ and integrate over each of these sets separately. First

$$\begin{aligned}
&\left|n\int_{G(p,n)^c}\left[f\left(\exp_p\left\{\frac{1}{n}\mu_{p,n}+\frac{1}{\sqrt{n}}(Y-\mu_{p,n})\right\}\right)-f(p)\right]\nu_p^{(n)}(dY)\right|\\
&\qquad\leqq 2\|f\|\,n\,\nu_p^{(n)}(G(p,n)^c)\xrightarrow[n\to\infty]{u}0
\end{aligned}$$

according to (2.2). Next, break the integral over $G(p,n)$ up in accordance with the decomposition (2.6) and estimate each term individually.

Note that because f has compact support

$$L=\sup_p\left[\max\left(\|(Df)_p\|,\|(D^2f)_p\|\right)\right]<\infty.$$

Now

$$\int_{H^1(p,n)}\langle Df,Y\rangle_p\,\nu_p^{(n)}(dY)=\langle Df,\mu_{p,n}\rangle_p$$

consequently

$$\begin{aligned}
\text{(a)}\quad &\left|\int_{G(p,n)}\sqrt{n}\langle Df,Y-\mu_{p,n}\rangle_p\,\nu_p^{(n)}(dY)\right|\\
&=\sqrt{n}\left|-\int_{H^1(p,n)\setminus G(p,n)}\langle Df,Y\rangle_p\,\nu_p^{(n)}(dY)+\langle Df,\mu_{p,n}\rangle_p\cdot\left(1-\nu_p^{(n)}(G(p,n))\right)\right|\\
&\leqq\sqrt{n}\left[\|(Df)_p\|\sqrt{n}\,\nu_p^{(n)}(G(p,n)^c)+\|(Df)_p\|\,\|\mu_{p,n}\|\left(1-\nu_p^{(n)}(G(p,n))\right)\right]\\
&\leqq L\left[\left(1+\frac{1}{\sqrt{n}}M\right)n\,\nu_p^{(n)}(G(p,n)^c)+\frac{1}{\sqrt{n}}Mn\left(1-\nu_p^{(n)}(\mathcal{M}_p)\right)\right]\xrightarrow[n\to\infty]{u}0
\end{aligned}$$

because of (2) and the boundedness of $k(p)$.

$$\begin{aligned}
\text{(b)}\quad &\left|\int_{G(p,n)}\left\{\frac{1}{\sqrt{n}}\langle D^2f,\mu_{p,n}\otimes(Y-\mu_{p,n})\rangle_p+\frac{1}{2n}\langle D^2f,\mu_{p,n}\otimes\mu_{p,n}\rangle_p\right\}\nu_p^{(n)}(dY)\right|\\
&\leqq\left[\frac{1}{\sqrt{n}}L\|\mu_{p,n}\|\left\{\|\mu_{p,n}\|+\delta_n\sqrt{n}\right\}+\frac{1}{2n}L\|\mu_{p,n}\|^2\right]\nu_p^{(n)}(G(p,n))\xrightarrow[n\to\infty]{u}0
\end{aligned}$$

because $\delta_n\downarrow 0$ as $n\to\infty$.

---

[21] The sets $G(p,n)$ and $H^1(p,n)$ are defined in the beginning of this section.

(c) $$\left| n \int_{G(p,n)} R\left(p, \frac{1}{n}\mu_{p,n} + \frac{1}{\sqrt{n}}(Y-\mu_{p,n})\right) \nu_p^{(n)}(dY)\right|$$
$$\leqq nK \int_{G(p,n)} \left\| \frac{1}{n}\mu_{p,n} + \frac{1}{\sqrt{n}}(Y-\mu_{p,n})\right\|^{2+\alpha} \nu_p^{(n)}(dY)$$
$$\leqq K 2^\alpha \left[\left\| \frac{1}{n}\mu_{p,n} - \frac{1}{\sqrt{n}}\mu_{p,n}\right\|^\alpha + \delta_n^\alpha\right] \int_{G(p,n)} \left\| \left(\frac{1}{\sqrt{n}}-1\right)\mu_{p,n} + Y\right\|^2 \nu_p^{(n)}(dY)$$
$$\leqq \varepsilon_n \int_{H^1(p,n)} (M + \|Y\|^2)\, \nu_p^{(n)}(dY),$$

where $\lim_{n\to\infty} \varepsilon_n = 0$.

Also because of (4.7)

$$\int_{H^1(p,n)} \|Y\|^2 \nu_p^{(n)}(dY) \leqq \sqrt{N}\, \|\mu_{p,n}^{(2)}\| \leqq \sqrt{N}\, M$$

so altogether

$$s\text{-}\lim_{n\to\infty} \left| n \int_{G(p,n)} R\left(p, \frac{1}{n}\mu_{p,n} + \frac{1}{\sqrt{n}}(Y-\mu_{p,n})\right) \nu_p^{(n)}(dY)\right| = 0$$

(d) $$\int_{G(p,n)} \langle Df, \mu_{p,n}\rangle_p \nu_p^{(n)}(dY) = \langle Df, \mu_{p,n}\rangle_p \cdot \nu_p^{(n)}(G(p,n)) \xrightarrow[n\to\infty]{u} \langle Df, \mu\rangle_p.$$

(e) $$\tfrac{1}{2} \int_{G(p,n)} \langle D^2 f, (Y-\mu_{p,n}) \otimes (Y-\mu_{p,n})\rangle_p \nu_p^{(n)}(dY)$$
$$= \tfrac{1}{2} \int_{H^1(p,n)} \langle D^2 f, Y\otimes Y - \mu_{p,n}\otimes Y - Y \otimes \mu_{p,n} + \mu_{p,n}\otimes \mu_{p,n}\rangle_p \nu_p^{(n)}(dY) + \rho(p,n)$$
$$= \tfrac{1}{2}\langle D^2 f, \mu_{p,n}^{(2)}\rangle_p - \langle D^2 f, \mu_{p,n}\otimes\mu_{p,n}\rangle_p \cdot \left[1 - \tfrac{1}{2}\nu_p^{(n)}(H^1(p,n))\right] + \rho(p,n),$$

where

$$|\rho(p,n)| \leqq \tfrac{1}{2} L \left[\|\mu_{p,n}\| + \sqrt{n}\right]^2 \cdot \nu_p^{(n)}(G(p,n)^c) \leqq L_1 n\, \nu_p^{(n)}(G(p,n)^c) \xrightarrow[n\to\infty]{u} 0.$$

Moreover

$$s\text{-}\lim_{n\to\infty} \nu_p^{(n)}(H^1(p,n)) = 1$$

so all in all the above expression has the limit

$$\tfrac{1}{2}\langle D^2 f, \mu^{(2)} - \mu\otimes\mu\rangle_p = \tfrac{1}{2}\langle D^2 f, \sigma^2\rangle_p$$

uniformly in p as $n\to\infty$.

Combining (2.5), (2.5′) with (a)–(e) we obtain

$$s\text{-}\lim_{n\to\infty} A_n f = \tfrac{1}{2}\langle D^2 f, \sigma^2\rangle + \langle Df, \mu\rangle - k\cdot f$$

whenever $f \in C^{(2,\alpha)} \cap C_K$, and this finishes the proof of Proposition 2.3.

Let K be a compact subset of $\mathscr{M}$ and let $\varepsilon > 0$ be given. According to axiom $\mathscr{A}$ there exists a C^∞ function f with compact support such that

$$\chi_K \leqq f \leqq \chi_F, \quad F = \mathrm{supp}(f)$$

and

$$\sup_{p\in\mathcal{M}}[\max\{\|(Df)_p\|, \|(D^2f)_p\|\}]\leqq\frac{\varepsilon}{4M}.$$

So it follows from Proposition 2.1 (b) when $p\in F^c$, $t\geqq 0$:

$$\begin{aligned}0&\leqq P_n(t,p,K)\leqq \overset{0}{P}_n(t,p,K)\leqq \overset{n}{S}_t f(p)=\overset{n}{S}_t f(p)-f(p)\\&=\left(\int_0^t \overset{n}{S}_s \overset{0}{A}_n f\,ds\right)(p)\leqq\left\|\int_0^t \overset{n}{S}_s \overset{0}{A}_n f\,ds\right\|\leqq t\|\overset{0}{A}_n f\|\leqq t\{\|\overset{0}{\mathscr{A}}f\|+\|\overset{0}{A}_n f-\overset{0}{\mathscr{A}}f\|\}\\&\leqq t\{\sup_p[\tfrac{1}{2}\|(D^2f)_p\|\cdot\|\sigma_p^2\|+\|(Df)_p\|\cdot\|\mu_p\|]+\|\overset{0}{A}_n f-\overset{0}{\mathscr{A}}f\|\}\\&\leqq t\left\{\frac{\varepsilon}{2}+\|\overset{0}{A}_n f-\overset{0}{\mathscr{A}}f\|\right\}.\end{aligned}$$

Also Proposition 2.3 tells us we can choose n_ε so

$$\|\overset{0}{A}_n f-\overset{0}{\mathscr{A}}f\|<\frac{\varepsilon}{2}\qquad\text{for } n\geqq n_\varepsilon,$$

hence we have

Lemma 2.2. *For every compact subset K of $\mathcal{M}$ and every $\varepsilon>0$ there exists a compact subset F of $\mathcal{M}$ and an integer n_ε such that $K\subset F$ and*

$$0\leqq P_n(t,p,K)\leqq \overset{0}{P}_n(t,p,K)\leqq t\,\varepsilon$$

for $n\geqq n_\varepsilon$, $p\in F^c$, and all $t\geqq 0$.

Now, we can verify (a) of Lemma 1.5. Namely, if this condition were not satisfied then there would exist $f\in C_0$, $\varepsilon>0$, $t>0$, and sequences $n_k\uparrow\infty$, $p_k\to\infty$ so for all k

$$|\overset{n_k}{T}_t f(p_k)|\geqq\varepsilon,$$

where we may assume $f\geqq 0$. Set

$$K=\{p\colon p\in\mathcal{M}, 2f(p)\geqq\varepsilon\},$$

then K is compact because $f\in C_0$. Moreover

$$\varepsilon\leqq \overset{n_k}{T}_t f(p_k)=\int_{\mathcal{M}} f(q)\,P_{n_k}(t,p_k,dq)\leqq\|f\|\cdot P_{n_k}(t,p_k,K)+\frac{\varepsilon}{2};$$

so for all k

$$0<\frac{\varepsilon}{2\|f\|}\leqq P_{n_k}(t,p_k,K).$$

On the other hand, according to Lemma 2.2 there is a compact $F\subset\mathcal{M}$ and an integer n_0 such that

$$P_n(s,p,K)\leqq s\frac{\varepsilon}{4t\|f\|}$$

for $n\geqq n_0$, $p\in F^c$ and all $s\geqq 0$. Now, for k sufficiently large $n_k\geqq n_0$ and $p_k\in F^c$ so for such k

$$0<P_{n_k}(t,p_k,K)\leqq t\frac{\varepsilon}{4t\|f\|}\leqq\frac{1}{2}P_{n_k}(t,p_k,K)$$

which is absurd. Consequently, the sequence $\{(\overset{n}{T}_t, t \geqq 0)\}_{n=1}^{\infty}$ (and by a similar argument also the sequence $\{(\overset{n}{S}_t, t \geqq 0)\}_{n=1}^{\infty}$) must satisfy condition (a).

From Lemma 1.5 it now follows that every subsequence $\{m_l\}_{l=1}^{\infty}$ of $\{n\}_{n=1}^{\infty}$ has a further subsequence $\{n_k\}_{k=1}^{\infty}$ such that for every $f \in C_0$ and $t \geqq 0$ the limits

$$T_t f = s\text{-}\lim_{k \to \infty} \overset{n_k}{T}_t f; \qquad S_t f = s\text{-}\lim_{k \to \infty} \overset{n_k}{S}_t f \tag{2.7}$$

exist. The families $(T_t, t \geqq 0)$ and $(S_t, t \geqq 0)$ are C_0-semigroups with generators A and $\overset{0}{A}$ whose domains contain $\mathscr{D} = C^{(2,\alpha)} \cap C_K$ and which satisfy

$$Af = \mathscr{A} f, \qquad \overset{0}{A} f = \overset{0}{\mathscr{A}} f \quad \text{when } f \in \mathscr{D},$$

where $\mathscr{A}$ and $\overset{0}{\mathscr{A}}$ are the differential operators defined by (a) and (b) in Proposition 2.3. So in particular we have shown the existence of C_0-semigroups whose generators, when properly restricted, coincide with $\mathscr{A}$ and $\overset{0}{\mathscr{A}}$ without assuming that σ^2 be strictly positive definite. However, the limits obtained might depend upon the subsequences chosen and in particular the original sequences might not themselves converge. To show that this unpleasant situation does not occur we need the extra condition upon σ^2, which, as will be shown below, implies that the limits we get are independent of the particular subsequences from which they are derived.

An elementary argument based upon this observation then yields the result that for every $f \in C_0$ and $t \geqq 0$

$$s\text{-}\lim_{n \to \infty} \overset{n}{T}_t f = T_t f; \qquad s\text{-}\lim_{n \to \infty} \overset{n}{S}_t f = S_t f.$$

However, before we can prove this we need several auxiliary results concerning the behavior of the operators A and $\overset{0}{A}$ and of the sequences

$$\{(\overset{n_k}{T}_t, t \geqq 0)\}_{k=1}^{\infty} \quad \text{and} \quad \{(\overset{n_k}{S}_t, t \geqq 0)\}_{k=1}^{\infty}.$$

First

Lemma 2.3. *For arbitrary $f \in C_0$, $f \geqq 0$, $t \geqq 0$, $\Gamma \in \mathscr{B}(\mathscr{M})$, $p \in \mathscr{M}$ the following inequalities hold*

(a) $$0 \leqq T_t f \leqq S_t f$$

(b) $$P(t, p, \Gamma) \leqq \overset{0}{P}(t, p, \Gamma).$$

Here P and $\overset{0}{P}$ are the transition functions corresponding to $(T_t, t \geqq 0)$ resp. $(S_t, t \geqq 0)$. The proof of the lemma is simple. Namely, (a) is a consequence of (a) in Proposition 2.1 while (b) follows from (a) of the present lemma. One might note, that intuitively this result and its counterpart in Proposition 2.1 are obvious because the process corresponding to $(\overset{n}{T}_t, t \geqq 0)$ has a positive probability of being "killed" while the process corresponding to $(\overset{n}{S}_t, t \geqq 0)$ is conservative. In this connection also note that although we know $\overset{0}{P}_n(t, p, \mathscr{M}) \equiv 1$ for all n we do not yet have a similar result for $\overset{0}{P}(t, p, \mathscr{M})$.

In the following let $B(\varepsilon, p)$ denote the ball in $\mathscr{M}$ with center p and radius ε, and let $B^c(\varepsilon, p)$ denote its complement (in $\mathscr{M}$) then

Proposition 2.4. *For every compact set $\Gamma \subset \mathcal{M}$ and every pair of positive numbers (ε, η) there exists an integer k_0 and a number $\gamma > 0$ such that*

(a)
$$\sup_{\substack{0<t\leqq\gamma \\ p\in\Gamma}} \frac{1}{t}\{1-\overset{\circ}{P}_{n_k}(t,p,B(\varepsilon,p))\} \leqq \eta$$

(b)
$$\sup_{\substack{0<t\leqq\gamma \\ p\in\Gamma}} \frac{1}{t} P_{n_k}(t,p,B^c(\varepsilon,p)) \leqq \eta$$

when $k \geqq k_0$.

Corollary. *For every compact set $\Gamma \subset \mathcal{M}$ and every pair of positive numbers (ε, η) there exists a number $\gamma > 0$ such that*

(c)
$$\sup_{\substack{0<t\leqq\gamma \\ p\in\Gamma}} \frac{1}{t}\{1-\overset{\circ}{P}(t,p,B(\varepsilon,p))\} \leqq \eta.$$

(d)
$$\sup_{\substack{0<t\leqq\gamma \\ p\in\Gamma}} \frac{1}{t} P(t,p,B^c(\varepsilon,p)) \leqq \eta.$$

Proof of Corollary. If $F \subset \mathcal{M}$ is compact then (2.7) implies

$$\overline{\lim_{k\to\infty}} \overset{\circ}{P}_{n_k}(t,p,F) \leqq \overset{\circ}{P}(t,p,F).$$

Hence

$$\begin{aligned} P(t,p,B^c(\varepsilon,p)) &\leqq \overset{\circ}{P}(t,p,B^c(\varepsilon,p)) \\ &\leqq 1-\overset{\circ}{P}(t,p,B(\varepsilon,p)) \leqq 1-\overset{\circ}{P}\left(t,p,\overline{B\left(\frac{\varepsilon}{2},p\right)}\right) \\ &\leqq 1-\overline{\lim_{k\to\infty}}\overset{\circ}{P}_{n_k}\left(t,p,\overline{B\left(\frac{\varepsilon}{2},p\right)}\right) = \underline{\lim}_{k\to\infty}\left\{1-\overset{\circ}{P}_{n_k}\left(t,p,\overline{B\left(\frac{\varepsilon}{2},p\right)}\right)\right\}, \end{aligned}$$

and an application of Proposition 2.4(a) immediately yields the desired result.

In order to prove Proposition 2.4 note first that statement (b) follows from statement (a) and Lemma 2.3(b), so it suffices to prove (a)[22]. Now, if (a) were not true, there would exist a compact subset Γ_0 of $\mathcal{M}$, numbers $\varepsilon_0 > 0$, $\eta_0 > 0$, and sequences $n'_j \to \infty$, $t_j \downarrow 0$, $\{q_j\}_{j=1}^{\infty} \subset \Gamma_0$ so for all j

$$1-\overset{\circ}{P}_{n'_j}(t_j,q_j,B(\varepsilon_0,q_j)) > t_j\eta_0.$$

Because of the compactness of Γ_0 it may be assumed that $q_j \to q_0 \in \Gamma_0$ for $j \to \infty$. Hence, for j sufficiently large

$$B(\tfrac{1}{2}\varepsilon_0,q_0) \subset B(\varepsilon_0,q_j)$$

and thus for such j

$$1-\overset{\circ}{P}_{n'_j}(t_j,q_j,B(\tfrac{1}{2}\varepsilon_0,q_0)) > t_j\eta_0.$$

[22] The idea behind the proof is from [6] (p. 93–95).

Choose next $f \in C_K^\infty$ such that

$$\chi_{B(\frac{1}{4}\varepsilon_0,\, q_0)} \leqq f \leqq \chi_{B(\frac{1}{2}\varepsilon_0,\, q_0)}$$

then

$$\overset{n'_j}{S}_{t_j} f(p) = \int_{\mathcal{M}} f(q)\, \mathring{P}_{n'_j}(t_j, p, dq) \leqq \mathring{P}_{n'_j}(t_j, p, B(\tfrac{1}{2}\varepsilon_0, q_0)),$$

so

$$t_j \eta_0 < 1 - \overset{n'_j}{S}_{t_j} f(q_j) = 1 - S_{t_j} f(q_j) + S_{t_j} f(q_j) - \overset{n'_j}{S}_{t_j} f(q_j).$$

Also

$$0 = \mathring{A} f(q_0) = \lim_{t \downarrow 0} \frac{1}{t} [S_t f(q_0) - f(q_0)]$$

$$= \lim_{j \to \infty} \frac{1}{t_j} [S_{t_j} f(q_j) - f(q_j)] = \lim_{j \to \infty} \frac{1}{t_j} [S_{t_j} f(q_j) - 1],$$

because $\frac{1}{t}[S_t f - f]$ converges uniformly to $\mathring{A} f$; thus for j sufficiently large

$$t_j \eta_0 \leqq 2\, |S_{t_j} f(q_j) - \overset{n'_j}{S}_{t_j} f(q_j)|. \tag{2.8}$$

Let $\tau = \max_j t_j$ and choose [23] j_1 so for $j \geqq j_1$

$$\sup_{0 \leqq s \leqq \tau} \|\overset{n'_j}{S}_s(\mathring{A} f) - S_s(\mathring{A} f)\| < \tfrac{1}{8}\eta_0,$$

and next choose $j_2 \geqq j_1$ so for $j \geqq j_2$

$$\|\mathring{A}_{n'_j} f - \mathring{A} f\| < \tfrac{1}{8}\eta_0,$$

then if $j \geqq j_2$

$$\|\overset{n'_j}{S}_{t_j} f - S_{t_j} f\| = \left\| \int_0^{t_j} [\overset{n'_j}{S}_s [\mathring{A}_{n'_j} f - \mathring{A} f] + [\overset{n'_j}{S}_s - S_s](\mathring{A} f)]\, ds \right\|$$

$$\leqq t_j \|\mathring{A}_{n'_j} f - \mathring{A} f\| + t_j \cdot \sup_{0 \leqq s \leqq \tau} \|\overset{n'_j}{S}_s(\mathring{A} f) - S_s(\mathring{A} f)\| < \tfrac{1}{4} t_j \eta_0,$$

which in connection with (2.8) yields, when j is sufficiently large,

$$\tfrac{1}{2} t_j \eta_0 \leqq |S_{t_j} f(q_j) - \overset{n'_j}{S}_{t_j} f(q_j)| \leqq \|S_{t_j} f - \overset{n'_j}{S}_{t_j} f\| < \tfrac{1}{4} t_j \eta_0.$$

This is a contradiction and consequently we may conclude that (a) is true.

Note, that once it has been shown that $\{(\overset{n}{T}_t,\ t \geqq 0)\}_{n=1}^\infty$ rather than the subsequence $\{(\overset{n_k}{T}_t,\ t \geqq 0)\}_{k=1}^\infty$ converges then (a) and (b) become valid for the P_n, $\mathring{P}_n$ rather than for the P_{n_k}, $\mathring{P}_{n_k}$.

Next we shall look a little more closely at the operators A and $\mathring{A}$ [24]. Recall the remark following the proof of Lemma 1.5 according to which the actual convergence of the sequence $\{(\overset{n}{T}_t,\ t \geqq 0)\}_{n=1}^\infty$ would follow from what we already

---

[23] Use Lemma 1.6.

[24] On the next few pages we shall work only with the operators T_t, A, $\mathcal{A}$. The results we obtain will of course correspond to similar results relating to the operators S_t, $\mathring{A}$, $\mathring{\mathcal{A}}$.

know by now provided it is true that the set

$$\mathcal{R}=\{f\colon\ f\in C_0,\ f=g-Ag \text{ for some } g\in\mathcal{D}\}$$

is dense in C_0. Also when $g\in\mathcal{D}$ then $Ag=\mathcal{A}g$, so essentially we are faced with the problem of having to solve the differential equation $f=g-\mathcal{A}g$, and to show that the solution g is contained in $\mathcal{D}$ for f ranging within a dense subset of C_0. For general manifolds $\mathcal{M}$ we cannot handle this problem directly. So instead we proceed by solving the equation locally first, and then use the known properties of the generator A in order to glue the local solutions together. The method is lengthy and will be broken into several parts each of which will be formulated as a lemma. First we prove A is a local operator.

Lemma 2.4. *If $f\in\mathcal{D}_A$, $g\in\mathcal{D}_A$ and $f\equiv g$ on a neighborhood of the point p_0 then*

$$Af(p_0)=Ag(p_0).$$

Here $\mathcal{D}_A$ denotes the domain of A. To prove the lemma, let $h\in\mathcal{D}_A$ with $h(p)\equiv 0$ for $p\in B(r,p_0)$, then

$$\begin{aligned}|Ah(p_0)|&=\left|\lim_{t\downarrow 0}\frac{1}{t}\Big[\int_{\mathcal{M}} h(q)\,P(t,p_0,dq)-h(p_0)\Big]\right|\\&=\left|\lim_{t\downarrow 0}\frac{1}{t}\int_{B^c(r,p_0)} h(q)\,P(t,p_0,dq)\right|\le\|h\|\lim_{t\downarrow 0}\frac{1}{t}P(t,p_0,B^c(r,p_0))=0\end{aligned}$$

according to the corollary to Proposition 2.4. By setting $h=f-g$ we obtain the statement of the lemma.

Lemma 2.5. *If $f\in\mathcal{D}_A\cap C^{(2,\alpha)}_{\mathrm{loc}}$ then for each $p\in\mathcal{M}$*

$$Af(p)=\mathcal{A}f(p).$$

Proof. Let $f\in\mathcal{D}_A\cap C^{(2,\alpha)}_{\mathrm{loc}}$ and let $p_0\in\mathcal{M}$. Then, according to Lemma 4.2, there is a function $f_0\in C^{(2,\alpha)}\cap C_K$ such that $f_0(p)=f(p)$ for $p\in B(r_0,p_0)$ for some $r_0>0$, and because $f_0\in\mathcal{D}\subset\mathcal{D}_A$ Lemma 2.4 yields

$$Af(p_0)=Af_0(p_0)=\mathcal{A}f_0(p_0)=\mathcal{A}f(p_0).$$

Because A is a local operator it makes sense to talk about Af even when f is not defined everywhere on $\mathcal{M}$. To be specific, let $\mathcal{G}$ be an open subset of $\mathcal{M}$ and let f be a function which is bounded and continuous on $\mathcal{G}$. If now, for every point $p\in\mathcal{G}$ we can find a neighborhood V_p of p with $V_p\subset\mathcal{G}$ and a function $g\in\mathcal{D}_A$ such that $g\equiv f$ on V_p then we may set

$$Af(p)=Ag(p).$$

We denote the class of such functions f by the symbol $\mathcal{D}_A^{\mathcal{G}}$. In case $\mathcal{M}$ is compact $\mathcal{D}_A=\mathcal{D}_A^{\mathcal{M}}$, but in general $\mathcal{D}_A$ is a proper subset of $\mathcal{D}_A^{\mathcal{M}}$. Note, that if $f\in\mathcal{D}_A^{\mathcal{G}}$ and $p\in\mathcal{G}$ then

$$Af(p)=\lim_{t\downarrow 0}\frac{1}{t}\Big\{\int_{\mathcal{G}} f(q)\,P(t,p,dq)-f(p)\Big\}.$$

This follows immediately from the proof of Lemma 2.4.

Lemma 2.6. *Let $\mathcal{G} \subset \mathcal{M}$ be open and let f be a bounded continuous function on $\mathcal{G}$ which (on $\mathcal{G}$) belongs to class $C_{\text{loc}}^{(2,\alpha)}$. Then*

$$f \in \mathcal{D}_A^{\mathcal{G}} \quad \text{and} \quad Af(p) = \mathscr{A} f(p) \quad \text{for } p \in \mathcal{G}.$$

Proof. Let f be given and satisfy the conditions of the lemma, and let $p_0 \in \mathcal{G}$. As in the proof of Lemma 2.5 we can find a function f_0 with compact support in $\mathcal{G}$ such that $f_0 \in C^{(2,\alpha)}$ and $f_0(p) \equiv f(p)$ on a neighborhood of p_0; but then

$$f_0 \in \mathcal{D}_A \quad \text{and} \quad Af_0(p_0) = \mathscr{A} f_0(p_0),$$

which proves the lemma.

Lemma 2.7. *Let $\mathcal{G}$ be a bounded*[25] *open subset of $\mathcal{M}$ and let f be a function which is continuous on $\mathcal{G} \cup \partial\mathcal{G}$ and whose restriction to $\mathcal{G}$ is contained in $\mathcal{D}_A^{\mathcal{G}}$. If for some $\lambda > 0$, f satisfies the equation*

$$\lambda f = Af \text{ on } \mathcal{G}, \quad f = 0 \text{ on } \partial\mathcal{G},$$

then $f \equiv 0$ on $\mathcal{G} \cup \partial\mathcal{G}$.

Proof. If f does not vanish identically on $\mathcal{G}$ it either has a positive maximum or a negative minimum. Assume the former, and choose $p_0 \in \mathcal{G}$ so $f(p_0)$ is maximum then

$$Af(p_0) = \lim_{t \downarrow 0} \frac{1}{t} \{ \int_{\mathcal{G}} f(q)\, P(t, p_0, dq) - f(p_0) \} \leqq 0$$

$$0 = \lambda f(p_0) - Af(p_0) \geqq \lambda f(p_0) > 0,$$

which is a contradiction; thus $f \leqq 0$. But the assumption that f has a negative minimum also leads to a contradiction (consider $-f$) consequently we must have $f \equiv 0$.

Let $(R_\lambda, \lambda > 0)$ be the resolvent of the semigroup $(T_t, t \geqq 0)$. Then

Lemma 2.8. *Let $f \in C_0$ be Hölder continuous of order α, then*

$$R_\lambda f \in C_{\text{loc}}^{(2,\alpha)} \cap \mathcal{D}_A$$

when $\lambda > 0$.

Proof[26]. Let $\lambda > 0$ and $p_0 \in \mathcal{M}$ be given. Choose a coordinate system $(V, \varphi)(V \subset \mathcal{M}, \varphi\colon V \to E^N)$ on a neighborhood V of p_0 so the following conditions are satisfied

(a) $F(\bar{x}) = f \circ \varphi^{-1}(\bar{x})$ is Hölder continuous of order α on $\varphi(V) = U$.

(b) The coefficients in the coordinate expression for the operator $\mathscr{A}$:

$$\tilde{\mathscr{A}} G(\bar{x}) = a^{i,j}(\bar{x})\, \partial_i \partial_j G(\bar{x}) + b^k(\bar{x})\, \partial_k G(\bar{x}) + c(\bar{x})\, G(\bar{x})$$

are Hölder continuous of order α.

Let B be a ball (Euclidean metric) contained in U and with center $\varphi(p_0)$, and consider the differential equation

$$(\lambda - \tilde{\mathscr{A}})\, G(\bar{x}) = F(\bar{x}) \qquad \text{for } \bar{x} \in B,$$

$$G(\bar{x}) = (R_\lambda f)(\varphi^{-1}(\bar{x})) \quad \text{for } \bar{x} \in \partial B.$$

[25] That is, $\bar{\mathcal{G}}$ is compact.

[26] See also [6] p. 158.

According to the general theory [27] of such equations, there is a uniquely determined solution $G(\bar{x})$ which is continuous on $\bar{B}$ and of class $C^{(2,\alpha)}$ on B. Set $g(p)=G(\varphi(p))$, then g is of class $C^{(2,\alpha)}$ on $\varphi^{-1}(B)=V_0\subset V$ and continuous on $\bar{V}_0$. Also

$$\begin{aligned}(\lambda-\mathscr{A})\,g(p)&=f(p) &&\text{for } p\in V_0,\\ g(p)&=R_\lambda f(p) &&\text{for } p\in\partial V_0.\end{aligned}$$

Next, Lemma 2.6 yields

$$g\in\mathscr{D}_A^{V_0} \quad\text{and}\quad \mathscr{A}g=Ag \quad\text{on } V_0,$$

hence

$$f(p)=(\lambda-\mathscr{A})\,g(p)=(\lambda-A)\,g(p), \qquad p\in V_0.$$

On the other hand, if h is the restriction of $R_\lambda f$ to V_0 then

$$h\in\mathscr{D}_A^{V_0}, \quad\text{and}\quad (\lambda-A)\,h(p)=(\lambda-A)\,R_\lambda f(p)=f(p)$$

for $p\in V_0$, which by Lemma 2.7 implies $h(p)=g(p)$ on V_0.

Thus $R_\lambda f$ agrees on V_0 with a function in $C^{(2,\alpha)}$ and consequently $R_\lambda f\in C^{(2,\alpha)}_{\mathrm{loc}}$.

Now we can finally complete the proof of the convergence theorem. Let $\{(\overset{n_k}{T}_t,\, t\geqq 0)\}$ and $\{(\overset{n'_k}{T}_t,\, t\geqq 0)\}$ be two convergent subsequences for our original sequence with limits $(T_t,\, t\geqq 0)$ resp. $(T'_t,\, t\geqq 0)$, the corresponding generators and resolvents being denoted by A, A' resp. R_λ, R'_λ. Choose a function $f\in C_0$ which is Hölder continuous of order α, then according to Lemma 2.8 and Lemma 2.5 we have for $\lambda>0$

$$\begin{aligned}R_\lambda f&\in C^{(2,\alpha)}_{\mathrm{loc}}\cap\mathscr{D}_A &&\text{and}\quad AR_\lambda f=\mathscr{A}R_\lambda f,\\ R'_\lambda f&\in C^{(2,\alpha)}_{\mathrm{loc}}\cap\mathscr{D}_{A'} &&\text{and}\quad A'R'_\lambda f=\mathscr{A}R'_\lambda f.\end{aligned}$$

Hence

$$(\lambda-\mathscr{A})\,R_\lambda f=(\lambda-A)\,R_\lambda f=f=(\lambda-A')\,R'_\lambda f=(\lambda-\mathscr{A})\,R'_\lambda f$$

or

$$(\lambda-\mathscr{A})(R_\lambda f-R'_\lambda f)=0;$$

moreover

$$\lim_{p\to\infty} R_\lambda f(p)=\lim_{p\to\infty} R'_\lambda f(p)=0$$

and these two results taken together imply [28] $R_\lambda f=R'_\lambda f$. In other words, we have shown

$$R_\lambda f=R'_\lambda f,$$

for $\lambda>0$ and f ranging within a dense subset of C_0. But this in turn implies

$$T_t f=T'_t f, \qquad \text{for } t\geqq 0 \text{ and } f\in C_0,$$

so whenever a subsequence $\{(\overset{n'_k}{T}_t,\, t\geqq 0)\}_{k=1}^\infty$ converges it must converge to $(T_t,\, t\geqq 0)$.

On the other hand, we know already that any prescribed sequence $\{(\overset{n_k}{T}_t,\, t\geqq 0)\}_{k=1}^\infty$ has a convergent subsequence. Combining these results we may then conclude that the original sequence $\{(\overset{n}{T}_t,\, t\geqq 0)\}_{n=1}^\infty$ itself converges to $(T_t,\, t\geqq 0)$. Thus we have

[27] [17] p. 167, Theorem 36 IV.

[28] Argue as in the proof of Lemma 2.7.

Theorem 2.1. *Let $\mathcal{M}$ be a complete Riemannian manifold of class $\mathcal{A}\mathcal{B}$ and let $\{(\nu_p^{(n)})_{p\in\mathcal{M}}\}_{n=1}^{\infty}$ be a sequence of families of subprobability measures on the tangent spaces $\mathcal{M}_p$ satisfying conditions $\mathcal{C}0$–$\mathcal{C}6$. Then the corresponding C_0-semigroups $(\overset{n}{T}_t, t\geqq 0)$ and $(\overset{n}{S}_t, t\geqq 0)$ satisfy for $f\in C_0$, $t\geqq 0$*

$$s\text{-}\lim_{n\to\infty} \overset{n}{T}_t f = T_t f, \qquad s\text{-}\lim_{n\to\infty} \overset{n}{S}_t f = S_t f,$$

where $(T_t, t\geqq 0)$ and $(S_t, t\geqq 0)$ are C_0-semigroups whose generators A and $\overset{0}{A}$ when applied to a function f in $C^{(2,\alpha)}\cap C_K$ are determined by

$$A f = \mathcal{A} f = \tfrac{1}{2}\langle D^2 f, \sigma^2\rangle + \langle D f, \mu\rangle - k\cdot f,$$
$$\overset{0}{A} f = \overset{0}{\mathcal{A}} f = \tfrac{1}{2}\langle D^2 f, \sigma^2\rangle + \langle D f, \mu\rangle.$$

The corresponding transition functions satisfy the inequality

$$P(t,p,\Gamma)\leqq \overset{0}{P}(t,p,\Gamma) \quad \text{for } (t,p,\Gamma)\in(E^1)^+\times\mathcal{M}\times\mathcal{B}(\mathcal{M}).$$

Remark. On basis of Lemma 1.6 we can make the apparently stronger statements, that for $f\in C_0$ and $t_0>0$

(a) $$\lim_{n\to\infty}\sup_{0\leqq t\leqq t_0} \|\overset{n}{T}_t f - T_t f\| = 0$$

(b) $$\lim_{n\to\infty}\sup_{0\leqq t\leqq t_0} \|\overset{n}{S}_t f - S_t f\| = 0.$$

Corollary. *Let the conditions of Theorem 2.1 be satisfied and let $\{\alpha_n\}_{n=1}^{\infty}$ be a sequence of integers for which*

$$\lim_{n\to\infty}\frac{1}{n}\alpha_n = t>0.$$

Then for $f\in C_0$

$$s\text{-}\lim_{n\to\infty} \overset{n}{T}{}^{\alpha_n} f = T_t f, \qquad s\text{-}\lim_{n\to\infty} \overset{n}{S}{}^{\alpha_n} f = S_t f.$$

Recall that

$$\overset{n}{T} f(p) = \int_{\mathcal{M}} f(q)\, P_n(p,dq); \qquad \overset{n}{S} f(p) = \int_{\mathcal{M}} f(q)\, \overset{0}{P}_n(p,dq).$$

The corollary states that the iterates

$$P_n^{(\alpha_n)}(p,\cdot) \quad\text{and}\quad \overset{0}{P}{}_n^{(\alpha_n)}(p,\cdot)$$

converge vaguely to the measures $P(t,p,\cdot)$ and $\overset{0}{P}(t,p,\cdot)$ respectively. For the proof it suffices to consider the $(\overset{n}{T})$ only. So let the sequence $\{\alpha_n\}_{n=1}^{\infty}$ with

$$\lim_{n\to\infty}\frac{1}{n}\alpha_n = t$$

be given, and let $f\in C_0$. If $0<\varepsilon<1$ we have

$$e^{-nt}\sum_{|k-nt|>\varepsilon nt}\frac{(nt)^k}{k!}\leqq\frac{nt}{\varepsilon^2(nt)^2}=\frac{1}{\varepsilon^2 nt}$$

according to Chebyshev's inequality, hence

$$\begin{aligned}\|\overset{n}{T}{}^{\alpha_n}f-\overset{n}{T}_t f\| &\leqq e^{-nt}\sum_{k=0}^{\infty}\frac{(nt)^k}{k!}\|\overset{n}{T}{}^{\alpha_n}f-\overset{n}{T}{}^k f\|\\ &\leqq \frac{2\|f\|}{\varepsilon^2(nt)}+e^{-nt}\sum_{|1-\frac{k}{nt}|\leqq\varepsilon}\frac{(nt)^k}{k!}\|\overset{n}{T}{}^{\alpha_n}f-\overset{n}{T}{}^k f\|.\end{aligned}\tag{2.9}$$

Now, if $l>k$

$$\overset{n}{T}{}^l f-\overset{n}{T}{}^k f=\overset{n}{T}{}^k\left\{\sum_{j=0}^{l-k-1}\overset{n}{T}{}^j(\overset{n}{T}f-f)\right\}=\frac{1}{n}\overset{n}{T}{}^k\left\{\sum_{j=0}^{l-k-1}\overset{n}{T}{}^j(A_n f)\right\}.$$

thus

$$\|\overset{n}{T}{}^l f-\overset{n}{T}{}^k f\|\leqq\frac{l-k}{n}\|A_n f\|$$

which after insertion in (2.9) yields

$$\begin{aligned}\|\overset{n}{T}{}^{\alpha_n}f-\overset{n}{T}_t f\| &\leqq \frac{2\|f\|}{\varepsilon^2 nt}+t\|A_n f\|\,e^{-nt}\sum_{|1-\frac{k}{nt}|\leqq\varepsilon}\frac{(nt)^k}{k!}\left|\frac{\alpha_n}{nt}-\frac{k}{nt}\right|\\ &\leqq \frac{2\|f\|}{\varepsilon^2 nt}+t\|A_n f\|\cdot\left\{\varepsilon+\left|1-\frac{\alpha_n}{nt}\right|\right\},\end{aligned}$$

hence

$$\overline{\lim_{n\to\infty}}\|\overset{n}{T}{}^{\alpha_n}f-\overset{n}{T}_t f\|\leqq\varepsilon t\,\overline{\lim_{n\to\infty}}\|A_n f\|.$$

Also, the left-hand side is independent of ε, thus it must be zero whenever the right-hand term is finite, in particular it must be zero when $f\in\mathscr{D}$.

But $\mathscr{D}=C_K\cap C^{(2,\alpha)}$ is dense in C_0 and so we may conclude

$$\lim_{n\to\infty}\|\overset{n}{T}{}^{\alpha_n}f-\overset{n}{T}_t f\|=0$$

for all $f\in C_0$. This immediately implies the statement of the corollary.

Theorem 2.2. *Assume the conditions listed in Theorem* 2.1 *are satisfied, and let* $(X_t^{(n)}, t\geqq 0)$, $(X_t^{(0)}, t\geqq 0)$ *be standard processes on* $\mathscr{M}_\Delta$ *associated with the semigroups* $(\overset{n}{T}_t, t\geqq 0)$ *resp.* $(T_t, t\geqq 0)$ *and satisfying: weak-*$\lim_{n\to\infty} X_0^{(n)}=X_0^{(0)}$; *then the sequence of processes* $\{(X_t^{(n)}, t\geqq 0)\}_{n=1}^{\infty}$ *converges weakly to the process* $(X_t^{(0)}, t\geqq 0)$.

In order to prove this theorem we need according to Lemma 1.4 only show that for every $\varepsilon>0$ there exists an $\alpha>0$ and an integer n_ε so

$$\sup_{\substack{p\in\mathscr{M}_\Delta\\ 0<t}}\frac{1}{t}\tilde{P}_n(t,p,B_\Delta^c(\varepsilon,p))\leqq\alpha,\qquad\text{for }n\geqq n_\varepsilon,\tag{2.10}$$

where $B_\Delta^c(\varepsilon,p)$ is the complement (in $\mathscr{M}_\Delta$) of the ball

$$B_\Delta(\varepsilon,p)=\{q: q\in\mathscr{M}_\Delta, \delta(p,q)<\varepsilon\}.$$

Also we retain the notation $B(\varepsilon, p)$ for the ball

$$B(\varepsilon, p)=\{q: q\in\mathcal{M}, d(p,q)<\varepsilon\},$$

and let $B^c(\varepsilon, p)$ denote the complement in $\mathcal{M}$ of $B(\varepsilon, p)$.

The proof of (2.10) will be based upon the following

Lemma 2.9. *For every compact subset Γ of $\mathcal{M}$ and every $\varepsilon>0$ there exists constants a, b such that for all $t>0$*

$$\sup_n \sup_{p\in\Gamma}\big(1-P_n(t,p,\mathcal{M})\big)\leqq at, \tag{2.11}$$

$$\overline{\lim_{n\to\infty}} \sup_{p\in\Gamma} P_n\big(t,p,B^c(\varepsilon,p)\big)\leqq bt. \tag{2.12}$$

Proof of Lemma. According to Proposition 2.4(b) there exists an integer n_0 and a number $\gamma>0$ so

$$\sup_{p\in\Gamma} P_n\big(t,p,B^c(\varepsilon,p)\big)\leqq 1\cdot t$$

when $n\geqq n_0$ and $0\leqq t\leqq\gamma$. Set $b=\max\left(\frac{1}{\gamma},1\right)$ then $bt>1$ for $t\geqq\gamma$, consequently

$$\sup_{p\in\Gamma} P_n\big(t,p,B^c(\varepsilon,p)\big)\leqq bt$$

for $n\geqq n_0$ and all t. This proves (2.12). Next, choose $p_0\in\mathcal{M}$ and define the sequence

$$F_n=\{p: p\in\mathcal{M}, d(p_0,p)\leqq n\}, \qquad n=1,2,\ldots.$$

Each F_n is closed and bounded, hence compact. Also $F_n\uparrow\mathcal{M}$ as $n\uparrow\infty$. Select functions $f_n\in C_K^\infty(\mathcal{M})$ which satisfy

$$\chi_{F_n}\leqq f_n\leqq 1;\qquad \sup_{p\in\mathcal{M}}[\|(Df)_p\|+\|(D^2f)_p\|]\leqq 1,$$

this can be done according to axiom $\mathcal{A}$. Write K_n for the support of f_n, then K_n is compact and $F_n\subset K_n$. Also, because Γ is compact, there exists an n_0 such that $\Gamma\subset F_{n_0}$. Then, for $p\in\Gamma$ we have

$$\begin{aligned} 1-P_l(t,p,\mathcal{M}) &\leqq 1-P_l(t,p,K_{n_0})\leqq f_{n_0}(p)-\overset{l}{T}_t f_{n_0}(p)\\ &=\left[-\int_0^t (\overset{l}{T}_s A_l f_{n_0})\,ds\right](p)\leqq t\cdot\|A_l f_{n_0}\|\leqq t\cdot\{\|Af_{n_0}\|+\|A_l f_{n_0}-Af_{n_0}\|\}\\ &\leqq t\cdot\{\sup_{p\in\mathcal{M}}[\|\sigma_p^2\|+\|\mu_p\|]+\|k\|+\|A_l f_{n_0}-Af_{n_0}\|\}\leqq at, \end{aligned}$$

with a independent of p and l, and this proves (2.11).

As an immediate consequence of the above proof we also obtain

Corollary.

$$\sup_{\substack{p\in\mathcal{M}\\ 0<t}}\frac{1}{t}\{1-P(t,p,\mathcal{M})\}<\infty.$$

To see this, let $p \in \mathcal{M}$ and choose n_0 so $p \in F_{n_0}$, then

$$\begin{aligned}1-P(t,p,\mathcal{M}) &\leqq 1-P(t,p,K_{n_0})\\ &\leqq 1-\overline{\lim_{l\to\infty}} P_l(t,p,K_{n_0}) = \underline{\lim}_{l\to\infty}[1-P_l(t,p,K_{n_0})]\\ &\leqq t\cdot\{\|k\|+\sup_{q\in\mathcal{M}}[\|\sigma_q^2\|+\|\mu_q\|]\}\leqq a_0 t,\end{aligned}$$

where a_0 does not depend upon p, thus

$$\sup_{\substack{p\in\mathcal{M}\\ 0<t}} \frac{1}{t}\{1-P(t,p,\mathcal{M})\}\leqq a_0.$$

Now, we are able to prove (2.10). Let $\varepsilon>0$ be given and define

$$\Gamma_1=\left\{p: p\in\mathcal{M},\ \delta(p,\Delta)\geqq\frac{2\varepsilon}{3}\right\},\qquad \Gamma_2=\left\{p: p\in\mathcal{M},\ \delta(p,\Delta)\geqq\frac{\varepsilon}{3}\right\}.$$

These sets are compact in $\mathcal{M}$ and

$$\Gamma_1\subset\operatorname{interior}(\Gamma_2);$$

also we may assume that ε is so small that Γ_1 is non-empty. If $p\in\Gamma_2^c$ and $q\in B_\Delta^c(\varepsilon,p)$ then

$$\delta(q,\Delta)\geqq\delta(q,p)-\delta(\Delta,p)>\varepsilon-\frac{\varepsilon}{3}=\frac{2\varepsilon}{3}$$

so $q\in\Gamma_1$ and consequently

$$\sup_{p\in\Gamma_2^c}\tilde{P}_n(t,p,B_\Delta^c(\varepsilon,p))\leqq\sup_{p\in\Gamma_2^c}\tilde{P}_n(t,p,\Gamma_1)=\sup_{p\in\Gamma_2^c}P_n(t,p,\Gamma_1).$$

Choose $f\in C_K^\infty(\mathcal{M})$ such that

$$\chi_{\Gamma_1}\leqq f\leqq\chi_{\Gamma_2}$$

then for $p\in\Gamma_2^c$

$$\begin{aligned}P_n(t,p,\Gamma_1)&\leqq \overset{n}{T}_t f(p)=\overset{n}{T}_t f(p)-f(p)=\left(\int_0^t \overset{n}{T}_s A_n f\,ds\right)(p)\\ &\leqq t\cdot\|A_n f\|\leqq ct,\quad\text{with}\quad c=\sup_n\|A_n f\|,\end{aligned}$$

thus

$$\sup_{p\in\Gamma_2^c}\tilde{P}_n(t,p,B_\Delta^c(\varepsilon,p))\leqq ct,\qquad\text{for all } n,t>0. \tag{2.13}$$

Next, according to Lemma 1.1, there exists a $\rho>0$ so for all $p\in\mathcal{M}$

$$B(\rho,p)\subset B_\Delta(\varepsilon,p)$$

and this in conjunction with (2.12) yields

$$\sup_{p\in\Gamma_2}P_n(t,p,B_\Delta^c(\varepsilon,p))\leqq\sup_{p\in\Gamma_2}P_n(t,p,B^c(\rho,p))\leqq bt$$

for $n \geqq n_\varepsilon$ and all $t>0$; hence

$$\begin{aligned}\sup_{p\in\Gamma_2} \tilde{P}_n(t,p,B_\Delta^c(\varepsilon,p)) &\leqq \sup_{p\in\Gamma_2} P_n(t,p,B_\Delta^c(\varepsilon,p)) + \sup_{p\in\Gamma_2}[1-P_n(t,p,\mathcal{M})]\\ &\leqq (a+b)t, \qquad \text{for } n\geqq n_\varepsilon \text{ and all } t>0.\end{aligned} \tag{2.14}$$

Combining (2.13) and (2.14) we obtain (2.10), and this completes the proof of Theorem 2.2.

Corollary. *Almost all sample-paths of the process* $(X_t^{(0)}, t\geqq 0)$ *are continuous on the interval* $[0,\zeta)$.

Proof. Immediate from the corollary of Proposition 2.4.

Remarks. (I) It should be noted that the real difficulty in the proof of Theorem 2.1 is to establish that the graph of the differential operator $(\mathscr{A}, \tilde{\mathscr{D}})$, $\tilde{\mathscr{D}} = C_{\text{loc}}^{(2,\alpha)} \cap \mathscr{D}_A$, is dense in the graph of $(A, \mathscr{D}_A)$. To see that this result is a consequence of our proof let $f_0 \in \mathscr{D}_A$ and let $\varepsilon>0$. Set

$$g_0 = f_0 - Af_0$$

and choose a $g_\varepsilon \in C_0$ which is Hölder continuous of order α and satisfies

$$\|g_0 - g_\varepsilon\| < \frac{\varepsilon}{3}.$$

Then according to Lemma 2.8

$$f_\varepsilon = R_1\, g_\varepsilon \in \tilde{\mathscr{D}},$$

so

$$\|f_0 - f_\varepsilon\| = \|R_1(g_0 - g_\varepsilon)\| \leqq \|g_0 - g_\varepsilon\| < \frac{\varepsilon}{3},$$

and consequently

$$\begin{aligned}\|(f_0, Af_0) - (f_\varepsilon; \mathscr{A}f_\varepsilon)\| &= \|f_0 - f_\varepsilon\| + \|Af_0 - \mathscr{A}f_\varepsilon\|\\ &= \|f_0 - f_\varepsilon\| + \|(f_0 - g_0) - (f_\varepsilon - g_\varepsilon)\| \leqq 2\,\|f_0 - f_\varepsilon\| + \|g_0 - g_\varepsilon\| < \varepsilon\end{aligned}$$

which proves our statement.

(II) We know already that $C^\infty \cap C_K \subset \mathscr{D}_A$, but by using the fact that the weak and the strong generators for a C_0-semigroup coincide we may even prove

Lemma 2.10. *If* $f\in C^2 \cap C_K$ *then* $f\in\mathscr{D}_A$ *and*

$$Af = \mathscr{A}f.$$

Proof. Recall ([6] p. 20) that if $\{f_n\}_{n=1}^\infty$ is a sequence of functions in C_0 such that

$$f(p) = \lim_{n\to\infty} f_n(p) \text{ exists for every } p\in\mathcal{M}, \quad \text{and} \quad \sup_n \|f_n\| < \infty$$

then $\{f_n\}_{n=1}^\infty$ is said to converge weakly to f. We denote this by

$$f_n \xrightarrow{w} f, \qquad \text{for } n\to\infty.$$

The weak infinitesimal generator $\tilde{A}$ of the semigroup $(T_t, t\geqq 0)$ is then defined by referring to this type of convergence rather than to uniform convergence. However, because $(T_t, t \geqq 0)$ is a C_0-semigroup: $\tilde{A} = A$, and also $\tilde{A}$ is a closed operator; thus we have:

$$\begin{gathered}\text{If } \{f_n\}_{n=1}^{\infty} \subset \mathscr{D}_A \quad \text{and} \quad f_n \xrightarrow{w} f, \quad Af_n \xrightarrow{w} g \in C_0 \\ \text{for } n \to \infty, \text{ then } f \in \mathscr{D}_A \text{ and } Af = g.\end{gathered} \tag{2.15}$$

Now, let $f \in C^2 \cap C_K$ then there exists [29] a sequence $\{f_n\}_{n=1}^{\infty} \subset C^{\infty} \cap C_K$ such that

$$f_n \xrightarrow{w} f \quad \text{and} \quad \mathscr{A} f_n \xrightarrow{w} \mathscr{A} f \quad \text{for } n \to \infty.$$

But for each n

$$f_n \in \mathscr{D}_A \quad \text{and} \quad Af_n = \mathscr{A} f_n,$$

so from (2.15) we may conclude

$$f \in \mathscr{D}_A \quad \text{and} \quad Af = \mathscr{A} f.$$

We remark that we by now have shown that our limit process $(X_t^{(0)}, t \geqq 0)$ is a diffusion process in the sense of: [6] Section 5.18.

(III) Next, let us look a little more closely at the function $P(t, p, \mathscr{M}) =$ "the probability that the process starting from p at time zero is still alive at time t". Set

$$Q(t, p) = P(t, p, \mathscr{M})$$

then

$$Q(t+s, p) = \int_{\mathscr{M}} P(t, q, \mathscr{M})\, P(s, p, dq) \leqq Q(s, p)$$

so for p fixed $Q(t, p)$ is a nonnegative decreasing function with (see the corollary to Lemma 2.9)

$$Q(0, p) = 1 = \operatorname*{s\text{-}lim}_{t \to 0} Q(t, p).$$

Let us first show

$$\frac{\partial Q}{\partial t}(0, p) = -k(p). \tag{2.16}$$

To see this, let $\Gamma \subset \mathscr{M}$ be compact and let $\varepsilon > 0$, define

$$\Gamma_\varepsilon = \{p : d(p, \Gamma) \leqq \varepsilon\}$$

then Γ_ε is a compact neighborhood of Γ. Let $\eta > 0$, then according to the corollary to Proposition 2.4 there is a $\gamma > 0$ such that

$$\sup_{p \in \Gamma,\, 0 < t \leqq \gamma} \frac{1}{t} P\big(t, p, B^c(\varepsilon, p)\big) \leqq \eta$$

[29] This is the approach used in [6] (p. 165). The existence of the sequence $\{f_n\}_{n=1}^{\infty}$ is also shown there in the case $\mathscr{M} = E^N$. To take care of our situation it suffices to cover the support, F, of f with coordinate balls $\{V_i\}_{i=1}^{L}$ and then select functions $\varphi_i \in C^{\infty}$ satisfying $0 \leqq \varphi_i \leqq 1$ with $\operatorname{supp}(\varphi_i)$ being compact and contained in V_i, and such that

$$\sum_{i=1}^{L} \varphi_i(p) = 1 \qquad \text{for } p \in F.$$

We can then find sequences $[\{f_{in}\}_{n=1}^{\infty}]_{i=1}^{L}$ of C^{∞} functions with compact support so

$$f_{in} \xrightarrow{w} \varphi_i \cdot f, \quad \mathscr{A} f_{in} \xrightarrow{w} \mathscr{A}(\varphi_i \cdot f) \quad \text{for } n \to \infty, \ 1 \leqq i \leqq L.$$

Set $f_n = \sum_{i=1}^{L} f_{in}$, then $\{f_n\}_{n=1}^{\infty}$ has the desired properties.

and then also

$$\sup_{p\in\Gamma,\,0<t\leqq\gamma} \frac{1}{t} P(t,p,\Gamma_\varepsilon^c) \leqq \eta .$$

Thus, if f is any bounded measurable function

$$\lim_{t\downarrow 0} \sup_{p\in\Gamma} \left[\frac{1}{t} \int_{\Gamma_\varepsilon^c} f(q)\, P(t,p,dq)\right] = 0 .$$

Now, choose $f\in C^\infty \cap C_K$ with $\chi_{\Gamma_\varepsilon} \leqq f \leqq 1$, then for $p\in\Gamma$

$$\begin{aligned} -k(p) = Af(p) &= \lim_{t\downarrow 0} \frac{1}{t} \left\{ \int_{\mathscr{M}} f(q)\, P(t,p,dq) - 1 \right\} \\ &= \lim_{t\downarrow 0} \frac{1}{t} \left\{ P(t,p,\mathscr{M}) - 1 - \int_{\Gamma_\varepsilon^c} (1 - f(q))\, P(t,p,dq) \right\} \\ &= \lim_{t\downarrow 0} \frac{1}{t} \{ Q(t,p) - Q(0,p) \} = \frac{\partial Q}{\partial t}(0,p), \end{aligned}$$

and this proves (2.16).

Next, according to the corollary to Lemma 2.9 the function

$$\frac{1}{t}\{1 - Q(t,p)\}, \qquad t>0,\ p\in\mathscr{M},$$

is bounded. Hence we have

$$\begin{aligned} \frac{\partial^+ Q}{\partial t}(t,p) &= \lim_{h\downarrow 0} \frac{1}{h} \{ Q(t+h,p) - Q(t,p) \} \\ &= \lim_{h\downarrow 0} \int_{\mathscr{M}} \frac{1}{h} [Q(h,q) - 1]\, P(t,p,dq) \\ &= \int_{\mathscr{M}} \frac{\partial Q}{\partial t}(0,q)\, P(t,p,dq), \end{aligned}$$

so

$$\frac{\partial^+ Q}{\partial t}(t,p) = -\int_{\mathscr{M}} k(q)\, P(t,p,dq). \tag{2.17}$$

Set $\lambda = \sup_{p\in\mathscr{M}} k(p)$, then (2.17) yields

$$\frac{\partial^+ Q}{\partial t}(t,p) \geqq -\lambda \cdot Q(t,p).$$

Assume for a moment that we know $Q(\cdot,p)$ is left continuous, then this result in conjunction with $Q(0,p)=1$ gives

$$Q(t,p) \geqq e^{-\lambda t},$$

and in particular if $k\equiv 0$ we obtain $P(t,p,\mathscr{M})\equiv 1$ so in this case the process $(X_t^{(0)}, t\geqq 0)$ considered as a process on $\mathscr{M}$, is conservative.

Finally, the left continuity of $Q(\cdot, p)$ follows from the inequalities $(h>0)$

$$0 \leqq Q(t-h, p) - Q(t, p) = \int_{\mathcal{M}} [1 - Q(h, q)]\, P(t-h, p, dq)$$

$$= h \cdot \int_{\mathcal{M}} \frac{1}{h} [1 - Q(h, q)]\, P(t-h, p, dq) \leqq h \cdot K,$$

where

$$K = \sup_{\substack{p \in \mathcal{M} \\ 0 < t}} \frac{1}{t} [1 - Q(t, p)] < \infty.$$

We can make a similar construction for each of the processes $(X_t^{(n)}, t \geqq 0)$, that is if we set

$$Q_n(t, p) = P_n(t, p, \mathcal{M}) = e^{-nt} \sum_{j=0}^{\infty} \frac{(nt)^j}{j!} P_n^{(j)}(p, \mathcal{M}),$$

$$k_n(p) = n[1 - \nu_p^{(n)}(\mathcal{M}_p)]$$

then

$$\frac{\partial Q_n}{\partial t}(t, p) = -\int_{\mathcal{M}} k_n(q)\, P_n(t, p, dq), \tag{2.17'}$$

as one easily shows. If now $k(\cdot)$ is contained in C_0 rather than in $\tilde{C}$ we may use (2.17) and (2.17′) to show

$$s\text{-}\lim_{n \to \infty} \overset{n}{T}_t f = T_t f, \qquad \text{for all } f \in \tilde{C}.$$

To do this, it apparently suffices to verify

$$s\text{-}\lim_{n \to \infty} P_n(t, p, \mathcal{M}) = P(t, p, \mathcal{M}) \tag{2.18}$$

because every $f \in \tilde{C}$ is of the form $f_0 + c$ where $f_0 \in C_0$ and c is a constant. But

$$\begin{aligned}
\sup_p |P(t, p, \mathcal{M}) - P_n(t, p, \mathcal{M})| &= \sup_p |Q(t, p) - Q_n(t, p)| \\
&= \sup_p \left| \int_0^t \frac{\partial^+ Q}{\partial t}(s, p)\, ds - \int_0^t \frac{\partial Q_n}{\partial t}(s, p)\, ds \right| \\
&= \sup_p \left| \int_0^t \{\overset{n}{T}_s k_n(p) - \overset{n}{T}_s k(p) + \overset{n}{T}_s k(p) - T_s k(p)\}\, ds \right| \\
&\leqq \int_0^t \|k_n - k\|\, ds + \int_0^t \|\overset{n}{T}_s k - T_s k\|\, ds \\
&\leqq t\Big[\|k_n - k\| + \sup_{0 \leqq s \leqq t} \|\overset{n}{T}_s k - T_s k\|\Big] \xrightarrow[n \to \infty]{} 0
\end{aligned}$$

and so (2.18) follows.

In particular we may now claim, that if the measures $\nu_p^{(n)}$ are all genuine probability measures, that is if

$$\nu_p^{(n)}(\mathcal{M}_p) \equiv 1 \quad \text{or} \quad k_n(p) \equiv 0,$$

then all the processes $\{(X_t^{(n)}, t \geqq 0)\}_{n=0}^{\infty}$ may be taken to be conservative processes on $\mathscr{M}$ rather than on $\mathscr{M}_\Delta$. Also, referring to the corollary of Theorem 2.1, for each $p \in \mathscr{M}$ the sequence $\{P_n^{(\alpha_n)}(p, \cdot)\}_{n=1}^{\infty}$ of probability measures on $\mathscr{B}(\mathscr{M})$ converges *weakly* to the probability measure $P(t, p, \cdot)$ when $\alpha_n \sim nt$.

3. Brownian Motions

3.1. Consider for a moment the situation where our manifold $\mathscr{M}$ is N-dimensional Euclidean space E^N. Ordinary random walks in E^N may then be visualized as being a special kind of geodesic random walks, namely, as being geodesic random walks with identically distributed steps. Also, the corresponding limit processes (obtained as in Section 2) may be thought of as making up a special class of processes canonically associated with the geometric structure of E^N. We shall call these processes Brownian motions in E^N.

On basis of this point of view we now proceed to define the concept of a Brownian motion in an arbitrary Riemannian manifold $\mathscr{M}$. Namely, we define: a Brownian motion in $\mathscr{M}$ is a diffusion process which may be obtained as the limit process [30] for a sequence of geodesic random walks with "identically distributed steps." Here we say that a random walk has identically distributed steps if it is constructed on basis of a family of probability measures $\{\nu_p\}_{p \in \mathscr{M}}$ on the $\mathscr{M}_p$ with the property that for every p, q in $\mathscr{M}$

$$\nu_q = \tau_{pq}\, \nu_p$$

where τ_{pq} denotes parallel translation along any broken C^∞ curve joining p and q.

Consequently, if

$$\mathscr{A} f = \tfrac{1}{2}\langle D^2 f, \sigma^2 \rangle + \langle Df, \mu \rangle \tag{3.1}$$

is the differential generator of a Brownian motion then the tensorfields μ and σ^2 are invariant under parallel translations.

That is, if τ_{pq} is defined as above then

$$\tau_{pq}\, \sigma_p^2 = \sigma_q^2 \quad \text{and} \quad \tau_{pq}\, \mu_p = \mu_q. \tag{3.2}$$

On the other hand, if we are given tensorfields μ and σ^2 satisfying (3.2) and also $\mu_p \in \mathscr{M}_p$, $\sigma_p^2 \in \mathscr{M}_p \otimes \mathscr{M}_p$, σ_q^2 is symmetric and strictly positive definite; then there exists a Brownian motion in $\mathscr{M}$ with differential generator given by (3.1).

To see this, fix $p_0 \in \mathscr{M}$, let $\{e_i\}_{i=1}^N$ be an orthonormal basis for $\mathscr{M}_{p_0}$ such that

$$\sigma_{p_0}^2 = \sum_{i=1}^{N} \lambda_i^2\, e_i \otimes e_i, \qquad 0 < \lambda_1 \leqq \lambda_2 \leqq \cdots \leqq \lambda_N,$$

and let ν_{p_0} be the measure on $\mathscr{M}_{p_0}$ corresponding to a uniform mass distribution of total mass 1 on the ellipsoid $\mathscr{E}_{p_0}$ with center μ_{p_0} and principal axes along the e_i with lengths $l_i = \alpha \lambda_i$, where the value of the constant α will be determined later. First, we show that if γ is a broken C^∞ curve starting and ending at p_0 and τ denotes parallel translation along γ then

$$\tau\, \mathscr{E}_{p_0} = \mathscr{E}_{p_0}, \qquad \text{and consequently } \tau\, \nu_{p_0} = \nu_{p_0}.$$

[30] In the sense of Section 2.

So let $X \in \mathscr{E}_{p_0}$. This means

$$X = \mu_{p_0} + \sum_{i=1}^{N} x^i e_i, \quad \text{where} \quad \sum_{i=1}^{N} \left(\frac{x^i}{\lambda_i}\right)^2 \leqq \alpha^2,$$

and we must show $\tau X \in \mathscr{E}_{p_0}$, that is

$$\tau X = \mu_{p_0} + \sum_{i=1}^{N} \xi^i e_i, \quad \text{with} \quad \sum_{i=1}^{N} \left(\frac{\xi^i}{\lambda_i}\right)^2 \leqq \alpha^2.$$

But $\tau \mu_{p_0} = \mu_{p_0}$ per assumption so

$$\tau \left(\mu_{p_0} + \sum_{i=1}^{N} x^i e_i\right) = \mu_{p_0} + \sum_{i=1}^{N} x^i \tau e_i = \mu_{p_0} + \sum_{i=1}^{N} \xi^i e_i$$

where

$$\tau e_i = \sum_{j=1}^{N} \beta_i^j e_j, \qquad \xi^i = \sum_{j=1}^{N} x^j \beta_j^i,$$

thus it remains to prove

$$\sum_{i=1}^{N} \left[\frac{1}{\lambda_i} \sum_{j=1}^{N} x^j \beta_j^i\right]^2 \leqq \alpha^2. \tag{3.3}$$

To do this, let $\bar{x}$ be the column vector (x^i), and introduce the matrices

$$\Lambda = \{\lambda_{ij}\}, \qquad B = \{b_{ij}\}$$

where

$$\lambda_{ij} = \lambda_i \cdot \delta_{ij} \quad \text{and} \quad b_{ij} = \beta_j^i,$$

then the sum in (3.3) equals:

$$\|\Lambda^{-1} B \bar{x}\|^2 = \bar{x}^* B^* \Lambda^{-2} B \bar{x}.$$

The fact that τ is an isometry implies that the matrix B is orthogonal. Furthermore, according to (3.2)

$$\sum_{i=1}^{N} \lambda_i^2 e_i \otimes e_i = \sigma_{p_0}^2 = \tau \sigma_{p_0}^2 = \sum_{i,j,k=1}^{N} \lambda_i^2 \beta_i^j \beta_i^k e_j \otimes e_k$$

so

$$\sum_{i=1}^{N} \lambda_i^2 \beta_i^j \beta_i^k = \begin{cases} \lambda_j^2 & \text{for } j = k \\ 0 & \text{for } j \neq k \end{cases}$$

which means

$$\Lambda^2 = B \Lambda^2 B^*,$$

and consequently

$$\Lambda^{-2} = B^* \Lambda^{-2} B.$$

Altogether

$$\sum_{i=1}^{N} \left[\frac{1}{\lambda_i} \sum_{j=1}^{N} x^j \beta_j^i\right]^2 = \|\Lambda^{-1} B \bar{x}\|^2 = \|\Lambda^{-1} \bar{x}\|^2 = \sum_{i=1}^{N} \frac{1}{\lambda_i^2} (x^i)^2 \leqq \alpha^2$$

and this proves (3.3).

We have thus shown that for every parallel translation, τ, around a closed circuit starting and ending at p_0 it holds

$$\tau \mathscr{E}_{p_0} \subset \mathscr{E}_{p_0}$$

and this implies

$$\tau \mathscr{E}_{p_0} = \mathscr{E}_{p_0}$$

as claimed.

Next, we may define v_p for arbitrary $p \in \mathscr{M}$ by setting

$$v_p = \tau_{p_0 p} \, v_{p_0}$$

where $\tau_{p_0 p}$ denotes parallel translation along any broken C^∞ curve γ joining p_0 and p.

What we have just shown implies that the definition is independent of the choice of γ. Moreover, it is clear that

$$\int_{\mathscr{M}_p} X \, v_p(dX) = \mu_p, \qquad \int_{\mathscr{M}_p} X \otimes X \, v_p(dX) - \mu_p \otimes \mu_p = \sigma_p^2 \tag{3.4}$$

if this just holds at $p = p_0$. So let us prove that by computing the integrals at $p = p_0$.

Let E be the ellipsoid in E^N centered at the origin and with axes in the directions of the coordinate axes and with lengths l_i, so

$$E = \left\{ \bar{x} : \bar{x} = (x^i) \in E^N, \ \sum_{i=1}^N \left(\frac{x^i}{\lambda_i} \right)^2 \leqq \alpha^2 \right\}.$$

We can then identify E and $\mathscr{E}_{p_0}$ via the map

$$\bar{x} \to \mu_{p_0} + \sum_{i=1}^N x^i \, e_i.$$

If V denotes the volume of E and $d\bar{x}$ Lebesgue measure in E^N we may then write

$$\int_{\mathscr{M}_{p_0}} X \, v_{p_0}(dX) = \mu_{p_0} + \sum_{i=1}^N \left[\frac{1}{V} \int_E x^i \, d\bar{x} \right] e_i = \mu_{p_0},$$

$$\begin{aligned}
\int_{\mathscr{M}_{p_0}} X \otimes X \, v_{p_0}(dX) - \mu_{p_0} \otimes \mu_{p_0} &= \int_{\mathscr{M}_{p_0}} (X - \mu_{p_0}) \otimes (X - \mu_{p_0}) \, v_{p_0}(dX) \\
&= \sum_{i,j=1}^N \left[\frac{1}{V} \int_E x^i x^j \, d\bar{x} \right] e_i \otimes e_j = \frac{1}{V} \sum_{i=1}^N \left(\int_E (x^i)^2 \, d\bar{x} \right) e_i \otimes e_i \\
&= \frac{1}{V} \sum_{i=1}^N \left(\int_E \left(\frac{x^i}{\lambda_i} \right)^2 d\bar{x} \right) \lambda_i^2 \, e_i \otimes e_i.
\end{aligned}$$

Set [31] $y^i \lambda_i \alpha = x^i$ then in these coordinates

$$E \sim \{ \bar{y} : \bar{y} = (y^i) \in E^N, \sum y_i^2 \leqq 1 \} = S$$

so

$$\int_E \left(\frac{x^i}{\lambda_i} \right)^2 d\bar{x} = \int_S (y^i)^2 \, d\bar{y} \cdot D$$

[31] No summation!

with $D=(\lambda_1 \cdot \cdots \cdot \lambda_N) \cdot \alpha^{N+2}$, and

$$\int_S (y^i)^2 \, d\bar{y} = \frac{1}{N} \sum_{i=1}^{N} \int_S (y^i)^2 \, d\bar{y} = \frac{1}{N} \int_S |\bar{y}|^2 \, d\bar{y} = K$$

where K only depends upon N. Altogether

$$\frac{1}{V} \sum_{i=1}^{N} \left(\int_E \left(\frac{x^i}{\lambda_i} \right)^2 d\bar{x} \right) \lambda_i^2 \, e_i \otimes e_i = \frac{KD}{V} \sigma_{p_0}^2,$$

and as V is proportional to α^N while $D \sim \alpha^{N+2}$ we may choose α such that $KD=V$. This proves (3.4). Now set $v_p^{(n)} = v_p$ for all $n \geqq 1$, $p \in \mathcal{M}$, then it is easily verified that the sequence $\{(v_p^{(n)})_{p \in \mathcal{M}}\}_{n=1}^{\infty}$ satisfies conditions $\mathscr{C}0$–$\mathscr{C}6$, and consequently the sequence of random walks constructed from these measures converges weakly to a diffusion process in $\mathcal{M}$ with differential generator $\mathscr{A}$ given by (3.1).

Moreover, this diffusion process is a Brownian motion because the corresponding random walks have identically distributed steps. We have thus shown

Proposition 3.1. *The operator*

$$\mathscr{A} f = \tfrac{1}{2} \langle D^2 f, \sigma^2 \rangle + \langle Df, \mu \rangle$$

is the differential generator of a Brownian motion if and only if μ and σ^2 are constant.

"constant" here means "invariant under parallel translations", and it is of course still assumed that σ^2 is strictly positive definite.

Example. For each $p \in \mathcal{M}$ choose an orthonormal basis $\{e_i(p)\}_{i=1}^N$, and then set (sum over repeated indices)

$$\mu_p = 0, \qquad \sigma_p^2 = e_i(p) \otimes e_i(p), \tag{3.5}$$

then these fields satisfy (3.2). What μ is concerned this is obvious, and it is also clear from the above computations that σ_p^2 is the covariance[32] of a uniform distribution on a sphere in $\mathcal{M}_p$ centered at 0, and thus is invariant under rotations of $\mathcal{M}_p$ about 0. But parallel translation around a closed loop starting and ending at p is such a rotation, and consequently σ_p^2 satisfies (3.2).

It thus follows, that there always exists at least one Brownian motion in $\mathcal{M}$, namely, the one[33] whose differential generator is

$$\mathscr{A} f = \tfrac{1}{2} \langle D^2 f, \sigma^2 \rangle$$

with σ^2 determined by (3.5). We shall call this process the Wiener process in $\mathcal{M}$.

The differential generator $\mathscr{A}$ of the Wiener process is $\frac{1}{2} \nabla^2$, which may be seen as follows. For $p \in \mathcal{M}$ let as before $\{e_i(p)\}_{i=1}^N$ be an orthonormal basis for $\mathcal{M}_p$ and let $\{\varepsilon^j(p)\}_{j=1}^N$ be the corresponding dual basis for $\mathcal{M}_p^*$. Then

$$\sigma_p^2 = e_i(p) \otimes e_i(p), \qquad g_p = \varepsilon^j(p) \otimes \varepsilon^j(p)$$

[32] Which in the present setting is contravariant.

[33] Processes whose generators are proportional to each other will not be considered as being essentially different. It is of course assumed $\mathcal{M}$ is complete and of class $\mathscr{A}\mathscr{B}$.

where g_p is the metric tensor "evaluated" at p. Using the definition of ∇^2 (see 4.1) we now have

$$\langle D^2 f, \sigma_p^2 \rangle = \langle g_p, G^2 f \rangle = (\nabla^2 f)(p)$$

for $f \in C^2$, and this proves our claim.

The transition function $W(t, p, \Gamma)$ of the Wiener process is the natural counterpart to the normal distribution $N(t, \bar{x}, \bar{y})\, d\bar{y}$ in E^N, where

$$N(t, \bar{x}, \bar{y}) = (2\pi t)^{-\frac{N}{2}} \cdot e^{-\frac{\|\bar{x}-\bar{y}\|^2}{2t}},$$

and just as $N(t, \bar{x}, \bar{y})$ is invariant under rotations and translations of E^N so $W(t, p, \Gamma)$ is invariant under isometries of $\mathcal{M}$, i.e., if $\varphi: \mathcal{M} \to \mathcal{M}$ is an isometry then

$$W(t, p, \Gamma) = W\big(t, \varphi^{-1}(p), \varphi^{-1}(\Gamma)\big). \tag{3.6}$$

To see this, set for $t \geqq 0$, $p \in \mathcal{M}$, $\Gamma \in \mathcal{B}(\mathcal{M})$:

$$Q(t, p, \Gamma) = W\big(t, \varphi^{-1}(p), \varphi^{-1}(\Gamma)\big),$$

then Q is a transition function. The corresponding semigroup $(S_t, t \geqq 0)$ is easily seen to be a C_0-semigroup, and if $(T_t, t \geqq 0)$ is the semigroup associated with W then we have, for $f \in C_0$ and $t \geqq 0$

$$\varphi^*(S_t f) = T_t(\varphi^* f) \tag{3.7}$$

where

$$(\varphi^* f)(p) = f\big(\varphi(p)\big).$$

Let the operator A with domain $\mathcal{D} \subset C_0$ be the infinitesimal generator of $(T_t, t \geqq 0)$, and correspondingly let $(A_1, \mathcal{D}_1)$ be the infinitesimal generator of $(S_t, t \geqq 0)$, then (3.7) yields

$$\varphi^* f \in \mathcal{D} \Rightarrow f \in \mathcal{D}_1 \quad \text{and} \quad \varphi^*(A_1 f) = A(\varphi^* f).$$

Now, if $f \in C^2 \cap C_K$ then also $\varphi^* f \in C^2 \cap C_K \subset \mathcal{D}$, so

$$A_1 f(p) = (\varphi^{-1})^* \circ A(\varphi^* f)(p) = \tfrac{1}{2}(\varphi^{-1})^* \circ \nabla^2(\varphi^* f)(p) = \tfrac{1}{2}\nabla^2 f(p) = Af(p),$$

from which we may conclude (see Remark I at the end of Section 2) that

$$\mathcal{D} = \mathcal{D}_1 \quad \text{and} \quad Af = A_1 f \quad \text{for } f \in \mathcal{D},$$

and then also

$$T_t f = S_t f \quad \text{for all } f \in C_0 \text{ and } t \geqq 0.$$

This in conjunction with

$$Q(t, p, \mathcal{M}) = W\big(t, \varphi^{-1}(p), \mathcal{M}\big) = 1$$

implies

$$Q(t, p, \Gamma) = W(t, p, \Gamma)$$

which is the statement (3.6).

In case the class of isometries of $\mathcal{M}$ is "large", (3.6) apparently imposes strong restrictions upon the possible behavior of $W(t, p, \Gamma)$, and might even in some cases be expected to characterize $W(t, p, \Gamma)$ uniquely. This point of view has been taken by Yosida in several of his papers ([24, 25] and also [26] p. 398 ff.). Specifically, he has considered the situation where $\mathscr{M}$ is a homogeneous space and has defined a Brownian motion on such a space to be a temporally homogeneous Markov process whose transition function $P(t, p, \Gamma)$ satisfies (3.6) and the condition [34]

$$\lim_{t \downarrow 0} \frac{1}{t} P(t, p, V_p^c) = 0$$

for all $p \in \mathscr{M}$ and all neighborhoods V_p of p. On basis of this, he proves that the corresponding semigroup is a $\tilde{C}$-semigroup [35] whose infinitesimal generator, A, in local coordinates is given by the expression

$$Af(\bar{x}) = \frac{1}{2} b^{ij}(\bar{x}) \frac{\partial^2 f}{\partial x^i \partial x^j} + a^i(\bar{x}) \frac{\partial f}{\partial x^i}, \qquad \text{for } f \in C^2 \cap \mathscr{D}_A,$$

where the "infinitesimal drift" and "variance" are determined by

$$\begin{aligned} a^i(\bar{x}_0) &= \lim_{t \downarrow 0} \frac{1}{t} \int_{d(\bar{x}, \bar{x}_0) < \varepsilon} (x^i - x_0^i)\, P(t, \bar{x}_0, d\bar{x}) \\ b^{ij}(\bar{x}_0) &= \lim_{t \downarrow 0} \frac{1}{t} \int_{d(\bar{x}, \bar{x}_0) < \varepsilon} (x^i - x_0^i)(x^j - x_0^j)\, P(t, \bar{x}_0, d\bar{x}) \end{aligned} \tag{3.8}$$

the integrals being independent of the choice of $\varepsilon > 0$. Now, (3.6) implies

$$A \circ \varphi^* = \varphi^* \circ A, \qquad \text{on } C^2 \cap \mathscr{D}_A,$$

whenever φ is an isometry of $\mathscr{M}$, and this relation may be used to set up differential equations which must be satisfied by the a^i and b^{ij}. From this Yosida proved ([24]) that there is essentially only one Brownian motion on S^2, namely, the Wiener process with generator $\frac{1}{2}\nabla^2$.

It should be pointed out, that in general "invariance under isometries" is not the same as "invariance under parallel translations." Thus, our definition of a Brownian motion is not the same as the usual one, but nevertheless the Wiener process is a Brownian motion according to both definitions. In terms of the infinitesimal characteristics μ and σ^2, where as usual

$$\mathscr{A} f = \tfrac{1}{2}\langle D^2 f, \sigma^2 \rangle + \langle Df, \mu \rangle,$$

the difference between the two definitions shows up as follows. It follows from Proposition 3.1, that the process with generator $\mathscr{A}$ is a Brownian motion according to our definition if

$$\tau_{pq}\, \sigma_p^2 = \sigma_q^2, \qquad \tau_{pq}\, \mu_p = \mu_q, \qquad \text{for all } p, q;$$

while it is a Brownian motion according to Yosida's definition if

$$\varphi_*(\mu_p) = \mu_q, \qquad \varphi_*(\sigma_p^2) = \sigma_q^2, \qquad \text{for all } p, q,$$

whenever φ is an isometry mapping p into q, and φ_* denotes the differential of φ.

---

[34] Lindeberg's condition. Compare with the corollary of our Proposition 2.4.

[35] It is assumed that $P(t, p, \mathscr{M}) = 1$.

In order to prove this statement we first give an invariant interpretation of (3.8). Let $p \in \mathscr{M}$ be given, then there exists a ball B_p in $\mathscr{M}_p$ centered at 0, and a ball B'_p in $\mathscr{M}$ centered at p such that the map

$$\exp_p\colon\ B_p \to B'_p$$

is a diffeomorphism. Define for Borel sets $\Gamma \subset B_p$

$$m_p(t, \Gamma) = P(t, p, \exp_p \Gamma),$$

then for each $t \geqq 0$ $m_p(t, \cdot)$ is a subprobability measure on B_p, and for f bounded and measurable

$$\int_{B'_p} f(q)\, P(t, p, dq) = \int_{B_p} f(\exp_p X)\, m_p(t, dX).$$

Thus, if f agrees on B'_p with a function in $\mathscr{D}_A$ the Lindeberg condition implies

$$Af(p) = \lim_{t \downarrow 0} \frac{1}{t} \left[\int_{B_p} f(\exp_p X)\, m_p(t, dX) - f(p) \right].$$

Now, let $\omega \in \mathscr{M}_p^*$ and define the function f on B'_p by

$$f(\exp_p X) = \omega(X).$$

It is no restriction to assume (see [26] the proof given on p. 400–403) that f agrees on a neighborhood, which we may assume is B'_p, of p with a function in $C^2 \cap \mathscr{D}_A$. Consequently

$$Af(p) = \lim_{t \downarrow 0} \frac{1}{t} \left[\int_{B_p} \omega(X)\, m_p(t, dX) - 0 \right] = \omega \left(\lim_{t \downarrow 0} \frac{1}{t} \int_{B_p} X\, m_p(t, dX) \right).$$

To find $Af(p) = \mathscr{A} f(p)$ introduce normal coordinates on B'_p. That is, the point $\exp_p(\alpha^i e_i)$ where $\{e_i\}_{i=1}^N$ is some fixed orthonormal basis for $\mathscr{M}_p$, gets coordinates $(\alpha^1, \ldots, \alpha^N)$. Then we may write

$$\mathscr{A} f(p) = \frac{1}{2} \sigma^{ij} \left(\frac{\partial^2 \bar{f}}{\partial \alpha^i\, \partial \alpha^j} - \Gamma_{ij}^k \frac{\partial \bar{f}}{\partial \alpha^k} \right) + \mu^j \frac{\partial \bar{f}}{\partial \alpha^j}$$

where

$$\bar{f}(\alpha^1, \ldots, \alpha^N) = f(\exp_p(\alpha^i e_i)) = \omega(\alpha^i e_i) = \alpha^i \omega(e_i).$$

Because of the choice of coordinate system, the Christoffel symbols vanish at p [36], hence

$$\mathscr{A} f(p) = \mu^j \omega(e_j) = \omega(\mu_p)$$

[36] For each set $(\beta^1, \ldots, \beta^N) \in E^N$ the curve

$$\bar{x}(t) = (t\beta^1, \ldots, t\beta^N) \sim \exp_p(t\beta^i e_i)$$

is a geodesic. Consequently for $1 \leqq k \leqq N$

$$0 = \frac{d^2 x^k}{dt^2} + \Gamma_{ij}^k(\bar{x}(t)) \frac{dx^i}{dt} \frac{dx^j}{dt} = \Gamma_{ij}^k(\bar{x}(t))\, \beta^i \beta^j,$$

so for $t = 0$ and all i, j, k

$$\Gamma_{ij}^k(p)\, \beta^i \beta^j = 0 \qquad \text{so } \Gamma_{ij}^k(p) = 0.$$

which taken together with the result above yields

$$\omega(\mu_p)=\omega\left(\lim_{t\downarrow 0}\frac{1}{t}\int_{B_p} X\, m_p(t, dX)\right),$$

and thus because $\omega\in\mathscr{M}_p^*$ was arbitrary

$$\mu_p=\lim_{t\downarrow 0}\frac{1}{t}\int_{B_p} X\, m_p(t, dX). \tag{3.8 i}$$

A similar argument, using the function

$$f(\exp_p X)=\langle\omega\otimes\omega, X\otimes X\rangle$$

gives

$$\sigma_p^2=\lim_{t\downarrow 0}\frac{1}{t}\int_{B_p} X\otimes X\, m_p(t, dX). \tag{3.8 ii}$$

Now, let $\varphi\colon\mathscr{M}\to\mathscr{M}$ be an isometry with $q=\varphi(p)$ and put

$$\varphi_*(B_p)=B_q,\qquad \varphi(B'_p)=B'_q,$$

then the map

$$\varphi\circ\exp_p=\exp_q\circ\varphi_*$$

is a diffeomorphism sending B_p onto B'_q. So, constructing the measure $m_q(t,\cdot)$ on B_q the same way as we constructed $m_p(t,\cdot)$ on B_p, and using that $P(t,p,\Gamma)$ per assumption is invariant under φ we obtain

$$m_p(t,\Gamma)=m_q(t,\varphi_*(\Gamma)),\qquad \text{for } \Gamma\in\mathscr{B}(B_p).$$

Finally, this result combined with (3.8 i + ii) is easily seen to imply

$$\varphi_*(\mu_p)=\mu_q,\qquad \varphi_*(\sigma_p^2)=\sigma_q^2$$

which is what we wanted to prove.

3.2. After this detour, we return to our original definition of a Brownian motion as being a diffusion process in $\mathscr{M}$ that may be obtained as the limit of a sequence of geodesic random walks with identically distributed steps. As shown in Proposition 3.1, such a process is essentially characterized by the requirement that it must satisfy (3.2), and consequently we may classify Brownian motions by classifying the corresponding tensorfields μ and σ^2. We shall in the following use this approach to obtain a description of which Brownian motions there may exist in a given manifold $\mathscr{M}$. Our motivation for considering this type of problem has been the result by Yosida mentioned above, that the Wiener process is the only Brownian motion in S^2. We shall prove a corresponding result for general manifolds.

Because of the way things have been set up, our results will follow rather easily from certain standard theorems from differential geometry[37]. Consequently, we have omitted the proofs and tried to keep the exposition as brief as possible.

[37] Our basic reference here is [14].

If p is an arbitrary point of $\mathscr{M}$ and γ is a closed curve starting and ending at p, then parallel translation along γ induces an automorphism of $\mathscr{M}_p$. One defines $\Psi(p)$, the homogeneous holonomy group at p, to be the group of automorphisms of $\mathscr{M}_p$ that may be obtained by such parallel translations. The connectedness of $\mathscr{M}$ implies that for any pair of points, (p, q), $\Psi(p)$ and $\Psi(q)$ are isomorphic, and consequently we may speak of the holonomy group $\Psi(=\Psi(p)=\Psi(q))$ of $\mathscr{M}$ without referring to any particular point p.

Now, if $\mathscr{N}$ is a subspace of $\mathscr{M}_p$ which is invariant under the action of Ψ and which contains no proper Ψ-invariant subspace then $\mathscr{N}$ is said to be irreducible, and, in particular, if $\mathscr{M}_p$ is irreducible we say that the manifold itself is irreducible. Generally, $\mathscr{M}_p$ can be written as a direct sum

$$\mathscr{M}_p=\mathscr{M}_p^0\oplus\mathscr{M}_p^1\oplus\cdots\oplus\mathscr{M}_p^L \tag{3.9}$$

of mutually orthogonal Ψ-invariant subspaces where $\mathscr{M}_p^1, \ldots, \mathscr{M}_p^L$ are irreducible and

$$\mathscr{M}_p^0=\{X: X\in\mathscr{M}_p \text{ and } \tau X=X \text{ for all } \tau\in\Psi(p)\}.$$

It may of course happen that $\mathscr{M}_p^0=\{0\}$.

(3.9) is a so-called canonical decomposition of $\mathscr{M}_p$, and it is unique up to a reordering of the $\mathscr{M}_p^j$ if $\mathscr{M}$ is simply connected. A basis $\{e_i\}_{i=1}^N$ for $\mathscr{M}_p$ with the property that

$$e_1, \ldots, e_{k_1}\in\mathscr{M}_p^0;\ e_{k_1+1}, \ldots, e_{k_1+k_2}\in\mathscr{M}_p^1;\ \ldots;\ e_{k_1+\cdots+k_L}, \ldots, e_N\in\mathscr{M}_p^L$$

will be said to be adapted to the given decomposition.

Recall that the concept of parallelism is defined in terms of the Riemannian connexion ∇ rather than in terms of the metric tensor g. Consequently, if g and $\tilde{g}$ are metric tensors generating the same affine connexion then one would expect that canonical decompositions relative to these metrics could be chosen so as to be closely related. And indeed, using elementary linear algebra one may prove.

Lemma 3.1. *Let g and $\tilde{g}$ be two metric tensors on $\mathscr{M}$ which generate the same affine connexion ∇, then for each $p\in\mathscr{M}$ there exists a decomposition*

$$\mathscr{M}_p=\mathscr{M}_p^0\oplus\cdots\oplus\mathscr{M}_p^L$$

which is canonical relative to both metrics, invariant under parallel translations, and satisfies

(I) *For each i, $1\leqq i\leqq L$, there is a constant $\lambda_i>0$ such that*

(a) $\tilde{g}_p(X, Y)=\lambda_i\, g_p(X, Y)$, *for $X, Y\in\mathscr{M}_p^i$.*

(II) *$\mathscr{M}_p^0$ may be written as a direct sum*

$$\mathscr{M}_p^0=\mathscr{H}_p^1\oplus\cdots\oplus\mathscr{H}_p^K$$

of mutually orthogonal subspaces such that

(b) $\tilde{g}_p(X, Y)=l_j\, g_p(X, Y)$, *for $X, Y\in\mathscr{H}_p^j$ for some constants $l_j>0$, $1\leqq j\leqq K$.*

On the other hand, let (3.9) *be a given g-canonical*[38] *decomposition of* $\mathcal{M}_{p_0}$, $p_0 \in \mathcal{M}$, *and let*

$$\mathcal{M}_{p_0}^0 = \mathcal{H}_{p_0}^1 \oplus \cdots \oplus \mathcal{H}_{p_0}^K$$

where the $\mathcal{H}_{p_0}^i$ *are g-orthogonal. Define for each* $p \in \mathcal{M}$ *a corresponding decomposition of* $\mathcal{M}_p$ *by parallel transport of the given decomposition from* p_0 *to p. Choose positive numbers* $\lambda_1, \ldots, \lambda_L, l_1, \ldots, l_K$ *and define* $\tilde{g}$ *by conditions* (a) *and* (b) *together with the conditions that* $\tilde{g}(X, Y)$ *must be linear in X and Y and that the spaces* $\mathcal{H}_p^1, \ldots, \mathcal{H}_p^K$, $\mathcal{M}_p^1, \ldots, \mathcal{M}_p^L$ *must be* $\tilde{g}$*-orthogonal. Then* $\tilde{g}$ *is a metric tensor which is invariant under parallel translations, and thus the connexion generated by* $\tilde{g}$ *is* ∇.

We are now ready to describe the different Brownian motions in $\mathcal{M}$ on basis of the requirement that σ^2 and μ must be "constant". The two terms are independent of each other so we may treat them separately. First, the remarks above imply immediately

Proposition 3.2. *There exists Brownian motions with nonzero drift,* μ, *on* $\mathcal{M}$ *if and only if the holonomy group* $\Psi(p)$ *at any point p leaves fixed some nonzero element of* $\mathcal{M}_p$. *Conversely, every such element* X_p *can be used to define a drift term* μ *by setting* $\mu_q = \tau_{pq} X_p$. τ_{pq} *denoting parallel translation from p to q.*

Specifically, the proposition says that if we consider a canonical decomposition (3.9) of $\mathcal{M}_p$ then the condition for the existence of a nonzero drift term μ is that

$$\mathcal{M}_p^0 \neq \{0\}.$$

Next, consider the tensorfield σ^2. It is a symmetric and strictly positive definite tensorfield of type (2, 0) satisfying condition (3.2), and we shall use this to define a new metric tensor $\tilde{g}$ on $\mathcal{M}$. First, for $p \in \mathcal{M}$ define the map

$$F\colon \mathcal{M}_p \to \mathcal{M}_p^*$$

by

$$\langle F(X), Y \rangle = g_p(X, Y)$$

for all X, Y in $\mathcal{M}_p$. Then set

$$\tilde{g}_p(X, Y) = \langle \sigma_p^2, F(X) \otimes F(Y) \rangle$$

for X, Y in $\mathcal{M}_p$ (we are "lowering indices"). It is straightforward to check that this defines $\tilde{g}$ as a metric tensor on $\mathcal{M}$, which is invariant under parallel translations, and consequently Lemma 3.1 applies. Using the notation introduced in that lemma we thus have, for $p \in \mathcal{M}$,

$$\mathcal{M}_p = \mathcal{M}_p^0 \oplus \cdots \oplus \mathcal{M}_p^L,$$
$$\mathcal{M}_p^0 = \mathcal{H}_p^1 \oplus \cdots \oplus \mathcal{H}_p^K,$$

these spaces being mutually orthogonal relative to both metrics, and

$$\begin{aligned} \tilde{g}_p(X, Y) &= \lambda_i\, g_p(X, Y), \quad \text{for } X, Y \text{ in } \mathcal{M}_p^i, \quad 1 \leq i \leq L; \\ \tilde{g}_p(X, Y) &= l_j\, g_p(X, Y), \quad \text{for } X, Y \text{ in } \mathcal{H}_p^j. \end{aligned} \tag{3.10}$$

[38] That is, canonical relative to the metric g.

It is clearly no restriction to assume that the $\mathscr{H}_p^j$ are one-dimensional, so this we shall do also. Now, choose for each $p\in\mathscr{M}$ a basis $\{e_i(p)\}_{i=1}^N$ for $\mathscr{M}_p$ which is adapted to the given decomposition, and such that, for $1\leq i\leq K$, $e_i(p)$ spans $\mathscr{H}_p^i$. We may assume that this basis is orthonormal relative to the metric g and orthogonal relative to the metric $\tilde{g}$. Set for $1\leq i\leq K$

$$(s_p^2)_i = e_i(p)\otimes e_i(p),$$

and if $e_{k_i+1}(p),\ldots,e_{k_{i+1}}(p)$ are those vectors among the $e_i(p)$ which span $\mathscr{M}_p^i$, $1\leq i\leq L$, set

$$(s_p^2)^i = \sum_{j=k_i+1}^{k_{i+1}} e_j(p)\otimes e_j(p).$$

The $(s^2)^i$ and $(s^2)_j$ are symmetric contravariant tensorfields that are invariant under parallel translations, and for each $p\in\mathscr{M}$ we have the decomposition

$$\sigma_p^2 = \sum_{i=1}^K l_i\cdot(s_p^2)_i + \sum_{i=1}^L \lambda_i\cdot(s_p^2)^i. \tag{3.11}$$

If, in particular, $\mathscr{M}$ is irreducible there is only one term, i.e.,

$$\sigma_p^2 = \lambda\sum_{i=1}^N e_i(p)\otimes e_i(p),$$

and this means (compare with (3.5)) that the corresponding Brownian motion has differential generator

$$\mathscr{A}f = \tfrac{1}{2}\langle D^2 f, \sigma^2\rangle = \frac{\lambda}{2}\nabla^2 f,$$

so it is (essentially) the Wiener process.

In the general case we derive from (3.11):

Proposition 3.3. *Let $\mathscr{M}$ be a complete Riemannian manifold of class $\mathscr{A}\mathscr{B}$.*

(I) *If $\mathscr{M}$ is irreducible there exists only one Brownian motion on $\mathscr{M}$, namely the Wiener process.*

(II) *In case $\mathscr{M}$ is reducible, the Brownian motions on $\mathscr{M}$ may be classified as follows. Let $p_0\in\mathscr{M}$ and let*

$$\mathscr{M}_{p_0} = \mathscr{M}_{p_0}^0\oplus\cdots\oplus\mathscr{M}_{p_0}^L$$

be a canonical decomposition of $\mathscr{M}_{p_0}$. Also, let $\{e_i\}_{i=1}^N$ be an orthonormal basis for $\mathscr{M}_{p_0}$ adapted to this decomposition with

$$e_1,\ldots,e_{k_1}\in\mathscr{M}_{p_0}^0;\qquad e_{k_1+1},\ldots,e_{k_2}\in\mathscr{M}_{p_0}^1\quad etc.,$$

let $l_1,\ldots,l_{k_1}$, $\lambda_1,\ldots,\lambda_L$ be positive numbers and let $\alpha_1,\ldots,\alpha_{k_1}$ be arbitrary real numbers. Set

$$\mu_{p_0} = \sum_{j=1}^{k_1}\alpha_j e_j;\qquad \sigma_{p_0}^2 = \sum_{j=1}^{k_1} l_j e_j\otimes e_j + \sum_{i=1}^L\lambda_i\left[\sum_{j=k_i+1}^{k_{i+1}} e_j\otimes e_j\right],$$

$$\mu_q = \tau_{p_0q}\,\mu_{p_0},\qquad \sigma_q^2 = \tau_{p_0q}\,\sigma_{p_0}^2,\quad q\in\mathscr{M},$$

where $\tau_{p_0 q}$ denotes parallel translation along any C^∞ curve joining p_0 and q. Then there exists a Brownian motion on $\mathcal{M}$ with differential generator

$$\mathscr{A}f=\tfrac{1}{2}\langle D^2 f, \sigma^2\rangle+\langle Df, \mu\rangle. \tag{3.12}$$

Moreover, every Brownian motion on $\mathcal{M}$ may be obtained in this fashion.

The geometric content of this result is perhaps most easily comprehended if one compares with what happens if $\mathcal{M}=E^N$. A Brownian motion on E^N has generator

$$\mathscr{A}f=a^{ij}\frac{\partial^2 f}{\partial x^i\,\partial x^j}+b^k\frac{\partial f}{\partial x^k}$$

with constant coefficients. Here it may be assumed that the matrix $\{a^{ij}\}$ has zero entries outside the diagonal, and thus the process may be viewed as being a "sum" of one-dimensional motions along the coordinate axes. Now, Proposition 3.3 expresses the corresponding result for a general manifold $\mathcal{M}$, namely, that locally a Brownian motion on $\mathcal{M}$ may be thought of as being the "sum" of certain basic Brownian motions which cannot themselves be further reduced. If $\mathcal{M}$ is simply connected we can obtain a complete analogy with the case $\mathcal{M}=E^N$ by using the de Rham decomposition theorem [39] according to which $\mathcal{M}$ is isometric to a direct product

$$\mathcal{M}^0\times\mathcal{M}^1\times\cdots\times\mathcal{M}^L$$

where $\mathcal{M}^0=E^K$, $K=\dim\mathcal{M}_p^0\geqq 0$, and $\mathcal{M}^1,\ldots,\mathcal{M}^L$ are simply connected, complete, irreducible Riemannian manifolds, each one being of class $\mathscr{A}\mathscr{B}$ (because $\mathcal{M}$ itself is assumed to be of class $\mathscr{A}\mathscr{B}$). Identifying $\mathcal{M}$ and $\mathcal{M}^0\times\cdots\times\mathcal{M}^L$ we get

$$\mathcal{M}_p=(\mathcal{M}^0)_p\oplus(\mathcal{M}^1)_p\oplus\cdots\oplus(\mathcal{M}^L)_p$$

the point being that $\mathcal{M}_p^j=(\mathcal{M}^j)_p$ for each j, where as usual the $\mathcal{M}_p^j$ are given by (3.9). Let $X_0(t),\ldots,X_L(t)$ be independent Brownian motions in $\mathcal{M}^0,\ldots,\mathcal{M}^L$ with generators

$$\mathscr{A}_0 f_0=\sum_{j=1}^{K}\left[l_j\frac{\partial^2 f_0}{\partial x^j\,\partial x^j}+\alpha_j\frac{\partial f_0}{\partial x^j}\right]$$

resp.

$$\mathscr{A}_i f_i=\lambda_i\cdot\tfrac{1}{2}\nabla_i^2 f_i,\qquad \text{for } 1\leqq i\leqq L,$$

where ∇_i^2 denotes the Laplace operator in $\mathcal{M}^i$.

Then the process

$$X(t)=\big(X_0(t),\ldots,X_L(t)\big)$$

is the Brownian motion in $\mathcal{M}$ whose differential generator is given by (3.12).

The general version of the theorem by Yosida mentioned earlier now follows from Proposition 3.3:

Corollary. *Let $\mathcal{M}$ be a 2-dimensional complete Riemannian manifold of class $\mathscr{A}\mathscr{B}$ and with nonvanishing Gaussian curvature K, then there exists only one Brownian motion on $\mathcal{M}$, namely, the Wiener process.*

---

[39] [14] p. 192, Theorem 6.2.

Proof. The restricted[40] holonomy group Ψ^0 of $\mathscr{M}$ is a closed connected subgroup of SO(2) so either it contains only the identity map or it is all of SO(2). Now, the existence of a point p at which the curvature is different from zero implies (use the Gauss-Bonnet formula) that $\Psi^0(p) \neq \{\text{identity map}\}$ and consequently $\Psi^0 = \mathrm{SO}(2)$, which means $\mathscr{M}$ is irreducible.

The result now follows from (I) of Proposition 3.3.

4. Appendix

The purpose of the present section is to establish the general notation relating to some basic concepts from differential geometry[41] which are being used throughout this paper.

4.1. $\mathscr{M}$ will always denote a fixed N-dimensional complete Riemannian manifold with metric tensor g and associated Riemannian connexion ∇. In particular, $\mathscr{M}$ is assumed to be connected and separable. For $p \in \mathscr{M}$ the tangent and cotangent spaces at p are denoted $\mathscr{M}_p$ resp. $\mathscr{M}_p^*$, and in general we write $(\mathscr{M}_p)_s^r$ for the space of tensors of type (r, s) over $\mathscr{M}_p$. Occasionally we shall need to view an object $T \in (\mathscr{M}_p)_s^r$ either as being a linear functional on the space

$$\underbrace{\mathscr{M}_p^* \otimes \cdots \otimes \mathscr{M}_p^*}_{r\,\text{times}} \otimes \underbrace{\mathscr{M}_p \otimes \cdots \otimes \mathscr{M}_p}_{s\,\text{times}}$$

or as being an $(r+s)$-linear functional on the space

$$\mathscr{M}_p^* \times \cdots \times \mathscr{M}_p^* \times \mathscr{M}_p \times \cdots \times \mathscr{M}_p .$$

In both cases we write

$$\langle T, S \rangle \quad \text{or} \quad \langle S, T \rangle$$

for the action of T on S.

An inner product and a corresponding norm is defined on $(\mathscr{M}_p)_s^r$ in the following fashion. Choose an orthonormal basis $\{e_i\}_{i=1}^N$ for $\mathscr{M}_p$ and let $\{\varepsilon^j\}_{j=1}^N$ be the corresponding dual basis for $\mathscr{M}_p^*$; then the collection

$$\{e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes \varepsilon^{j_1} \otimes \cdots \otimes \varepsilon^{j_s}\} \begin{array}{l} 1 \leqq j_l \leqq N \\ 1 \leqq i_k \leqq N \end{array}$$

forms a basis for $(\mathscr{M}_p)_s^r$ which now is taken to be orthonormal. This yields the desired inner product, and we obtain the counterpart of Schwarz's inequality

$$|\langle T, S \rangle| \leqq \|T\| \cdot \|S\| .$$

Let f be a real valued differentiable function on $\mathscr{M}$. With f we associate two tensorfields, namely, (1) Df = the differential of f, and (2) $D^2 f$ which is defined as follows[42].

[40] This is defined in the same way as Ψ, but referring only to parallel translations along nullhomotopic curves.

[41] Our basic references on this subject are the books: [8] and [14]. In order to avoid misunderstandings we remark that the word "differentiable" here means "of class C^∞" unless otherwise indicated.

[42] In general we define tensorfields T by exhibiting for each $p \in \mathscr{M}$ the "value" $T_p \in (\mathscr{M}_p)_s^r$ of T at p. The degree of smoothness of T then depends upon the smoothness of this assignment and is found by expressing the definition in terms of local coordinates.

Let $p \in \mathcal{M}$, $X \in \mathcal{M}_p$, $Y \in \mathcal{M}_p$. Choose a curve $\gamma(t)$, $0 \leqq t \leqq 1$, with

$$\gamma(0) = p; \qquad \gamma_*(0) = \text{tangent to } \gamma(\cdot) \text{ at } 0 = X,$$

and define the vectorfield Y_t along $\gamma(t)$ by requiring

$$Y_0 = Y, \qquad \nabla_{\gamma_*(t)} Y_t = 0,$$

i.e., Y_t is the parallel translate of Y along $\gamma(\cdot)$. Finally set

$$\varphi(t) = \langle (Df)_{\gamma(t)}, Y_t \rangle;$$

then $D^2 f$ is defined at p by the conditions

$$(D^2 f)_p \in (\mathcal{M}_p)_2^0, \qquad \langle (D^2 f)_p, X \otimes Y \rangle = \varphi'(0).$$

To see that this definition is independent of $\gamma(\cdot)$ introduce a local coordinate system (V, α)[43] around p and set

$$\alpha(p) = (x_0^i), \quad X = \xi_0^i \partial_i, \quad Y = \eta_0^i \partial_i,$$
$$\alpha(\gamma(t)) = (x^i(t)), \quad Y_t = \eta^i(t) \partial_i.$$

Then[44]

$$\alpha(\gamma(0)) = (x_0^i), \quad \gamma_*(0) = \xi_0^i \partial_i,$$
$$\frac{d\eta^k}{dt} + \Gamma_{ij}^k \eta^i \frac{dx^j}{dt} = 0, \qquad \text{for } 1 \leqq k \leqq N$$

so

$$\langle (Df)_{\gamma(t)}, Y_t \rangle = \eta^k(t) \partial_k f(\gamma(t)),$$
$$\langle (D^2 f)_p, X \otimes Y \rangle = \left[\frac{d}{dt} \eta^k(t) \partial_k f(\gamma(t)) \right]_{t=0} = \left[\frac{d\eta^k}{dt} \partial_k f + \eta^j (\partial_i \partial_j f) \frac{dx^i}{dt} \right]_{t=0}$$
$$= (\partial_i \partial_j f - \Gamma_{ij}^k \partial_k f) \eta_0^i \xi_0^j.$$

Thus, the coordinate expressions are

$$\begin{aligned} Df &= (\partial_i f)\, dx^i \\ D^2 f &= (\partial_i \partial_j f - \Gamma_{ij}^k \partial_k f)\, dx^i \otimes dx^j, \end{aligned} \tag{4.1}$$

from which one sees that $D^2 f$ is symmetric.

If we write for f and h differentiable

$$2(Df)(Dh) = (Df) \otimes (Dh) + (Dh) \otimes (Df)$$

[43] $V \subset \mathcal{M}$, α: $V \to E^N = N$-dimensional Euclidean space. We describe points in E^N by the letters $\bar{x} = (x^i)$ and differentiation in the coordinate directions by

$$\partial_i f = \frac{\partial}{\partial x^i} (f \circ \alpha^{-1}).$$

The ∂_i are viewed as being tangent vectors to $\mathcal{M}$, and form at each point p a basis for $\mathcal{M}_p$ whose corresponding dual basis is $\{dx^i\}_{i=1}^N$.

[44] The Γ_{ij}^k are the Christoffel symbols:

$$\nabla_{\partial_i} \partial_j = \Gamma_{ij}^k \partial_k,$$

where repetition of the index k means: sum over k, $1 \leqq k \leqq N$.

we obtain the useful formalism

$$D(fh)=fDh+hDf,$$
$$D^2(fh)=fD^2h+2(Df)(Dh)+hD^2f.$$

The Laplace operator ∇^2 may now be defined in the following invariant fashion. If $I_p(\cdot,\cdot)$ denotes inner product on $(\mathcal{M}_p)_0^2$ define $(G^2f)_p\in(\mathcal{M}_p)_0^2$ by

$$I_p\big((G^2f)_p,\, X\otimes Y\big)=\langle (D^2f)_p,\, X\otimes Y\rangle$$

for all X, Y in $\mathcal{M}_p$. Then we can set

$$(\nabla^2 f)(p)=\langle g_p,(G^2f)_p\rangle$$

where g_p is the metric tensor evaluated at p, and f as usual is a differentiable function on $\mathcal{M}$.

In coordinates this reads

$$\nabla^2 f=g^{rs}(\partial_r\partial_s f-\Gamma^k_{rs}\partial_k f)=|g|^{-\frac{1}{2}}\partial_k(|g|^{\frac{1}{2}}g^{kl}\partial_l f),$$

where

$$g=g_{ij}dx^i\otimes dx^j,$$

$\{g^{ij}\}$ is the inverse of the matrix $\{g_{ij}\}$, and $|g|$ is the determinant of $\{g_{ij}\}$.

4.2. The distance between a pair of points p,q in $\mathcal{M}$ will be denoted $d(p,q)$. There is always a geodesic segment $\gamma(t)$, $a\leqq t\leqq b$, joining p and q for which

$$\text{length of } \gamma=|\gamma|=d(p,q).$$

Such a segment will be called a short geodesic segment and denoted $\widehat{pq}$. Although in general the choice of γ is not unique unless p and q are close together this convention will not cause any difficulties in the following. We can now define the concept of Hölder continuity of a tensorfield:

Definition 4.1. Let $0<\alpha\leqq 1$. *A tensorfield* T *of type* (r,s) *defined on* $\mathcal{M}$ *is said to be Hölder continuous of order* α *if there exists a constant* $K>0$ *so*

$$\sup\big[|\langle T_q,\tau_{pq}\theta\rangle-\langle T_p,\theta\rangle|:\ \theta\in(\mathcal{M}_p^*)_s^r,\ \|\theta\|\leqq 1\big]\leqq K d^\alpha(p,q), \tag{4.2}$$

for all p,q *in* $\mathcal{M}$.

Here τ_{pq} denotes parallel translation along any short geodesic segment $\widehat{pq}$, and part of the requirement is that the inequality must hold for every choice of $\widehat{pq}$. If we are dealing with a function f rather than a tensorfield we substitute the condition

$$|f(p)-f(q)|\leqq K d^\alpha(p,q) \tag{4.3}$$

for (4.2). In case every point p_0 has a neighborhood V_{p_0} such that (4.2) (or (4.3)) holds with $K=K(V_{p_0})$ we say that T (or f) is locally Hölder continuous of order α. The situation that will concern us the most is the one where f is twice continuously differentiable ($f\in C^2$) and D^2f is Hölder continuous of order α. We then write $f\in C^{(2,\alpha)}$ or $f\in C^{(2,\alpha)}_{\text{loc}}$ depending upon whether this holds globally or locally. The following lemmas are easily proved

Lemma 4.1. *The function $f \in C^2$ is contained in $C_{\mathrm{loc}}^{(2,\alpha)}$ if and only if every point p_0 in $\mathcal{M}$ is contained in a coordinate neighborhood V_0 such that*[45]

$$|\partial_i \partial_j f(\bar{x}) - \partial_i \partial_j f(\bar{y})| \leqq K \cdot |\bar{x} - \bar{y}|^\alpha$$

for some constant K, $1 \leqq i, j \leqq N$, and all $\bar{x}, \bar{y}$ in $\beta(V_0)$, where β denotes the coordinate map: $V_0 \to E^N$.

Lemma 4.2. *Let $f \in C_{\mathrm{loc}}^{(2,\alpha)}$ and let $p_0 \in \mathcal{M}$. Then there exists a neighborhood V of p_0 and a function f_0 with compact support such that $f_0 \in C^{(2,\alpha)}$ and $f_0(p) = f(p)$ for $p \in V$.*

Lemma 4.3. *Let $0 < \alpha \leqq 1$ and let $f \in C^{(2,\alpha)}$ and have compact support. Then there exists a constant K so*

$$|f(\exp_p X) - f(p) - \langle (Df)_p, X \rangle - \tfrac{1}{2} \langle (D^2 f)_p, X \otimes X \rangle| \leqq K \cdot \|X\|^{2+\alpha}$$

for all $p \in \mathcal{M}$ and $X \in \mathcal{M}_p$.

Here "$\exp_p$" is the exponential map: $\mathcal{M}_p \to \mathcal{M}$, and the estimate in Lemma 4.3 is obtained from the Taylor expansion of the function

$$\varphi(t) = f(\exp_p tX), \qquad -\infty < t < \infty,$$

about $t = 0$.

4.3. In this section we shall take a look at the "good" property of Riemannian manifolds on which the proof of our main theorem (Theorem (2.1)) is based. Let $p \in \mathcal{M}$ and let $X, Y \in \mathcal{M}_p$, set

$$\gamma(t, 0) = \exp_p(tX), \qquad -\infty < t < \infty,$$

and let τ_t denote parallel translation along $\gamma(\cdot, 0)$ from $p = \gamma(0,0)$ to $p_t = \gamma(t, 0)$, then define

$$\gamma(t, s) = \exp_{p_t}(s \cdot \tau_t Y), \qquad -\infty < s < \infty.$$

The problem[46] to be considered is that of finding an upper bound for the distance between the points $\gamma(0,1)$ and $\gamma(1,1)$ in terms of X and Y. We shall prove

Lemma 4.4. *To every point $p \in \mathcal{M}$ and every compact neighborhood V of p there exists a constant $c \geqq 0$ such that for $q \in V$ and $X, Y \in \mathcal{M}_q$ with $\|X\| \leqq 1$ and $\|Y\| \leqq 1$ we have*

$$d(\gamma(0,1), \gamma(1,1)) \leqq \|X\| \cdot (1 + c\|Y\|^2).$$

In order to prove the lemma we need a few results concerning the geometry of the tangent bundle $\mathcal{T}\mathcal{M}$ of $\mathcal{M}$. These we list first, referring the reader to the paper: [19] for proofs and further information.

Let (V, α), $\alpha: V \to U \subset E^N$, be a coordinate system on $\mathcal{M}$ and let $TV = \bigcup_{p \in V} \mathcal{M}_p$ be the part of $\mathcal{T}\mathcal{M}$ lying above V. TV is identified with $U \times E^N$ by letting a pair (p, X), $p \in V$ and $X \in \mathcal{M}_p$, correspond to a pair $(\bar{x}, \bar{v})$ where $\bar{x} = \alpha(p)$ and $X \sim v^i \partial_i$. This makes TV into a coordinate patch in $\mathcal{T}\mathcal{M}$. If $\gamma(t)$, $a \leqq t \leqq b$, is a curve in V with $\gamma(a) = q$ and X is an element of $\mathcal{M}_q$, then the lift $\hat{\gamma}$ of γ to (q, X) is the curve in

[45] $|\bar{x} - \bar{y}|^2 = \sum_{i=1}^N (x^i - y^i)^2$, where $\bar{x} = (x^i)$ and $\bar{y} = (y^i)$.

[46] See also [3] p. 247ff.

$\mathcal{TM}$ which satisfies

$$\hat{\gamma}(a)=(q,X), \qquad \pi\circ\hat{\gamma}(t)=\gamma(t), \qquad a\leqq t\leqq b,$$
$$\frac{dv^i(t)}{dt}=-\Gamma^i_{jk}(\bar{x}(t))\frac{dx^j(t)}{dt}\cdot v^k(t)$$

where we have written in coordinates

$$\gamma(t)\sim\bar{x}(t)=(x^i(t)); \qquad \hat{\gamma}(t)\sim(\bar{x}(t),\bar{v}(t)),$$

and where $\pi\colon \mathcal{TM}\to\mathcal{M}$ is the projection map, i.e., $\pi(p,X)=p$. By piecing together this definition extends to all of $\mathcal{M}$, so it makes sense to talk about the lift of an arbitrary curve in $\mathcal{M}$. Note that $\hat{\gamma}$ may be visualized as consisting of a pair $(\gamma(\cdot), X.)$ where $\gamma(\cdot)$ is a curve in $\mathcal{M}$ and $X.$ is a parallel vectorfield along γ. In particular, if γ has the property that

$$\hat{\gamma}(t)=(\gamma(t),\gamma_*(t)) \qquad \text{for all } t$$

which in coordinates reads

$$\frac{dx^i}{dt}=v^i, \qquad \frac{dv^i}{dt}=-\Gamma^i_{jk}v^jv^k \tag{4.4}$$

then γ is a geodesic. Let $P=(p,X)\in\mathcal{TM}$ then there exists a unique integral curve $\Phi(t,P)$, $-\infty<t<\infty$, of the system (4.4) satisfying

$$\Phi(0,P)=P.$$

$\Phi(\cdot,P)$ is the lift to P of the geodesic $\pi\circ\Phi(\cdot,P)$, and for all t, s

$$\Phi(t+s,P)=\Phi(t,\Phi(s,P)).$$

Considered as a function: $E^1\times\mathcal{TM}\to\mathcal{TM}$ $\Phi(\cdot,\cdot)$ is of class C^∞, while for t fixed the map

$$T_t\colon P\to T_tP=\Phi(t,P)$$

is a diffeomorphism of $\mathcal{TM}$, and as one easily sees the family $(T_t, -\infty<t<\infty)$ constitutes a one-parameter group of diffeomorphisms of $\mathcal{TM}$, the so-called geodesic flow. For $p\in\mathcal{M}$, $X\in\mathcal{M}_p$ the geodesic γ with

$$\gamma(0)=p, \qquad \gamma_*(0)=X$$

is now given by

$$\gamma(t)=\exp_p(tX)=\pi\circ T_tP, \qquad P=(p,X),$$

and in particular we get the exponential map factored

$$\exp_pX=\pi\circ T(P), \qquad P=(p,X),$$

where we have set $T=T_1$.

Finally, let $P=(p,X)\in\mathscr{T}\mathscr{M}$ and let $Y\in\mathscr{M}_p$. The horizontal lift $\mathscr{H}_P(Y)$ of Y to P is then defined to be the tangent at P to the lift $\hat\gamma$ to P of any curve γ with

$$\gamma(0)=p, \qquad \gamma_*(0)=Y, \tag{4.5}$$

thus

$$\mathscr{H}_P(Y)=\hat\gamma_*(0)\in(\mathscr{T}\mathscr{M})_P,$$

this definition being independent of the choice of γ satisfying (4.5).

We now have the tools needed for the proof of Lemma 4.4. Let $p\in\mathscr{M}$ and let $X_0, Y_0\in\mathscr{M}_p$ with $\|X_0\|=\|Y_0\|=1$. For $-\infty<u,v<\infty$ set

$$c(t,u)=\exp_p(tuX_0)=\pi\circ\Phi(t,(p,uX_0))$$

and let $G(\cdot\,;u,v)$ denote the lift of $c(\cdot\,;u)$ to the point (p,vY_0) in $\mathscr{T}\mathscr{M}$. The curve $\gamma(t,1)$ corresponding to[47] the vectors $X=uX_0$ and $Y=vY_0$ is then

$$\gamma(t,1)=\pi\circ T\circ G(t;u,v)$$

and consequently

$$d(\gamma(0,1),\gamma(1,1))\leqq\int_0^1\|\gamma_*(t,1)\|\,dt\leqq\left[\int_0^1\|\gamma_*(t,1)\|^2\,dt\right]^{\frac12}$$

with[48]

$$\gamma_*(t,1)=\pi_*\circ T_*\circ G_*(t;u,v)$$

and

$$G_*(t;u,v)=\mathscr{H}_{G(t;u,v)}(c_*(t;u,v)).$$

The integral

$$E(u,v)=\int_0^1\|\gamma_*(t,1)\|^2\,dt$$

is a C^∞ function of the variables (p,X_0,Y_0,u,v). For the moment we shall keep (p,X_0,Y_0) fixed and expand $E(u,v)$ in terms of u and v. To this end note first that $c(t;u)=c(ut;1)$ so

$$c_*(t;u)=u\,c_*(ut;1); \qquad G(t;u,v)=G(ut;1,v)$$

thus

$$\|\gamma_*(t,1)\|^2=u^2\left\|\pi_*\circ T_*\circ\mathscr{H}_{G(ut;1,v)}(c_*(ut;1))\right\|^2$$

and

$$E(u,v)=u\int_0^u\left\|\pi_*\circ T_*\circ\mathscr{H}_{G(s;1,v)}(c_*(s;1))\right\|^2ds.$$

Set

$$F(u,v)=\begin{cases}\|\pi_*\circ T_*\circ\mathscr{H}_{(p,vY_0)}(X_0)\|^2 & \text{for } u=0\\ u^{-2}E(u,v) & \text{for } u\neq0\end{cases}$$

then

$$F(u,v)=\int_0^1\left\|\pi_*\circ T_*\circ\mathscr{H}_{G(ut;1,v)}(c_*(ut;1))\right\|^2dt,$$

[47] In the notation of the lemma.

[48] For a map Ψ, Ψ_* denotes the differential of Ψ. The functions $c(\cdot\,;u)$, $G(\cdot\,;u,v)$ etc. are thought of as being functions of t parametrized by u, v. So for instance, $c_*(t;u)$ means $\frac{\partial}{\partial t}c(t;u)$.

so F is of class C^∞ in the variables (p, X_0, Y_0, u, v), and the same can then also be said about the function

$$D(u,v) = \int_0^1 (1-\theta) \frac{\partial^2 F}{\partial v^2}(u, \theta v)\, d\theta$$

which is the remainder in the expansion

$$F(u,v) = F(u,0) + v \frac{\partial F}{\partial v}(u,0) + v^2 D(u,v).$$

On the other hand, it will be shown in a moment that

$$F(u,0) \equiv 1, \qquad \frac{\partial F}{\partial v}(u,0) \equiv 0. \tag{4.6}$$

Consequently

$$E(u,v) = u^2 [1 + v^2 D(u,v)]$$

and the conclusion of the lemma follows ($u^2 \sim \|X\|^2$ and $v^2 \sim \|Y\|^2$). So it only remains to prove (4.6). First

$$T \circ G(t;\, u, 0) = G(t;\, u, 0)$$

so

$$\pi \circ T \circ G(t; u, 0) = \pi \circ G(t; u, 0) = c(t; u),$$

that is for $v=0$

$$\|\gamma_*(t,1)\|^2 = \|c_*(t;u)\|^2 = u^2 \|X_0\|^2 = u^2,$$

hence

$$F(u,0) = u^{-2} E(u,0) = u^{-2} \int_0^1 \|\gamma_*(t,1)\|^2\, dt = 1,$$

which proves the first part of (4.6). To prove the second part it obviously suffices to show

$$\frac{\partial}{\partial v} \|\gamma_*(t,1)\|^2 \big|_{v=0} = 0.$$

So consider for u fixed the expression

$$\gamma(t,1) = \pi \circ T \circ G(t; u, v)$$

as a map from the (t,v)-plane into $\mathcal{M}$. Letting ∇ denote covariant differentiation we have (with slight abuse of notation)

$$\frac{\partial}{\partial v} \|\gamma_*(t,1)\|^2 = \nabla_{\frac{\partial}{\partial v}} g\left(\frac{\partial \gamma}{\partial t}, \frac{\partial \gamma}{\partial t}\right) = 2g\left(\frac{\partial \gamma}{\partial t}, \nabla_{\frac{\partial}{\partial v}} \frac{\partial \gamma}{\partial t}\right) = 2g\left(\frac{\partial \gamma}{\partial t}, \nabla_{\frac{\partial}{\partial t}} \frac{\partial \gamma}{\partial v}\right).$$

Moreover, for $v=0$, the vectorfield $\frac{\partial \gamma}{\partial v}$ is parallel along the curve

$$\pi \circ T \circ G(t;\, u, 0) = \exp_p(t u X_0),$$

and consequently

$$\nabla_{\frac{\partial}{\partial t}} \frac{\partial \gamma}{\partial v}\bigg|_{v=0} = 0.$$

Inserting this in the above equation we obtain

$$\frac{\partial}{\partial v}\|\gamma_*(t,1)\|^2|_{v=0}=0,$$

as claimed. This finishes the proof of the lemma.

It will at times be necessary to assume that the manifold $\mathscr{M}$ possesses certain extra properties besides those already mentioned. These we have listed below as axioms $\mathscr{A}$ and $\mathscr{B}$. A manifold satisfying axiom $\mathscr{A}$, or axiom $\mathscr{B}$, or both will be referred to as being of class $\mathscr{A}$, or class $\mathscr{B}$, or class $\mathscr{A}\mathscr{B}$ respectively.

Axiom $\mathscr{A}$. *For every compact set $K\subset\mathscr{M}$ and every $\varepsilon>0$ there exists a C^∞ function f with compact support, so*

$$0\leqq f\leqq 1,\quad f(p)=1 \text{ for } p\in K,$$

and

$$\sup_{p\in\mathscr{M}}[\max\{\|(Df)_p\|,\|(D^2f)_p\|\}]\leqq\varepsilon.$$

In the formulation of the next axiom the notation is the one introduced in connection with Lemma 4.4:

Axiom $\mathscr{B}$. *There exists a constant $c\geqq 0$ such that for $p\in\mathscr{M}$, $X\in\mathscr{M}_p$, $Y\in\mathscr{M}_p$ with $\|X\|\leqq 1$ and $\|Y\|\leqq 1$ we have*

$$d(\gamma(0,1),\gamma(1,1))\leqq\|X\|(1+c\|Y\|^2),$$

where

$$\gamma(t,0)=\exp_p(tX)=p_t,\qquad \gamma(t,s)=\exp_{p_t}(sY_t),$$

Y_t being the parallel translate of Y along $\gamma(\cdot,0)$.

We should like to point out that it is possible that every complete manifold is of class $\mathscr{A}$. At least we do not know of any counterexample. Under any circumstances, both axioms are satisfied in the following important special cases:

(I) $\mathscr{M}=N$-dimensional Euclidean space,

(II) $\mathscr{M}$ is compact.

Here (I) is obvious, while (II) is a consequence of Lemma 4.4. $\mathscr{B}$ is also satisfied if $\mathscr{M}$ is a homogeneous space. This follows in a similar fashion from Lemma 4.4.

4.4. Let U be an N-dimensional real vector space furnished with an inner product $I(\cdot,\cdot)$, and let ν be a Borel measure[49] on U with $\nu(U)\leqq 1$. Such a measure will be called a subprobability measure on U. For $m\geqq 1$, the m-th absolute moment $|\mu|^{(m)}$ of ν is defined by

$$|\mu|^{(m)}=\int_U\|v\|^m\nu(dv),$$

where $\|v\|^2=I(v,v)$, and also if $|\mu|^{(m)}<\infty$ we define, $\mu^{(m)}$, the m-th moment of ν, to be the element of $U\otimes\cdots\otimes U$ (m times) which satisfies

$$\langle\mu^{(m)},\theta^1\otimes\cdots\otimes\theta^m\rangle=\int_U\left[\prod_{j=1}^m\langle v,\theta^j\rangle\right]\nu(dv)$$

[49] U is furnished with the topology associated with the metric induced by $I(\cdot,\cdot)$. In the following "measurable" means "Borel measurable."

for all $(\theta^1, \dots, \theta^m)$ in $U^* \times \cdots \times U^*$. Here U^* is the dual of U. Clearly

$$\left| \prod_{j=1}^{m} \langle v, \theta^j \rangle \right| \leqq \left[\prod_{j=1}^{m} \|\theta^j\| \right] \cdot \|v\|^m = \|\theta^1 \otimes \cdots \otimes \theta^m\| \cdot \|v\|^m$$

so the integral is well-defined and also

$$|\langle \mu^{(m)}, \theta^1 \otimes \cdots \otimes \theta^m \rangle| \leqq |\mu|^{(m)} \cdot \|\theta^1 \otimes \cdots \otimes \theta^m\|.$$

We use the notation

$$\mu^{(m)} = \int_U v \otimes \cdots \otimes v \, v(dv),$$

and also if $A \subset U$ is measurable

$$\int_A v \otimes \cdots \otimes v \, v(dv) = \int_U v \otimes \cdots \otimes v \, \tilde{v}(dv)$$

with $\tilde{v}$ given by: $\tilde{v}(B) = v(AB)$. If $\{e_i\}_{i=1}^N$ is a basis for U then

$$v = v^i e_i, \qquad v \otimes \cdots \otimes v = v^{i_1} \cdots v^{i_m} e_{i_1} \otimes \cdots \otimes e_{i_m}$$

and if follows that

$$\int_A v \otimes \cdots \otimes v \, v(dv) = \left[\int_A v^{i_1} \cdots v^{i_m} v(dv) \right] e_{i_1} \otimes \cdots \otimes e_{i_m}.$$

In particular, if the basis is orthonormal

$$\left\| \int_A v \otimes \cdots \otimes v \, v(dv) \right\|^2 = \sum_{i_1, \dots, i_m} \left[\int_A v^{i_1} \cdots v^{i_m} v(dv) \right]^2$$

from which one easily obtains

$$\|\mu^{(1)}\|^2 \leqq |\mu^{(2)}| \leqq \sqrt{N} \, \|\mu^{(2)}\| \leqq \sqrt{N} \, |\mu|^{(2)}. \tag{4.7}$$

If V is another real vector space and τ is a linear map from U to V then τ induces in a natural way a linear map, which will be denoted by τ also, from the space $U \otimes \cdots \otimes U$ to the space $V \otimes \cdots \otimes V$, namely,

$$\tau(a^{j_1 \dots j_k} u_{j_1} \otimes \cdots \otimes u_{j_k}) = a^{j_1 \dots j_k} (\tau u_{j_1}) \otimes \cdots \otimes (\tau u_{j_k}).$$

Moreover, the measure v on U induces a measure τv on V

$$\tau v(A) = v(\tau^{-1}(A)),$$

and if $\mu_v^{(k)}$ and $\mu_{\tau v}^{(k)}$ are the k-th moments of v resp. τv we have the formula

$$\mu_{\tau v}^{(k)} = \tau \mu_v^{(k)}. \tag{4.8}$$

Let us now return to our manifold $\mathcal{M}$. Assume there is defined for each $p \in \mathcal{M}$ a subprobability measure v_p on $\mathcal{M}_p$ satisfying

$$\int_{\mathcal{M}_p} \|v\|^2 v_p(dv) < \infty;$$

then we may construct tensorfields μ, $\mu^{(2)}$ and σ^2 by setting

$$(\mu)_p = \int_{\mathcal{M}_p} v \, v_p(dv); \qquad (\mu^{(2)})_p = \int_{\mathcal{M}_p} v \otimes v \, v_p(dv); \qquad \sigma^2 = \mu^{(2)} - \mu \otimes \mu.$$

These tensorfields need of course not exhibit any kind of regularity in their dependence upon p unless further conditions are imposed upon the ν_p. Such conditions will be set up next.

To this end, recall that for measures m_1, m_2 defined on the same space U the quantity[50]

$$|m_1 - m_2| = \sup \left[\left| \int_U f(v)\big(m_1(dv) - m_2(dv)\big) \right| : f \in C(U), \|f\| \leqq 1 \right] \tag{4.9}$$

is used to measure the deviation between m_1 and m_2. Now, in our case, the measures ν_p are defined on different spaces and thus cannot be compared directly by the use of (4.9). But, by introducing certain isomorphisms τ_{pq} between the spaces $\mathcal{M}_p$ and $\mathcal{M}_q$ we can instead compare ν_q and $\tau_{pq}\nu_p$. So fix p and q in $\mathcal{M}$ and let τ_{pq} denote parallel translation along a short geodesic $\widehat{pq}$. Then define $|\nu_p - \nu_q|$ to be[51] the supremum of the numbers $|\nu_q - \tau_{pq}\nu_p|$ taken over all such $\widehat{pq}$. We can now talk about continuity of the map $p \to \nu_p$ and in particular we can define Hölder continuity of order α by the condition that there shall exist a constant K so for every p, q in $\mathcal{M}$

$$|\nu_p - \nu_q| \leqq K \cdot d^\alpha(p, q). \tag{4.10}$$

Here $|\nu_p - \nu_q| \leqq 2$, so it suffices to verify (4.10) for $d(p, q) \leqq$ some positive constant. Local Hölder continuity is defined in a similar fashion.

References

1. Billingsley, P.: Convergence of Probability Measures. New York: Wiley 1968
2. Blumenthal, R. M., Getoor, R. K.: Markov Processes and Potential Theory. New York: Academic Press 1968
3. Cartan, É.: Lecons sur la Géométrie des Espaces de Riemann. Paris: Gauthier-Villars 1963
4. Chung, K. L.: A Course in Probability Theory. New York: Harcourt, Brace and World 1968
5. Dynkin, E. B.: Die Grundlagen der Theorie der Markoffschen Prozesse. Berlin-Heidelberg-New York: Springer-Verlag 1961
6. Dynkin, E. B.: Markov Processes. New York: Academic Press 1965
7. Gangolli, R.: On the construction of certain diffusions on a differentiable manifold. Z. Wahrscheinlichkeitstheorie verw. Gebiete **2**, 406–419 (1964)
8. Helgason, S.: Differential Geometry and Symmetric Spaces. New York: Academic Press 1962
9. Hunt, G. A.: Semi-groups of measures on Lie groups. Trans. Amer. Math. Soc. **81**, 264–293 (1956)
10. Itô, K.: Stochastic differential equations in a differentiable manifold. Nagoya Math. J. **1**, 35–47 (1950)
11. Itô, K.: Stochastic differential equations in a differentiable manifold (2). Mem. Coll. Sci. Kyôto Univ. **28**, 81–85 (1953)
12. Itô, K.: The Brownian Motion and Tensor Fields on Riemannian Manifold. Proc. Intern. Congr. Math., Stockholm, p. 536–539, 1963
13. Kelley, J. L.: General Topology. Princeton N.J.: Van Nostrand 1955
14. Kobayashi, S., Nomizu, K.: Foundations of Differential Geometry. New York: Wiley 1963
15. McKean, H. P., Jr.: Brownian motions on the 3-dimensional rotation group. Mem. Coll. Sci. Kyôto Univ. **33**, 25–38 (1960)
16. McKean, H. P., Jr.: Stochastic Integrals. New York: Academic Press 1969
17. Miranda, C.: Partial Differential Equations of Elliptic Type. 2nd ed.. Berlin-Heidelberg-New York: Springer-Verlag 1970
18. Prokhorov, Yu. V.: Convergence of random processes and limit theorems in probability theory. Theor. Probability Appl. **1**, 157–214 (1956)

[50] $C(U)$ is the class of bounded continuous functions on U with norm $\|f\| = \sup_{u \in U} |f(u)|$.

[51] It is easy to show that the definition is symmetric in p and q.

19. Sasaki, S.: On the differential geometry of tangent bundles of Riemannian manifolds. Tôhoku Math. J. **10**, 338–354 (1958)
20. Skorokhod, A. V.: Limit theorems for stochastic processes. Theor. Probability Appl. **1**, 261–290 (1956)
21. Skorokhod, A. V.: Limit theorems for Markov processes. Theor. Probability Appl. **3**, 202–246 (1958)
22. Stone, C.: Weak convergence of stochastic processes defined on semi-infinite time intervals. Proc. Amer. Math. Soc. **14**, 694–696 (1963)
23. Trotter, H. F.: Approximation of semi-groups of operators. Pacific J. Math. **8**, 887–919 (1958)
24. Yosida, K.: Brownian motion on the surface of the 3-sphere. Ann. Math. Statist. **20**, 292–296 (1949)
25. Yosida, K.: Brownian motion in a homogeneous Riemannian space. Pacific J. Math. **2**, 263–270 (1952)
26. Yosida, K.: Functional Analysis. New York: Academic Press 1965

Erik Jørgensen
Danmarks Ingeniørakademi
Danmarksgade 17
DK-9000 Aalborg
Denmark

(Received October 7, 1972)