The Annals of Statistics
1990, Vol. 18, No. 2, 832–849

NONPARAMETRIC ESTIMATION OF A PROBABILITY DENSITY ON A RIEMANNIAN MANIFOLD USING FOURIER EXPANSIONS

BY HARRIE HENDRIKS

Katholieke Universiteit Nijmegen

Supposing a given collection $y_1, \dots, y_N$ of i.i.d. random points on a Riemannian manifold, we discuss how to estimate the underlying distribution from a differential geometric viewpoint. The main hypothesis is that the manifold is closed and that the distribution is (sufficiently) smooth. Under such a hypothesis a convergence arbitrarily close to the $N^{-1/2}$ rate is possible, both in the L_2 and the L_∞ senses.

1. Introduction. Potential applications of differential geometry in statistics are at least two-fold: In a wide variety of situations the statistician is faced with a sample space that is no longer Euclidean but is more appropriately described by a manifold. On the other hand, the statistical model seems to be a geometrical object per se. Many statisticians have paid attention to formulating and explaining properties of statistical models in geometrical language. An interesting survey with a large number of references can be found in Barndorff-Nielsen, Cox and Reid (1986).

Although we are convinced that still a lot may be done in this area, in particular when general parametric models are considered, in this paper we focus on a special statistical problem where the sample space is a manifold.

Statistical theory on sample spaces like the circle and the two-dimensional sphere dates back to Watson and Williams (1956). A survey of the state of the art as well as many references can be found in Mardia (1972), Watson (1983) and Fisher, Lewis and Embleton (1987). Testing for uniformity on a compact homogeneous space has been considered in Beran (1968), and Giné (1975a), more generally, deals with the same problem for a compact Riemannian manifold. A two-sample permutation test on a compact Riemannian manifold is developed in Wellner (1979). For testing symmetry on such a manifold, see Jupp and Spurr (1983).

In this paper we focus on the discussion of Devroye and Györfi (1985) of an L_1-convergent nonparametric trigonometric series estimator of a density on the circle. We will generalize this discussion to other Riemannian manifolds comprising the large class of *closed* (i.e., compact and without boundary) Riemannian manifolds. Moreover the theory can be extended to the class of *homogeneous* manifolds, so that the group of isometries acts transitively.

We will get convergence results both in the L_2 and the L_∞ senses. (Note that for compact manifolds, L_2 convergence implies L_1 convergence; moreover

Received February 1987; revised June 1989.

AMS 1980 *subject classifications*. Primary 62G05; secondary 35P20, 58G11, 58G25.

Key words and phrases. Nonparametric density estimation, L_2 convergence, L_∞ convergence, closed manifolds, homogeneous manifolds, Laplace–Beltrami operator, Fourier theory, convergence of generalized zeta functions.

for L_∞ estimation a partial result on L_2 estimation is needed.) The necessary ingredients are made explicit for the circle, the two-dimensional sphere, the projective plane, the Lie group SO(3) and Euclidean space. As a special case, in Section 3.4, we find the Fourier integral estimate for the real line described by Konakov (1973) and Davis (1975, 1977).

It is well known that a Riemannian manifold is equipped with a symmetric positive second order partial differential operator, the Laplace–Beltrami operator. Given a positive self-adjoint extension, application of the spectral theorem will give rise to a Fourier transform defined on L_2 functions, and in particular on L_2 densities. For closed manifolds the spectrum is discrete and the Fourier transform naturally leads to an orthonormal basis of L_2, which may lead to an orthonormal series-type density estimator. In order to address an audience as large as possible, we will restrict the theoretical discussion to closed manifolds, although many of the results are formulated in such a way that they remain valid in the context of not necessarily closed, homogeneous manifolds. How to obtain the spectral function for nonclosed Riemannian manifolds is indicated in Section 3.4.

Since many practical examples of manifolds are closed or homogeneous (Euclidean space, linear groups, spheres, projective planes), our results will be widely applicable. However, the case of domains in Euclidean space with regular (nonempty) boundaries, which may be important in geology, escapes the present analysis. The statistician may notice that formally the treatment of the estimator is very similar to that of series-type estimators on a Euclidean space. For the geometer it might be interesting to observe that this formal similarity with the classical approach is based on a far from trivial generalization, due to Hörmander (1968), of the asymptotic behaviour of the spectral function and an interesting consequence for generalized zeta functions.

The statistical problem and the main results are formulated in Section 2, the geometrical facts are given in Section 3 and the proof of the main results is given in Section 4. In Section 5, an alternative nonparametric estimator is proposed, which makes the idea of "cooling down" the "hot" atomic density corresponding to a sample of observations by a conduction process explicit. It is (also) of the kernel type. For the reader unfamiliar with differential geometry, in the Appendix a survey is given of the concepts used in this paper. One may also consult Giné (1975a) for most of the concepts, as well as for explicit computations for the circle, the two-dimensional sphere and the projective plane.

2. Main results. Let $\mathbf{M}$ be an m-dimensional Riemannian manifold. For each $N \in \mathbb{N}$, let $Y_1, \dots, Y_N$ be a sample of i.i.d. random variables with values in $\mathbf{M}$. Let dvol denote the volume element of $\mathbf{M}$ associated with the Riemannian structure. It will be assumed throughout that the unknown probability distribution P of the Y_i has a density $f = dP/\mathrm{dvol} \in L_2(\mathbf{M}, \mathrm{dvol})$, with respect to the volume element. We propose to estimate the density f.

2.1. *Preliminaries.* In order to state our results we first introduce some notations from Fourier theory. Suppose that $\mathbf{M}$ is a closed Riemannian manifold. Then the Laplace–Beltrami operator Δ: $C^\infty(\mathbf{M}) \to C^\infty(\mathbf{M})$ is a positive essentially self-adjoint operator. See the Appendix for its definition. Its eigenvalues have finite multiplicity and may be enumerated by a function $\lambda: \mathbb{N} \to \mathbb{R}$, such that λ is positive (≥ 0) weakly increasing without upper bound and such that each eigenvalue occurs as often as its multiplicity. Given an orthonormal set of eigenfunctions $\{\phi_k\}_{k \in \mathbb{N}}$, for which $\Delta\phi_k = \lambda(k) \cdot \phi_k$, it is well known that $\{\phi_k\}$ is an orthonormal basis of $L_2(\mathbf{M})$. The eigenfunctions ϕ_k are smooth functions, i.e., $\phi_k \in C^\infty(\mathbf{M})$. These functions may be chosen to be real valued.

Associated with this decomposition, one has the *spectral function*

$$e(x,y,T) = \sum_{\lambda(k)<T} \phi_k(x) \cdot \overline{\phi_k(y)}\,. \tag{2.1}$$

This function does not depend on the particular choice of the eigenfunctions and we have the property

$$\int_{\mathbf{M}} e(x,y,T)g(y)\mathrm{dvol}(y) = \sum_{\lambda(k)<T} \alpha_k\phi_k(x), \quad \text{if } g(x) = \sum_k \alpha_k\phi_k(x).$$

Thus the integral operator with kernel $e(x,y,T)$ is the projection operator E_T: $L_2(\mathbf{M}) \to L_2(\mathbf{M})$ whose domain is the sum of the eigenspaces associated with the eigenvalues less than T; see Section 3.3 for explicit examples.

2.2. *Density estimation.* Consider the density f. Since $f \in L_2(\mathbf{M})$, by assumption it follows that $f = \lim_{T\to\infty} E_T f$, with convergence and equality in the L_2 sense. As $E_T f(x) = \int_{\mathbf{M}} e(x,y,T)f(y)\mathrm{dvol}(y)$, one sees that $E_T f$ is the expectation with respect to $P = f \cdot \mathrm{dvol}$ in the y coordinate of the random function $e(\cdot, y, T)$.

Given $T > 0$ and the observations $y_1, \dots, y_N$, we may estimate f by the empirical density

$$f_T^*(x) = \frac{1}{N} \cdot \sum_{j=1}^{N} e(x, y_j, T), \tag{2.2}$$

whose expectation is the projected density

$$f_T = E_T f.$$

Note that the estimator f_T^* is not necessarily a nonnegative function. Our aim is to choose T so that the distance between f and f_T^* is minimal in the L_2 sense [respectively, in the supremum sense (L_∞ sense)]. We obtain the following theorems, valid in case $\mathbf{M}$ is a closed (or homogeneous) Riemannian manifold.

THEOREM 2.1 (L_2 Estimate). *Suppose f is s times differentiable with square integrable derivatives* (*see Section* 4.2 *for the precise hypothesis*). *Let $T_0 > 0$. There are constants A* (*depending on* $\mathbf{M}$) *and B* (*depending on the* *density f) so that for $T \geq T_0$,*

$$\mathbf{E}\Big(\| f - f_T^* \|_{L_2}^2\Big) \leq A \cdot \frac{T^{m/2}}{N} + B \cdot T^{-s}.$$

In particular, for a suitable choice of T, one may obtain as dependence on N,

$$\mathbf{E}\Big(\| f - f_T^* \|_{L_2}^2\Big) \leq O\big(N^{m/(2s+m)} \cdot N^{-1}\big).$$

THEOREM 2.2 (L_∞ Estimate). *Suppose f is s times differentiable with square integrable derivatives and $s > m/2$. Let $T_0 > 0$. There are constants A' (depending on $\mathbf{M}$) and B' (depending on $\mathbf{M}$ and f) so that for $T \geq T_0$,*

$$\Big\{\mathbf{E}\Big(\| f - f_T^* \|_{L_\infty}^2\Big)\Big\}^{1/2} \leq A' \cdot \frac{T^{m/2}}{N^{1/2}} + B' \cdot T^{(m/2-s)/2}.$$

In particular, for a suitable choice of T, one may obtain as dependence on N,

$$\mathbf{E}\Big(\| f - f_T^* \|_{L_\infty}^2\Big) \leq O\big(N^{2m/(2s+m)} \cdot N^{-1}\big).$$

See Section 4 for more details about the constants A, B, A' and B' and the optimal choices for T. Asymptotically, for growing T_0, the dependence on A, B, A' and B' on $\mathbf{M}$ is only through the dimension m of $\mathbf{M}$. The above raw form is sufficient to deduce how to change N and T in order to reduce the error by a certain factor. Our inspiration for this type of theorem was Theorem 3 of Devroye and Györfi (1985, page 308), where the L_1 error is estimated in the situation where $\mathbf{M}$ is the circle S^1.

3. The spectral function and the zeta function. In order to prove Theorems 2.1 and 2.2 it is necessary to consider the behaviour of the spectral function, and the related zeta function.

3.1. *The general facts.* Consider the spectral function $e(x, y, T)$ of the Laplace–Beltrami operator. It has the following properties:

1. $\overline{e(x, y, T)} = e(y, x, T)$ (symmetry).
2. For each x, $e(x, \cdot\,, T)$ is an L_2 function.
3. $\int_{\mathbf{M}} e(x, y, T) e(y, z, T)\, dy = e(x, z, T)$ (idempotency).

From these properties follows as an example:

4. $|f_T(x)| = |\int_{\mathbf{M}} e(x, y, T) f(y) \mathrm{dvol}(y)| \leq e(x, x, T)^{1/2} \| f \|_{L_2}$.

We will make use of the following known facts about the asymptotic behaviour of the spectral function. Recall that $\mathbf{M}$ is supposed to be closed. Let

v_m denote the quantity

$$v_m = \frac{1}{(2\sqrt{\pi})^m \Gamma(m/2+1)}$$

and let $C(x,T)$ be defined by

$$e(x,x,T) = C(x,T) \cdot v_m \cdot T^{m/2}. \tag{3.1}$$

EXAMPLE. For the circle S^1, $m = 1$ and $v_1 = 1/\pi$. Moreover $C(x,T) = (k + \frac{1}{2})/T^{1/2}$, where k denotes the largest integer less than $T^{1/2}$ (see Section 3.3).

FACTS. It is one of the beautiful facts that the geometry of $\mathbf{M}$ shows up only in the lower order behaviour of $C(x,T)$. In particular $\lim_{T\to\infty} C(x,T) = 1$ [Minakshisundaram and Pleijel (1949), page 243] and even sharper $\sup_{x\in\mathbf{M}}|C(x,T) - 1| = O(T^{-1/2})$ [see Hörmander (1968), page 194, or Duistermaat and Guillemin (1975), (2.25)].

The *zeta function* of the Laplace–Beltrami operator is the function

$$Z(x,s) = \sum_{\lambda(k)>0} |\phi_k(x)|^2 \lambda(k)^{-s}.$$

It is known that the sum converges absolutely for complex numbers s with $\mathrm{Re}(s) > m/2$ and that it has a meromorphic extension with, among other poles, a simple pole in $m/2$ [see Minakshisundaram and Pleijel (1949) or Duistermaat and Guillemin (1975)].

In order to prove Theorem 2.2 we shall need to consider the rate of convergence of the zeta function. Consider the function

$$Y(x,s,T) = \sum_{\lambda(k)\geq T} |\phi_k(x)|^2 \lambda(k)^{-s}$$

and define $D(x,s,T)$ by

$$Y(x,s,T) = D(x,s,T) \cdot v_m \cdot \frac{m/2}{s - m/2} \cdot T^{m/2-s}. \tag{3.2}$$

EXAMPLE. For the circle S^1, a spectral decomposition is given in Section 3.3 and we have

$$Z(x,s) = \frac{1}{\pi}\sum_{k=1}^{\infty} k^{-2s} = \frac{1}{\pi}\zeta(2s), \qquad Y(x,s,T) = \frac{1}{\pi}\sum_{k\geq\sqrt{T}} k^{-2s}$$

and

$$D(x,s,T) = (2s-1)T^{s-1/2} \sum_{k\geq\sqrt{T}} k^{-2s}.$$

THEOREM 3.1.

$$\lim_{T\to\infty} \sup_{x\in\mathbf{M}} \sup_{s>m/2} \frac{|D(x,s,T)-1|}{s} = 0.$$

The proof is inspired by the method of Mandelbrojt (1969), Theorem I.2.1, to prove convergence of Dirichlet series.

PROOF. Let $\varepsilon > 0$. Choose T_ε such that for $T > T_\varepsilon$ and $x \in \mathbf{M}$, one has $|C(x,T) - 1| < \varepsilon$. Let $Q > T > T_\varepsilon$ and let $l(i)$, $i = 1,\dots,n$, be a strictly increasing enumeration of the set $\{Q,T\} \cup \{\lambda(k); Q > \lambda(k) \geq T\}$. Let $A(i) = \sum_{\lambda(k)<l(i)} |e_k(x)|^2$ and $t(i)$, $i = 1,\dots,n-1$, be such that $l(i) < t(i) < l(i+1)$ and that

$$\frac{l(i)^{-s} - l(i+1)^{-s}}{l(i)^{m/2-s} - l(i+1)^{m/2-s}} = \frac{s}{s-m/2} \cdot t(i)^{-m/2}.$$

Then

$$\begin{aligned}
&\sum_{Q>\lambda(k)\geq T} |e_k(x)|^2 \lambda(k)^{-s} \\
&\quad= \sum_{i=1}^{n-1} (A(i+1) - A(i)) l(i)^{-s} \\
&\quad= \sum_{i=1}^{n-1} A(i+1)\big(l(i)^{-s} - l(i+1)^{-s}\big) - A(1)l(1)^{-s} + A(n)l(n)^{-s} \\
&\quad= \sum_{i=1}^{n-1} e(x,x,t(i))\big(l(i)^{-s} - l(i+1)^{-s}\big) - e(x,x,T)T^{-s} \\
&\qquad+ e(x,x,Q)Q^{-s} \\
&\quad= \sum_{i=1}^{n-1} C(x,t(i))v_m t(i)^{m/2}\big(l(i)^{-s} - l(i+1)^{-s}\big) \\
&\qquad- C(x,T)v_m T^{m/2-s} + C(x,Q)v_m Q^{m/2-s} \\
&\quad= \sum_{i=1}^{n-1} C(x,t(i))v_m \frac{s}{s-m/2}\big(l(i)^{m/2-s} - l(i+1)^{m/2-s}\big) \\
&\qquad- C(x,T)v_m T^{m/2-s} + C(x,Q)v_m Q^{m/2-s}.
\end{aligned}$$

Notice that if $C(\cdot,\cdot)$ were identical to 1, this expression exactly equals $v_m \cdot (m/2)/(s - m/2) \cdot (T^{m/2-s} - Q^{m/2-s})$. Thus

$$\left| \sum_{Q > \lambda(k) \geq T} |e_k(x)|^2 \lambda(k)^{-s} - v_m \cdot \frac{m/2}{s - m/2} \cdot T^{m/2-s} \right|$$

$$\leq \varepsilon \cdot v_m \left\{ \frac{s}{s - m/2} + 1 \right\} \cdot T^{m/2-s} + \varepsilon \cdot v_m \left\{ -\frac{s}{s - m/2} + 1 \right\} \cdot Q^{m/2-s}$$

$$\leq \varepsilon \cdot \frac{2s - m/2}{m/2} \cdot v_m \cdot \frac{m/2}{s - m/2} \cdot T^{m/2-s}.$$

Now taking the limit for $Q \to \infty$, $|D(x, s, T) - 1| \leq \varepsilon \cdot \{(2s - m/2)/(m/2)\}$, from which the theorem immediately follows. □

REMARK. Using Hörmander (1968), the above results can easily be modified to the situation of any positive elliptic symmetric partial differential operator instead of the Laplace–Beltrami operator.

3.2. *Specialization to homogeneous closed spaces.* In this section we suppose that the group of isometries of $\mathbf{M}$ acts transitively on $\mathbf{M}$, i.e., $\mathbf{M}$ is homogeneous.

LEMMA 3.1. *If $\mathbf{M}$ is homogeneous, the function $e(x, x, T)$ is independent of x in $\mathbf{M}$. Moreover $C(x, T)$ and $D(x, s, T)$ are independent of x.*

PROOF. The argument is entirely classical. Let h be an isometry of $\mathbf{M}$. Then h acts on $L_2(\mathbf{M})$ via the action on $\mathbf{M}$ by $f \mapsto f \circ h$. Because h leaves the measure dvol invariant, it gives rise to an isometry of L_2. Moreover h leaves the Laplace–Beltrami operator invariant and therefore transforms a spectral decomposition into another one. But the spectral function is independent of the particular spectral decomposition so that $e(h(x), h(y), T) = e(x, y, T)$. Because the isometry group of $\mathbf{M}$ acts transitively on $\mathbf{M}$, the function $e(x, x, T)$ is independent of x. □

LEMMA 3.2. *If $\mathbf{M}$ is homogeneous and closed, then $\sum_{\lambda(k)=T} |\phi_k(x)|^2$ equals the multiplicity of T divided by the volume of $\mathbf{M}$.*

See Giné (1975b), which contains this result as an immediate consequence of his beautiful addition formula for the zonal eigenfunctions.

PROOF. Let λ be the smallest eigenvalue of the Laplace–Beltrami operator (strictly) greater than T. Then the sum expression equals $e(x, x, \lambda) - e(x, x, T)$. Thus it is independent of x. The result follows by integration. □

CONSEQUENCE. If $\mathbf{M}$ is homogeneous and closed, then:

1. $e(x, x, T) = e(T) = 1/(\mathrm{vol}(\mathbf{M})) \cdot \#\{k; \lambda(k) < T\}$.
2. $Y(x, s, T) = Y(s, T) = 1/(\mathrm{vol}(\mathbf{M})) \cdot \sum_{\lambda_k \geq T} 1/(\lambda(k)^s)$.

3.3. *Examples of spectral decompositions.* The techniques in finding the spectral decomposition for the most common homogeneous spaces have a long history. A very instructive exposition of some of them is given by Vilenkin (1968). A concise explanation of the spectral decompositions for the circle, the two-dimensional sphere and the projective plane can be found in Giné (1975a), Section 6.

In the case of the circle $\mathbf{M} = S^1$, we have the following: It can be parametrized by $\psi \mapsto (\cos\psi, \sin\psi)$, where $0 \leq \psi \leq 2\pi$. Its volume (length) is 2π. The eigenvalues are $T_j = (j-1)^2$. The multiplicity of T_1 is 1 and an orthonormal eigenfunction is $f(\psi) = 1/\sqrt{(2\pi)}$. The multiplicity of T_j $(j > 1)$ is 2 and an ortho-normal basis of the eigenspace is $f_1(\psi) = 1/\sqrt{\pi}\cos((j-1)\psi)$ and $f_2(\psi) = 1/\sqrt{\pi}\sin((j-1)\psi)$. The spectral function has the following expression, where k denotes the smallest integer less than $T^{1/2}$:

$$e(\psi_1, \psi_2, T) = \frac{1}{2\pi}\frac{\sin\left((k+\frac{1}{2})(\psi_1 - \psi_2)\right)}{\sin\left(\frac{1}{2}(\psi_1 - \psi_2)\right)}, \quad \text{if } \psi_1 \neq \psi_2,$$

$$e(\psi, \psi, T) = \frac{2k+1}{2\pi}.$$

In the case of the two-dimensional sphere $\mathbf{M} = S^2$, we have the following. It can be parametrized by $(\theta, \psi) \mapsto (\sin(\theta)\cos(\psi), \sin(\theta)\sin(\psi), \cos(\theta))$ where $0 \leq \theta \leq \pi$ and $0 \leq \psi \leq 2\pi$. Its volume (area) is 4π. The eigenvalues are $T_j = (j-1)j$. The multiplicity of T_j is $2j-1$ and an orthonormal basis of the eigenspace is

$$f_m(\theta, \psi) = 1/\sqrt{4\pi}\{(2j-1)(j-1-m)!/(j-1+m)!\}^{1/2}$$
$$\times \exp(im\psi)P_{j-1}^{(m)}(\cos(\theta)),$$

where $m = -j+1, \dots, j-1$. Here $P_{j-1}^{(m)}$ denotes the associated Legendre function [see Gradshteyn and Ryzhik (1965), formulae 8.752.1 and 8.910.2, for its definition].

In the case of the projective plane, considered as a quotient of S^2, the volume is 2π and the spectral decomposition is given by the eigenvalues T_{2j} of the two-dimensional sphere with the same multiplicity and the $\sqrt{2}$ multiples of their eigenfunctions.

In the case of the special orthogonal group of order 3, i.e., the rigid motions of the three-dimensional Euclidean space fixing the origin, one may refer to Vilenkin (1968) or Gel'fand, Minlos and Shapiro (1963). One has the following: The manifold $\mathbf{M} = \mathrm{SO}(3)$ can be provided with the parametrization by Euler angles [see Vilenkin (1968), page 106]. Its volume is $16\sqrt{2}\pi^2$. The eigenvalues are $T_j = (j-1)j/2$. The multiplicity of T_j is $(2j-1)^2$ and an orthonormal basis of the eigenspace is given by

$$f_{mn}(\phi,\theta,\psi) = \frac{\sqrt{2j-1}}{\sqrt{\mathrm{vol}(\mathrm{SO}(3))}} \cdot \exp(-i(m\phi + n\psi)) P_{mn}^{j-1}(\cos\theta),$$

where $m, n = -j+1, \dots, j-1$. Here P_{mn}^{j-1} is intimately related to the Jacobi polynomial $P_{j-1-m}^{(m-n, m+n)}$ [see Vilenkin (1968), page 125] from

$$P_{mn}^{l}(z) = 2^{-m} i^{m-n} \sqrt{\frac{(l-m)!(l+m)!}{(l-n)!(l+n)!}}$$
$$\times (1-z)^{(m-n)/2}(1+z)^{(m+n)/2} P_{l-m}^{(m-n, m+n)}(z).$$

[For the definition of Jacobi polynomials, see Gradshteyn and Ryzhik (1965), formula 8.960.]

3.4. *A remark on nonclosed manifolds.* In the general case, the Laplace–Beltrami operator Δ may be considered as an operator on $L_2(\mathbf{M})$ with the collection of smooth functions with compact support as domain and range. As such, it is a positive symmetric operator and by the theorem of Friedrichs it admits a positive self-adjoint extension. If $\mathbf{M}$ is geodesically complete (e.g., closed or homogeneous), the operator Δ is essentially self-adjoint, so that one has a natural self-adjoint extension [see Chernoff (1973)]. From the spectral theorem we have a measured space $\mathbf{S}$ together with a measurable real valued function λ: $\mathbf{S} \to \mathbb{R}$ and an isometry $\mathscr{F}$: $L_2(\mathbf{M}) \to L_2(\mathbf{S})$ so that $\mathscr{F}(\Delta g) = \lambda \cdot \mathscr{F}g$ (pointwise multiplication).

This equality holds at least for C^∞ functions with compact support in $\mathbf{M}$. The function λ is nonnegative.

Associated with the isometry $\mathscr{F}$, one has the spectral resolution $\{E_T\}$ of projection operators E_T: $L_2(\mathbf{M}) \to L_2(\mathbf{M})$, defined by $(\mathscr{F}E_T g)(v) = \mathscr{F}g(v)$ if $\lambda(v) < T$ and $(\mathscr{F}E_T g)(v) = 0$, otherwise.

Then E_T is an integral operator [see Agmon and Kannai (1967) and Hörmander (1968)]. Its kernel will be denoted by $e(x, y, T)$.

In the case where $\mathbf{M}$ is closed, $\mathbf{S}$ may be chosen to be the natural numbers $\mathbb{N}$, together with the count measure and λ as in Section 2.1. Then the points of $\mathbf{S}$ correspond to an orthonormal basis $\{\phi_k\}_k$ and we obtain formula (2.1).

For the Euclidean space $\mathbb{R}^n$, the spectral theorem gives rise to a space $\mathbf{S} = \mathbb{R}^n$, with Lebesgue measure and real valued function $\lambda(\xi) = \|\xi\|^2$. The Fourier transform is the classical one:

$$\mathscr{F}(f)(\xi) = \int (2\pi)^{-m/2} e^{-i\xi\cdot y} f(y)\, dy.$$

The associated spectral function is

$$e(x,y,T) = \frac{J_{m/2}(\|x-y\| \cdot \sqrt{T})}{(\|x-y\| \cdot \sqrt{T})^{m/2}} \cdot \frac{T^{m/2}}{(2\pi)^{m/2}}.$$

Here $J_{m/2}$ denotes the Bessel function of order $m/2$ [Gradshteyn and Ryzhik (1965), Section 8.4]. For $\|x - y\| \cdot \sqrt{T}$ tending to 0, this function converges to $e(x, x, T) = v_m \cdot T^{m/2}$. The corresponding estimate (2.2) in the case $\mathbf{M} = \mathbb{R}^1$ is the so-called Fourier integral estimate of Konakov (1973) and Davis (1975, 1977).

4. The Fourier expansion estimate. Suppose we have a sample $Y_1, \dots, Y_N$ of i.i.d. random variables with values in the Riemannian manifold $\mathbf{M}$. Let $T > 0$. Then we will estimate $f_T = E_T f$ by f_T^*, defined by formula (2.2): $f_T^*(x) = 1/N \cdot \sum_{j=1}^{N} e(x, y_j, T)$.

Our aim is to choose T so that the distance between f and f_T^* is minimal in the L_2 sense [respectively, in the supremum sense (L_∞ sense)]. Our analysis will be based on the decomposition $f - f_T^* = (f - f_T) + (f_T - f_T^*)$. We will refer to these terms as the projection error and the sampling error. Remark that the two error terms are orthogonal in the L_2 sense, because the sampling error lies in the sum of the eigenspaces associated with eigenvalues less than T and the projection error lies in the sum of the eigenspaces with eigenvalues not less than T.

Recall that $\mathbf{M}$ is supposed to be closed. Thus $C_T = \sup_{x \in \mathbf{M}} C(x, T)$ is finite and we have the equality [cf. formula (3.1)]

$$\sup_{x \in \mathbf{M}} e(x, x, T) = C_T \cdot v_m \cdot T^{m/2}. \tag{4.1}$$

Furthermore for $s > m/2$, the number $D_{s,T} = \sup_{x \in \mathbf{M}} D(x, s, T)$ is finite and we have the equality [cf. formula (3.2)]

$$\sup_{x \in \mathbf{M}} Y(x, s, T) = D_{s,T} \cdot v_m \cdot \frac{m/2}{s - m/2} \cdot T^{m/2 - s}. \tag{4.2}$$

4.1. *The (L_2) error due to sampling.* The variance of $f_T^*(x)$, whose expectation value is $f_T(x)$ (w.r.t. f dvol), is

$$\begin{aligned}\mathbf{E}\left(|f_T^*(x) - f_T(x)|^2\right) &= \frac{1}{N} \cdot \left(\mathbf{E}\left(|e(x,y,T)|^2\right) - |\mathbf{E}(e(x,y,T))|^2\right)\\ &= \frac{1}{N} \cdot \left(\mathbf{E}\left(|e(x,y,T)|^2\right) - |f_T(x)|^2\right).\end{aligned}$$

Thus, by exchange of the order of integration and using the idempotency property (Section 3.1, property 3) we obtain

$$\begin{aligned}\mathbf{E}\left(\| f_T - f_T^* \|_{L_2}^2\right) &= \frac{1}{N} \cdot \left(\mathbf{E}(e(y,y,T)) - \| f_T \|_{L_2}^2\right) \le \frac{1}{N} \cdot \left(\sup_y e(y,y,T)\right)\\ &\le \frac{1}{N} \cdot C_T \cdot v_m \cdot T^{m/2}.\end{aligned}$$

This leads to the following conclusion.

CONCLUSION.

$$\mathbf{E}\Big(\| f_T - f_T^* \|_{L_2}^2\Big) \le \frac{1}{N} \cdot C_T \cdot v_m \cdot T^{m/2}. \tag{4.3}$$

Recall from Section 3.1 that $\lim_{T \to \infty} C_T = 1$. Note also that for the uniform distribution [i.e., $f = 1/\mathrm{vol}(\mathbf{M})$] we have exactly the qualitative behaviour indicated by the inequality (4.3).

4.2. *The (L_2) error due to projection.* We will study $\| f - f_T \|_{L_2}$, under the assumption that f is s times differentiable, with L_2 derivatives. Because $\mathbf{M}$ is closed, it follows that $\Delta = -\operatorname{div}\operatorname{grad} = d^* d$ when applied to twice differentiable functions with square integrable derivatives (see the Appendix). Let $f^{(k)}$ be defined recursively, for $k \le s$, as

$$f^{(0)} = f, \qquad f^{(2j+1)} = df^{(2j)}, \qquad f^{(2j+2)} = d^* f^{(2j+1)} = \Delta f^{(2j)}.$$

Suppose moreover that $f^{(k)} \in L_2$ for $k = 1, \dots, s$. Then:

LEMMA 4.1. *If* $\mathbf{M}$ *is closed and* $f(x) = \sum_k \alpha_k \phi_k(x)$ *as in Section* 2.1, *then* $\int_{\mathbf{M}} |f^{(s)}|^2 \,\mathrm{dvol} = \sum \lambda(k)^s \cdot |\alpha_k|^2$.

PROOF. If s is even, then Δ may be applied $s/2$ times to f, giving an L_2 function. If s is odd, it is necessary to introduce the full exterior algebra of $\mathbf{M}$, together with the essentially self-adjoint operator $d + d^*$ and a corresponding spectral decomposition [see Chernoff (1973)]. Notice that the square of $d + d^*$ is exactly the Laplace–Beltrami operator (on the space of smooth functions).
As a consequence

$$\mathbf{E}\Big(\| f - f_T \|_{L_2}^2\Big) = \sum_{\lambda(k) \ge T} |\alpha_k|^2 \le \sum_{\lambda(k) \ge T} \lambda(k)^s |\alpha_k|^2 T^{-s} \le \| f^{(s)} \|_{L_2}^2 \cdot T^{-s}. \qquad \square$$

CONCLUSION.

$$\mathbf{E}\Big(\| f - f_T \|_{L_2}^2\Big) \le \| f^{(s)} \|_{L_2}^2 \cdot T^{-s}. \tag{4.4}$$

REMARK. In general, a function $f \in L_2(\mathbf{M})$ belongs to the Sobolev space $H_s(\mathbf{M})$ if and only if $\sum \lambda(k)^s \cdot |\alpha_k|^2 < \infty$. One may replace throughout this section the integrability condition on $f^{(s)}$ with the condition $f \in H_s(\mathbf{M})$ and the square integral of $f^{(s)}$ with $\| f \|_{H_s}^2 = \sum \lambda(k)^s \cdot |\alpha_k|^2$. Recall Sobolev's lemma, stating that for a positive integer k and $s > k + m/2$, a function $f \in H_s(\mathbf{M})$ is represented by a k times continuously differentiable function on $\mathbf{M}$ [see Aubin (1982), Theorem 2.21]. As a consequence, the alternative to the differentiability condition on f in Theorem 2.2, namely the Sobolev space condition $f \in H_s$ with $s > m/2$, still implies the continuity of the function f.

4.3. *Proof of Theorem* 2.1. Let $T_0 > 0$ be given. Let $A = \sup_{T \ge T_0} C_T \cdot v_m$ and $B = \| f^{(s)} \|_{L_2}^2$. Using the fact that the sampling error and the projection error are orthogonal in the L_2 sense, we deduce from inequalities (4.3) and (4.4) the error bound for $T \geq T_0$:

$$\mathbf{E}\Big(\| f - f_T^* \|_{L_2}^2\Big) \leq A \cdot \frac{T^{m/2}}{N} + B \cdot T^{-s}.$$

This expression attains its minimal value for $T = \max(T_0, \{(2sB/mA) \cdot N\}^{1/(s+m/2)})$ and (if $T > T_0$) the minimal value is

$$\left(1 + \frac{m}{2s}\right) \cdot \left(\frac{2sB}{mA}\right)^{m/(2s+m)} \cdot N^{m/(2s+m)} \cdot A \cdot N^{-1}. \qquad \square$$

4.4. *Pointwise bounds.* If we try to obtain pointwise bounds we find the following: First, notice that property 4 of Section 3.1 applies to $f_T - f_T^*$, so that

$$| f_T(x) - f_T^*(x) |^2 \leq e(x, x, T) \cdot \| f_T - f_T^* \|_{L_2}^2.$$

Therefore, using formulas (4.1) and (4.3), we conclude:

CONCLUSION.

$$\text{(4.5)} \qquad \mathbf{E}\Big(\sup_x | f_T(x) - f_T^*(x) |^2\Big) \leq \frac{1}{N} \cdot \{C_T \cdot v_m \cdot T^{m/2}\}^2.$$

Second, supposing that f is s times differentiable with $s > m/2$, one may derive the following inequality: If $R > S > T$, then

$$| f_R(x) - f_S(x) | = \sum_{R > \lambda(k) \geq S} \alpha_k \phi_k(x) \leq \sum_{R > \lambda(k) \geq S} \alpha_k \lambda(k)^{s/2} \phi_k(x) \lambda(k)^{-s/2}.$$

Thus

$$\begin{aligned} | f_R(x) - f_S(x) |^2 &\leq \sum_{R > \lambda(k) \geq S} |\alpha_k|^2 \lambda(k)^s \cdot \sum_{R > \lambda(k) \geq S} |\phi_k(x)|^2 \lambda(k)^{-s} \\ &\leq \| f^{(s)} \|_{L_2}^2 \cdot Y(x, s, T). \end{aligned}$$

A consequence of Theorem 3.1 is that $Y(x, s, T)$ converges uniformly in x to 0 as $T \to \infty$. Therefore f_T converges uniformly as $T \to \infty$ and because its L_2 limit is f, its pointwise limit is f.

We may thus conclude, using (4.2), that:

CONCLUSION.

$$\text{(4.6)} \qquad | f(x) - f_T(x) |^2 \leq \| f^{(s)} \|_{L_2}^2 \cdot D_{s,T} \cdot v_m \cdot \frac{m/2}{s - m/2} \cdot T^{m/2 - s}.$$

Recall from Theorem 3.1 that $\lim_{T \to \infty} D_{s,T} = 1$.

4.5. *Proof of Theorem* 2.2. Let $T_0 > 0$. Let $A' = \sup_{T \geq T_0} C_T \cdot v_m$ and

$$B' = \| f^{(s)} \|_{L_2} \cdot \left\{ \sup_{T \geq T_0} D_{s,T} \cdot v_m \cdot \frac{m/2}{(s - m/2)} \right\}^{1/2}.$$

Then we obtain from inequalities (4.5) and (4.6) the error bound for $T \geq T_0$:

$$\left\{ \mathbf{E}\left(\sup_x | f(x) - f_T^*(x) |^2 \right) \right\}^{1/2} \leq A' \cdot \frac{T^{m/2}}{N^{1/2}} + B' \cdot T^{(m/2 - s)/2}.$$

This expression attains its minimal value for

$$T = \max\left(T_0, \left\{ \frac{B' \cdot (s - m/2)}{A' \cdot m} \cdot N^{1/2} \right\}^{2/(s + m/2)} \right)$$

and (if $T > T_0$) the minimal value is

$$\left(1 + \frac{m}{(s - m/2)}\right) \cdot \left(\frac{(s - m/2) B'}{mA'} \right)^{m/(s+m/2)} \cdot N^{m/(2s+m)} \cdot A' \cdot N^{-1/2}. \quad \square$$

4.6. *A remark on the estimation of* $\| f^{(s)} \|_{L_2}$. As usual in statistics, any a priori information on the density is precious, as for example, knowledge of the mechanism that leads to the density from which the degree of differentiability or even an upper bound of $\| f^{(s)} \|_{L_2}$ may be inferred, or from which, in the context of this paper, e.g., symmetry properties may lead to (linear) relations between the coefficients of the Fourier expansion. This does not mean that there is no statistical way to estimate $\| f^{(s)} \|_{L_2}$.

We propose the following procedure to estimate $\| f_S^{(s)} \|_{L_2} = \sum_{\lambda(k) < S} |\alpha_k|^2 \lambda(k)^s$. The S used in this formula of course may be different and would have to be considerably larger than the upper eigenvalue T used in the Theorems 2.1 and 2.2. Notice that $\int_{\mathbf{M}} e(x, x, T)\mathrm{dvol}$, which behaves as $O(T^{m/2})$, is the number of coefficients α_k with $\lambda(k) < T$.

Given i.i.d. observations $x_1, \dots, x_n$, take

$$\Psi = \sum_{\lambda(k) < S} \frac{1}{n(n-1)} \sum_{i \neq j} \phi_k(x_i) \overline{\phi_k(x_j)}\, \lambda(k)^s.$$

Then it follows from Lemma 4.1 that $\mathbf{E}(\Psi) = \| f_S^{(s)} \|_{L_2}^2$. Of course it is important to know that the variance of the estimator Ψ is somehow under control. The following assertions depend on a study of the rate of divergence of the zeta function $Z(x, s)$ for $s < 0$.

CLAIM. If f belongs to $L_\infty(\mathbf{M}) \cap H_\alpha(\mathbf{M})$ for some $\alpha > 0$, then for $2s \leq \alpha$ the variance of Ψ can be estimated by

$$\frac{1}{N^2} A_f S^{2s + m/2} + \frac{B_f}{N} \tag{4.7}$$

and for $2s > \alpha$ the variance can be estimated by

$$\frac{1}{N^2} A_f S^{2s+m/2} + \frac{B_f'}{N} S^{2s-\alpha}. \tag{4.8}$$

The uniform density exhibits the behavior (4.7) with $B_f = 0$. In the dim zone where $f \notin L_\infty$, so that in particular $\alpha \le m/2$ and the Fourier series need not be absolutely summable, the variance of Ψ can be estimated by

$$\frac{1}{N^2} A_f S^{2s+m/2} + \frac{B_f''}{N} S^{2s-\alpha+m/4}. \tag{4.9}$$

We do not claim that the estimation (4.9) is sharp. In these inequalities A_f, B_f, B_f' and B_f'' depend on f. In particular A_f can be chosen proportional to $\|f\|_{L_2}^2$, with proportionality constant arbitrarily (depending on T) close to $2 \cdot v_m \cdot (m/2)/(m/2 + 2s)$.

5. Comparison to a method related to diffusion. Recall that we suppose that $\mathbf{M}$ is a closed manifold. Let $f \in L_2(\mathbf{M})$ be as in the introduction of Section 2 with $f = \sum a_k \phi_k$. Define

$$\tilde{f}_\tau^* = \sum \left(a_k^* e^{-\lambda(k)\tau}\right) \cdot \phi_k \quad \text{where} \quad a_k^* = \frac{1}{N} \cdot \sum_{j=1}^{N} \overline{\phi_k(y_j)}\,,$$

whose expectation is

$$\tilde{f}_\tau = \sum \left(a_k e^{-\lambda(k)\tau}\right) \cdot \phi_k.$$

Then τ functions as a cooling time or diffusion time, and we search for τ, for which the expected value of $\|f - \tilde{f}_\tau^*\|_{L_2}^2$ is minimal.

The expected variance due to sampling is

$$\mathbf{E}\left(\|\tilde{f}_\tau - \tilde{f}_\tau^*\|_{L_2}^2\right) = \frac{1}{N} \cdot \left(\mathbf{E}\left(\sum |\phi_k(x)|^2 e^{-2\lambda(k)\tau}\right) - \sum |a_k|^2 e^{-2\lambda(k)\tau}\right)$$
$$\le \frac{1}{N} \cdot \sup_x \sum |\phi_k(x)|^2 e^{-2\lambda(k)\tau}.$$

Let $\theta(x, x, \tau) = \sum |\phi_k(x)|^2 e^{-\lambda(k)\tau}$. Then according to Berger, Gauduchon and Mazet (1971), page 215, or Duistermaat and Guillemin (1975), Corollary 2.2′, one has $\lim_{\tau \to 0} \theta(x, x, \tau)(4\pi\tau)^{m/2} - 1 = 0$ and the limit is uniform in x. Let $C_\tau = \sup_x \theta(x, x, \tau)(4\pi\tau)^{m/2}$. Then $\tau \mapsto C_\tau$ is continuous and $\lim_{\tau \to 0} C_\tau = 1$. As a conclusion we have:

CONCLUSION.

$$\mathbf{E}\left(\|\tilde{f}_\tau - \tilde{f}_\tau^*\|_{L_2}^2\right) \le C_\tau \cdot \frac{1}{N} \cdot (8\pi\tau)^{-m/2}.$$

We have the bound

$$\| f - \tilde{f}_\tau \|_{L_2}^2 = \sum |a_k|^2 (1 - e^{-2\lambda(k)\tau})^2 \le \sum |a_k|^2 \lambda(k)^2 \tau^2 = \| f^{(2)} \|_{L_2}^2 \cdot \tau^2.$$

This behaviour remains the same if f is more than twice differentiable.

The choice of the optimal τ, based on these inequalities will lead to a convergence behaviour (for $N \to \infty$) comparable to the case of twice differentiable densities in Section 4. Thus the diffusion method will not lead to a convergence arbitrarily close to the $N^{-1/2}$ rate. It is in fact a manifold variant of the kernel estimate [see Devroye and Györfi (1985)] using the heat kernel.

REMARK. Uniform convergence can be proved if f belongs to H_s for some $s > 2 + m/2$. The convergence behaviour (for $N \to \infty$) is comparable to the case of $H_{(2+m/2)}$ functions in Section 4.

APPENDIX

The concepts of differential geometry. Our aim is to make explicit the definition of manifolds, Riemannian structure, integration and the Laplacian. A very rigorous treatment of these concepts may be found in Helgason (1962 or 1978). A warning should be given beforehand, that for each type of manifold, additional ad hoc and cunning tricks are needed to find formulas, if any, that are manageable by classical analysis [for example, the recommended use of polar (geodesic) coordinates when dealing with spheres or projective planes or spaces].

Let $\mathbf{M}$ be a metrizable topological space. Suppose there is given a collection $\mathbf{A}$ of maps ϕ: $U \to V$, where U is an open part of $\mathbf{M}$, V is an open part of $\mathbb{R}^m$ and ϕ is a homeomorphism onto. The space $\mathbf{M}$ is a C^∞ manifold with atlas $\mathbf{A}$ if

1. The domains U of the maps ϕ in $\mathbf{A}$ cover $\mathbf{M}$.
2. If ϕ: $U \to V$ and ϕ': $U' \to V'$ belong to $\mathbf{A}$, then the map $\phi' \circ \phi^{-1}$: $\phi(U \cap U') \to \phi'(U \cap U')$ is a C^∞ map.

The elements of $\mathbf{A}$ will be called charts of the manifold.

EXAMPLES.

1. An open part of the Euclidean space $\mathbb{R}^n$ or of the space of $k \times l$-matrices to be identified with $\mathbb{R}^{k \times l}$. An atlas may consist of one chart only.
2. The $(n-1)$-dimensional sphere S^{n-1} of points of $\mathbb{R}^n$ with Euclidean distance 1 to the origin. For each i, let $U_{i,\pm} = \{x \in S^{n-1}; \pm x_i > 0\}$. Let $V = \{x \in \mathbb{R}^{n-1}; \|x\| < 1\}$ and $\phi_{i,\pm}$: $U_{i,\pm} \to V$ given by $\phi_{i,\pm}(x_1, \dots, x_n) = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$. Then $\mathbf{A} = \{\phi_{i,\pm}\}$ is an atlas with $2n$ charts.

Two atlases give rise to the same C^∞ manifold structure of $\mathbf{M}$ if the union of the two atlases again is an atlas. As a matter of fact, each C^∞ manifold supports a unique maximal atlas.

A function $f: \mathbf{M} \to \mathbb{R}$ is r times differentiable if for each chart $\phi: U \to V$ in $\mathbf{A}$, the function $f \circ \phi^{-1}: V \to \mathbb{R}$ is r times differentiable in the classical sense. A function which is r times differentiable for all r is said to be a C^∞ function and belongs to the vector space $C^\infty(\mathbf{M})$. Let $p \in \mathbf{M}$, the tangent space $T_p(\mathbf{M})$ of $\mathbf{M}$ at p be the vector space

$$\{\xi: C^\infty(\mathbf{M}) \to \mathbb{R}; \xi \text{ is } \mathbb{R} \text{ linear and } \xi(f \cdot g) = f(p) \cdot \xi(g) + g(p) \cdot \xi(f)\}.$$

It may be considered as the set of all directional derivatives at p. Given a chart $\phi: U \to V$ with $p \in U$, a basis $\{\partial_i^{\phi, p}\}_i$ of the tangent space is given by $\partial_i^{\phi, p}(f) = ((\partial f \circ \phi^{-1})/\partial x_i)(\phi(p))$, where $i = 1, \dots, m$. Given another chart $\phi': U' \to V'$ with $p \in U'$, one has the transformation rule

$$\partial_i^{\phi, p} = \sum_j \frac{\partial(\phi' \circ \phi^{-1})_j}{\partial x_i}(\phi(p)) \cdot \partial_j^{\phi', p}.$$

A Riemannian structure for $\mathbf{M}$ is given by an inner product g_p on $T_p(\mathbf{M})$ for each $p \in \mathbf{M}$ with the following smoothness condition. For every chart $\phi: U \to V$ of $\mathbf{A}$ and indices i and j, the function $v \mapsto g_{ij}(v) = g_{\phi^{-1}(v)}(\partial_i, \partial_j)$ is a C^∞ function on V. Due to the above transformation rule, a density dvol may be defined on $\mathbf{M}$. If $\phi: U \to V$ is in $\mathbf{A}$, then dvol is determined by $\phi_*(\mathrm{dvol}|_U) = \sqrt{g} \cdot dx_1 \cdots dx_m$. Here $g: V \to \mathbb{R}$ denotes the function $g(v) = \det(g_{ij}(v))_{ij}$. If $f: U \to \mathbb{R}$ is a continuous function with compact support lying in U, then $\int_{\mathbf{M}} f \, \mathrm{dvol} = \int_V f \circ \phi^{-1}(v) \cdot \sqrt{g(v)} \, dx_1 \cdots dx_m$. Integration for real or complex valued continuous functions with compact support on $\mathbf{M}$ is now defined by linearity.

Corresponding to directional derivation, one has the total or exterior derivative of C^∞ functions. Let $T_p^*(\mathbf{M})$ denote the dual vector space of $T_p(\mathbf{M})$. The dual basis with respect to $\{\partial_i^{\phi, p}\}$ will be denoted by $dx_i^{\phi, p}$, $i = 1, \dots, m$. It naturally has the inner product $\tilde{g}_{\phi, p}$ with $\tilde{g}_{\phi, p}(dx_i, dx_j) = g^{ij}(\phi(p))$, where $(g^{ij}(v))_{ij}$ denotes the inverse matrix of $(g_{ij}(v))_{ij}$. Consider the space $T^*(\mathbf{M})$ as the disjoint union of the vector spaces $T_p^*(\mathbf{M})$. It is called the cotangent bundle. A 1-form is a mapping $s: \mathbf{M} \to T^*(\mathbf{M})$ such that for each $p, s(p) \in T_p^*(\mathbf{M})$. It is r times differentiable if for each chart $\phi: U \to V$, the function defined on V, $v \mapsto s(\phi^{-1}(v))(\partial_i^{\phi, \phi^{-1}(v)})$ is r times differentiable for each i. It is C^∞ if it is r times differentiable for each r. To a C^∞ function f is associated the total derivative df defined by $df(p)(\xi) = \xi(f)$ for $p \in \mathbf{M}$ and $\xi \in T_p(\mathbf{M})$. It is a C^∞ 1-form. This is our version of the classical gradient grad. Let C_0^∞ be the vector space of C^∞ functions with compact support and Γ_0^∞ be the vector space of C^∞ sections of $T^*(\mathbf{M})$ with compact support. Then we have $d: C_0^\infty \to \Gamma_0^\infty$. Moreover C_0^∞ is an inner product space by $(f_1, f_2) = \int_{\mathbf{M}} f_1 f_2 \, \mathrm{dvol}$ and Γ_0^∞ by $(s_1, s_2) = \int_{\mathbf{M}} \tilde{g}(s_1, s_2) \mathrm{dvol}$. The adjoint operator $d^*: \Gamma_0^\infty \to C_0^\infty$ is defined. It is our version of the opposite of the divergence div. The Laplace–Beltrami operator is defined as $\Delta = -\operatorname{div} \operatorname{grad} = d^*d: C_0^\infty \to C_0^\infty$.

If $\phi: U \to V$ is a chart, then

$$\Delta f(\phi^{-1}(v)) = -\frac{1}{\sqrt{g}} \sum_k \frac{\partial}{\partial x_k}\left(\sum_i g^{ik}\sqrt{g} \frac{\partial}{\partial x_i}(f \circ \phi^{-1})\right)(v).$$

Let Δ_1 denote the corresponding partial differential operator, with domain the twice differentiable functions. Suppose that $\mathbf{M}$ is geodesically complete, which is the case if $\mathbf{M}$ is closed or homogeneous.

Then Δ is an elliptic essentially self-adjoint partial differential operator of order 2 [see Chernoff (1973)]. Moreover the self-adjoint extension of Δ is defined on a twice differentiable function f, if f, df and $\Delta_1 f$ are square integrable, and its value then is $\Delta_1 f$.

Acknowledgment. The author is very grateful to F. Ruymgaart for his interest in this work and the stimulating discussions. Thanks are also due him for indicating the place of the subject in statistical theory.

REFERENCES

AGMON, S. and KANNAI, Y. (1967). On the asymptotic behavior of spectral functions and resolvent kernels of elliptic operators. *Israel J. Math.* **5** 1–30.

AUBIN, T. (1982). Nonlinear analysis on manifolds. Monge–Ampère equations. *Grundlehren der mathematischen Wissenschaften* **252**. Springer, Berlin.

BARNDORFF-NIELSEN, O. E., COX, D. R. and REID, N. (1986). The role of differential geometry in statistical theory. *Internat. Statist. Rev.* **54** 83–96.

BERAN, R. J. (1968). Testing for uniformity on a compact homogeneous space. *J. Appl. Probab.* **5** 177–195.

BERGER, M., GAUDUCHON, P. and MAZET, E. (1971). *Le Spectre d'une Variété Riemannienne. Lecture Notes in Math.* **194**. Springer, Berlin.

CHERNOFF, P. R. (1973). Essential self-adjointness of powers of generators of hyperbolic equations. *J. Funct. Anal.* **12** 401–414.

DAVIS, K. B. (1975). Mean square error properties of density estimates. *Ann. Statist.* **3** 1025–1030.

DAVIS, K. B. (1977). Mean integrated square error properties of density estimates. *Ann. Statist.* **5** 530–535.

DEVROYE, L. and GYŐRFI, L. (1985). *Nonparametric Density Estimation. The L_1-View.* Wiley, New York.

DUISTERMAAT, J. and GUILLEMIN, V. (1975). The spectrum of positive elliptic operators and periodic bicharacteristics. *Invent. Math.* **29** 39–79.

FISHER, N. I., LEWIS, T. and EMBLETON, B. J. J. (1987). *Statistical Analysis of Spherical Data.* Cambridge Univ. Press, Cambridge, England.

GEL'FAND, I. M., MINLOS, R. A. and SHAPIRO, Z. YA. (1963). *Representations of the Rotation and Lorentz Groups and their Applications.* Pergamon, Oxford.

GINÉ, E. M. (1975a). Invariant tests for uniformity on compact Riemannian manifolds based on Sobolev norms. *Ann. Statist.* **3** 1243–1266.

GINÉ, E. M. (1975b). The addition formula for the eigenfunctions of the Laplacian. *Adv. in Math.* **18** 102–107.

GRADSHTEYN, I. S. and RYZHIK, I. M. (1965). *Tables of Integrals, Series and Products.* Academic, New York.

HELGASON, S. (1962). *Differential Geometry and Symmetric Spaces.* Academic, New York.

HELGASON, S. (1978). *Differential Geometry, Lie Groups and Symmetric Spaces.* Academic, New York.

HÖRMANDER, L. (1968). The spectral function of an elliptic operator. *Acta Math.* **121** 193–218.

JUPP, P. E. and SPURR, B. D. (1983). Sobolev tests for symmetry of directional data. *Ann. Statist.* **11** 1225–1231.

KONAKOV, V. D. (1973). Nonparametric estimation of density functions. *Theory Probab. Appl.* **17** 361–362.

MANDELBROJT, S. (1969). *Séries de Dirichlet, Principes et Méthodes.* Gauthier–Villars, Paris.

MARDIA, K. V. (1972). *Statistics of Directional Data.* Academic, New York.

MINAKSHISUNDARAM, S. and PLEIJEL, Å. (1949). Some properties of the eigenfunctions of the Laplace operator on Riemannian manifolds. *Canad. J. Math.* **1** 242–256.

VILENKIN, N. J. (1968). *Special Functions and the Theory of Group Representations*. Amer. Math. Soc., Providence, R.I. (English translation.)

WATSON, G. S. (1983). *Statistics on Spheres*. Wiley, New York.

WATSON, G. S. and WILLIAMS, E. J. (1956). On the construction of significance tests on the circle and the sphere. *Biometrika* **43** 345–352.

WELLNER, J. A. (1979). Permutation tests for directional data. *Ann. Statist.* **7** 929–943.

MATHEMATISCH INSTITUUT
KATHOLIEKE UNIVERSITEIT
TOERNOOIVELD
NIJMEGEN
THE NETHERLANDS