DENSITIES AND DYNAMICS

LYLE NOAKES

Department of Mathematics, The University of Western Australia, Nedlands WA 6009, Australia

Received September 5, 1991; Revised December 16, 1991

Let $M$ be a closed surface without boundary, possibly a sphere or a torus. A repetitive process determines a large number of points $x_n$ distributed more or less uniformly over $M$. We cannot observe these points, nor do we know what $M$ is. There is a smooth function $f: M \to \mathbb{R}$ (not given) and we measure the numbers $y_n = f(x_n)$. The problem is to say something about the geometry of $M$ from the limited information that we have. We consider also a question concerning the existence of smooth densities of continuous random variables, and then the original problem reduces to a nonstandard problem in density estimation. This is by way of a mathematical result whose proof is given in Secs. 4, 5, and 6, but very little mathematical expertise is required to apply the method in practice. The method is described in Sec. 3 where we also give some examples.

1. Introduction

We are going to show that it is possible to get topological information about the underlying manifold of a dynamical process merely from a time series or a sample of points on it. There are genericity requirements, as in the classical embedding approach [Mees, 1991], but our method does not use an embedding and proceeds directly to the topology of the manifold. This approach is attractive because of its ease of implementation, and its decisiveness when large amounts of data are available. Another attraction is the potential generality of the method: for example there are likely to be applications to probability theory, as can already be seen in Sec. 2, as well as to dynamical systems.

We do not take applications very far in the present paper: We limit our manifolds to compact oriented surfaces, and the examples use very large amounts of noise-free data. These are severe limitations. Our present purpose is to reach as wide an audience as possible, while doing the necessary mathematical work in a transparent and serviceable way. In Noakes & Mees [1991] a start is made on generalizations to dynamical systems whose attractors have fractional dimensions. Other generalizations and careful treatments of applications are planned for subsequent papers.

Not all readers of the present paper will have the time and inclination to follow the mathematical details through, but a lot of what we do is uncomplicated, especially in Secs. 1, 2, and 3. The mathematical difficulties are finally faced in Secs. 4, 5, and 6, but a careful reading of the previous sections gives a good idea of what is going on. We make some additional comments within Secs. 4, 5, and 6.

Let $M$ be a compact $C^\infty$ 2-manifold without boundary: $M$ is locally $\mathbb{R}^2$ from the point of view of calculus. Let $M$ be oriented by a nowhere-zero $C^\infty$ area form $\mu$. A time series on $M$ is a sequence $S = \{x_n: n \geq 1\} \subseteq M$ where $n$ is thought of as discrete time. Let $f: M \to \mathbb{R}$ be a $C^\infty$ map and write $T$ for the time series $\{y_n = f(x_n): n \geq 1\}$ on $\mathbb{R}$. Suppose that $M$ (and therefore $f$) and also $S$ is unknown. Can we say anything about the geometry of $M$ by observing $T$? When $f$ is constant we cannot, but most maps are not constant: an open dense set of $C^\infty$ maps is restricted Morse in the following sense.

1. Call $x \in M$ regular when there is a $C^\infty$ system $(u, v)$ of local coordinates near $x$ in which $f$ takes the form $f(u, v) = f(x) + u$. When $x$ is not regular, call $x$ a critical point of $f$.

2. Call $c \in M$ a nondegenerate critical point of local
minimum (respectively, of local maximum) when there is a \( C^\infty \) system of local coordinates near \( c \) in which \( f \) takes the form \( f(u, v) = f(c) + u^2 + v^2 \) (respectively, \( f(u, v) = f(c) - u^2 - v^2 \)).

(3) Call \( c \) a saddle when \((u, v)\) can be chosen so that \( f(u, v) = f(c) + u^2 - v^2 \).

Call \( f \) restricted Morse when it is one-to-one on its set of critical points, and these critical points are either nondegenerate local maxima, nondegenerate local minima, or saddles. It follows easily from Corollary 6.8 in Milnor [1963] that the set of restricted Morse functions is open and dense in \( C^\infty(M, \mathbb{R}) \).

Let \( f \) be restricted Morse; then \( f \) has only finitely many critical points. The Euler characteristic

\[
\chi(M) = \#(\text{local minima}) - \#(\text{saddles}) + \#(\text{local maxima})
\]

is independent of \( f \) and depends only on the geometry of \( M \). Then \( \chi(M) = 2 - 2g \), where \( g \) is a non-negative integer called the genus of \( M \). Furthermore, \( \chi(M) \) characterises \( M \) up to a diffeomorphism, namely a \( C^\infty \) homeomorphism, as a sphere with \( g \) handles: this is discussed in detail in Milnor [1963] Part 1.

Examples. \( \chi(S^2) = 2 \), \( \chi(S^1 \times S^1) = 0 \).

If \( h : N \to M \) is a diffeomorphism then \( T \) is also defined by the time series \( S^t = \{ h^{-1} x(n) : n \geq 1 \} \) on \( N \) by means of the \( C^\infty \) map \( f' = f \circ h \). So \( \chi(M) \) is the most we can hope to determine about \( M \). This problem and others are discussed in a more general setting in Mees [1991]. There, a version of the Takens embedding theorem [Takens, 1981; Noakes, 1991] is used to embed \( S \) in a relatively high dimensional Euclidean space, and it is necessary to triangulate the resulting cloud of points. Here, we offer an alternative procedure when \( S \) is more or less uniform with respect to Lebesgue measure on \( C^\infty \) charts for \( M \), which leads to simple and efficient code. We do not say anything about the case where \( S \) is uniform with respect to a measure supported by a fractal set: this is the subject of joint work in progress with Alistair Mees. The limitation that \( M \) should be two-dimensional is less essential, at least in principle.

Because \( T \) is ordered, the Takens embedding theorem underwrites the classical approach by ensuring that \( T \) determines \( \chi(M) \). We are going to throw away some of our information: suppose from now on that \( T \) is given as an unordered set, rather than as a sequence.

Then Takens’ theorem does not apply and we do not yet know whether \( T \) determines \( \chi(M) \). It is helpful to approach the problem in an indirect way, by considering first the following question in probability theory.

2. Densities and Random Variables

Let \( M \) be a compact \( C^\infty \) two-manifold without boundary. Let \( \mu \) be a never zero \( C^\infty \) two-form on \( M \). A subset \( A \) of \( M \) is said to be measurable when its intersections with all \( C^\infty \) coordinate charts of \( M \) are Lebesgue measurable in the usual sense. Let \( X \) be a random variable with values in \( M \) whose probability distribution is given by an everywhere-positive \( C^\infty \) density \( \phi : M \to \mathbb{R} \). The probability that a value \( x \) of \( X \) lies in \( A \) is

\[
P(x \in A) = \int_A \phi(x) \mu(x)
\]

Let \( f : M \to \mathbb{R} \) be \( C^\infty \) and write \( y = f(x) \), and let \( Y \) be the corresponding real-valued random variable. If \( B \) is a measurable subset of \( \mathbb{R} \) then

\[
P(y \in B) = P(x \in f^{-1}(B)) = \int_{f^{-1}(B)} \phi(x) \mu(x)
\]

We ask whether \( Y \) also has a density function, namely whether for some \( C^\infty \) \( \psi : \mathbb{R} \to \mathbb{R} \) we have

\[
\int_{f^{-1}(B)} \phi(x) \mu(x) = \int_B \psi(y) \, dy
\]

The answer may be no: if \( f \) has the same value \( \text{const.} \) at every point we would have \( f(\text{const.}) \psi(y) \, dy = 1 \). This is impossible since \( \langle \text{const.} \rangle \) has measure zero. Most \( C^\infty \) functions are not like this however, and sometimes \( \psi \) is defined, at least on the complement of a finite subset of \( f(M) \). In the following examples \( M = \mathbb{R}^2 \) which is not compact: compensate by taking \( \mu \) to be the standard area form on \( \mathbb{R}^2 \) and \( \phi \) the bivariate normal given by \( \phi(u, v) = (1/(2\pi)) \exp(-(u^2 + v^2)/2) \).

Example 1 Let \( f : \mathbb{R}^2 \to \mathbb{R}^2 \) be projection onto the first coordinate. Then

\[
\int_{f^{-1}(y)} \phi(x) \, dx = \int_{B \cap \mathbb{R}^2} \phi(x) \, dx = \int_B \int_{\mathbb{R}} \phi(y, v) \, dv \, dy
\]

and \( \psi(y) = \int_{\mathbb{R}} \phi(y, v) \, dv \) is \( C^\infty \) in \( y \).
Example 2. Let $f: \mathbb{R}^2 \to \mathbb{R}$ be given by $f(u, v) = u^2 + v^2$. For $y \geq 0$ we have

$$
\int_{\gamma(y)} \phi(x) dx = \int_B \sqrt{y} \int_{-\pi}^\pi \phi(\sqrt{y} \cos \theta, \sqrt{y} \sin \theta) d\theta dy(y)
$$

and we set

$$
\psi(y) = \frac{1}{2} \int_{-\pi}^\pi \phi(\sqrt{y} \cos \theta, \sqrt{y} \sin \theta) d\theta.
$$

For $y \leq 0$ define $\psi(y) = 0$. Then $\psi$ is $C^\infty$ except possibly at $y = 0$. Indeed

$$
\lim_{y \to 0} \psi(y) = \pi \phi(0, 0) > 0
$$

whereas

$$
\lim_{y \to -\infty} \psi(y) = 0.
$$

So $\psi$ is not even continuous at $y = 0$.

Example 3. Let $f: \mathbb{R}^2 \to \mathbb{R}$ be given by $f(u, v) = u^2 - v^2$. For $y \geq 0$ write $u = \pm \sqrt{y} \cos \theta$, $v = \sqrt{y} \sin \theta$. If $y \to 0$ throughout $B$ then

$$
\int_{\gamma(y)} \phi(x) dx = \int_0^\infty \phi(\sqrt{y} \cos \theta, \sqrt{y} \sin \theta) d\theta dy(y).
$$

For $y > 0$, write $\psi(y) = 2 \int_0^\infty \phi(\sqrt{y} \cos \theta, \sqrt{y} \sin \theta) d\theta$. The integral is undefined if $y = 0$, but for $y > 0$

$$
2\pi \psi(y) = 2 \exp(-y/2) \int_0^\infty \exp(-v^2/2(y + v^2)) dv,
$$

which is $C^\infty$ in $y$, and whose order of magnitude as $y \to 0^+$ is that of

$$
2 \int_0^1 1/\sqrt{1 + v^2} dv = 2 \ln(1 + \sqrt{1 + v^2}) \bigg|_0^1 \approx -\ln |y/4|.
$$

When $y \leq 0$ define $\psi(y) = \int_0^\infty \phi(-\sqrt{y} \sin \theta, \sqrt{y} \cosh \theta) d\theta$. Then $\psi(y)$ is $C^\infty$ in all nonzero $y$, and $2\pi \psi(y)$ increases like $-\ln |y/4|$ as $y \to 0$.

When $M$ is compact, and $u$ and $\psi$ are not given, we apply techniques from differential topology to prove the following result.

**Theorem.** Let $f: M \to \mathbb{R}$ be restricted Morse, and let $C = \{c_1, c_2, \ldots, c_k\}$ be its set of critical points. Write $d_i = f(c_i)$ for $i = 1, 2, \ldots, k$. The random variable $Y = f(X)$ has a $C^\infty$ density $\psi: \mathbb{R} \to f(C)$ with the following properties.

1. If $c_i$ is a point of local minimum (respectively local maximum) of $f$ then

$$
\lim_{y \to d_i^-} \psi(y) > \lim_{y \to d_i^-} \psi(y)
$$

(respectively $\lim_{y \to d_i^-} \psi(y) < \lim_{y \to d_i^-} \psi(y)$).

2. If $c_i$ is a saddle of $f$ then

$$
\psi(y) = O(-\ln |y - d_i|)
$$

for $y$ near $d_i$.

**Corollary.** In the situation of the theorem, let $\text{jump}_e$ be the number of points of jump discontinuity of $\psi$, and let $\log_e$ be the number of points $d$ at which

$$
\psi(y) = O(-\ln |y - d|)
$$

for $y$ near $d$. Then

$$
(2 - \text{jump}_e + \log_e)/2
$$

is a non-negative integer $g$, and $M$ is diffeomorphic to a sphere with $g$ handles.

**Remarks.** A sphere with $0$ handles is the unit sphere $S^2$ in $\mathbb{R}^3$. A sphere with $1$ handle is diffeomorphic to a torus (the boundary of a tubular hoop). Note that $\psi$ has jump discontinuities at the supremum and infimum of the values for which it is nonzero.

There are good reasons why applied mathematicians, applied statisticians, and engineers do not often consider isolated points at which densities fail to be $C^\infty$. The problem of Sec. 1 may be exceptional in this respect, because in Sec. 3 our theorem is used to construct a numerical procedure that is simple to implement and effectively determines $M$ in many cases. The proof of the theorem is given in Secs. 4, 5, and 6, and each step in the argument is elementary, although it helps to have some sympathy with ideas from differential topology. There are many steps, and it may be helpful to keep in mind the following description of the underlying ideas.

Choose $y_0 \in \mathbb{R}$ and let $\varepsilon > 0$ be small. If $\psi$ exists and is $C^\infty$ on $(y_0 - \varepsilon, y_0 + \varepsilon)$ then

$$
\zeta\psi(y) \approx \int_{\Gamma_y}(\gamma, y - c_0) \phi(x) \mu(x)
$$

where $\zeta$ is positive and very small. Alternatively, $\psi$ is the derivative of a variable integral, as in Sec. 5.
If \( y_0 \notin f(C) \) let \( N = M \). If \( y_0 = d_i \) then \( i \) is unique because \( f \) is one to one on \( C \); take \( N \) to be the complement in \( M \) of a small diffeomorphic copy of an open disc centred on \( c_i \). Then \( N \) is a \( C^\infty \) surface with boundary, and we can arrange for \( f^{-1}(y) \) to cut the boundary transversely if at all. Consider separately the approximate contributions \( \int f^{-1}(y, x, \zeta) \cap N \phi(x) \mu(x) \) and \( \int f^{-1}(x, y, z, \zeta) \cap N \phi(x) \mu(x) \) to \( \phi(y) \).

We want to show that \( \int f^{-1}(y, x, \zeta) \cap N \phi(x) \mu(x) \) depends on \( y \) and \( \zeta \) in a \( C^\infty \) fashion, and in Sec. 5 we see that it suffices to show that the domain of integration \( f^{-1}(y, x, \zeta) \cap N \) is obtained from a standard domain by a diffeomorphism \( \tau \) depending smoothly on \( y \) and \( \zeta \). When \( N = M \) we might do this by quoting a classical theorem [Ehresmann, 1947] Proposition 1, but we always have \( N \neq M \) for some choices of \( y_0 \).

The construction of \( \tau \) is given in Sec. 4. There we surround \( f^{-1}(y, x, \zeta) \cap N \) by an open subset \( W \) of \( M \), and define a \( C^\infty \) vector field \( Z \) over \( W \) which points tangentially along \( \partial N \) and whose projection to \( R \) by \( df \) has a given value \( z \in R \). Because \( f^{-1}(y_0) \cap N \) is compact it is possible to define \( \tau \) in terms of the flow of \( Z \). More precisely, \( \tau \) is defined by moving from \( f^{-1}(y_0) \cap N \) along the solution curves of a system of \( m \) ordinary differential equations for variable points in \( M \). In order that \( \tau \) should map into \( N \) it is essential that \( Z \) be tangential to \( \partial N \).

The arguments of Secs. 4 and 5 are sufficient to prove that \( \psi \) exists and is \( C^\infty \) on \( R - f(C) \). It remains to prove assertions (1) and (2). In Sec. 6 we examine the values of \( \psi(y) \) as \( y \) approaches a point \( y_0 \in f(C) \). Because \( f \) is restricted Morse we are in a situation which closely resembles Example (1), (2), or (3) in the present section. This time neither \( \mu \) nor \( \phi \) is known, but we only have to integrate over a small disc rather than the whole of \( R^2 \). Because the disc is small the values of \( \mu \) and \( \phi \) will not vary greatly, and the calculations in Sec. 6 are similar in kind to those for Examples 2, 3. The theorem then follows.

The application of our theorem to the problem of Sec. 1 is a particularly simple matter, requiring little mathematical expertise, and we describe this next.

3. Time Series Again

For the problem of Sec. 1, let \( f \) be restricted Morse. Let the \( x_{(a)} \) be representative of a random variable \( X \) described by a never-zero \( C^\infty \) density \( \phi \); the \( y_{(a)} \) are representative of \( Y = f(X) \). Our theorem says that \( Y \) has a density \( \psi \) with jump discontinuities and logarithmic singularities at points in \( f(C) \). In practice we observe a finite subset \( F \) of \( T \) of size \#F, and a first estimate \( H \) of \( \psi \) may be obtained by rescaling the vertical axis of a histogram. Specifically, choose \( z_L \) and \( z_M \) so that \( T = (z_L, z_M) \). Choose a large integer \( J \) and let \( h = (z_M - z_L)/J \). For \( 0 \leq j \leq J \) write \( z_j = z_L + jh \) and set the initial values of count(\( j \)) to be 0. For each occurrence of \( x \in F \) find \( j \) so that \( x \in (z_{j-1}, z_j) \), and replace count(\( j \)) by count(\( j \)) + 1. Unless \#F is large the histogram may have a misleadingly ragged appearance: this is a familiar problem in density estimation, and there are smoothing techniques described in Silverman [1986] Chapter 3 designed to minimise mean integrated square error of estimates. Unfortunately, mean integrated square error is less important in the present context than determining numbers of points of jump discontinuity and logarithmic singularity. Classical smoothing can easily blur the features that concern us most.

The numbers of points of jump discontinuity and logarithmic singularity of \( \psi \) determine the numbers and types of critical points of the restricted Morse function \( f \); according to our theorem. This determines \( M \) up to diffeomorphism as a sphere with \( g \) handles, as in our corollary.

Example 1. Consider the difference equation on the two-dimensional torus \( S^1 \times S^1 = M \) given by

\[
\theta^{(a+1)} = \theta^{(a)} + \alpha \quad \theta^{(a+1)} = \theta^{(a)} + \theta^{(a)}
\]

where \( \alpha \) is irrational. Here \( \theta^{(a)} \) corresponds to \( x_{\theta^{(a)}} = (\cos \theta_{\theta^{(a)}} \sin \theta_{\theta^{(a)}} \sin \theta_{\theta^{(a)}} \in M \). For \( x \in M \) take \( f(x) = \theta \) be the square of the distance from \( x \) to the point \( (1, 0, 0, 3, 7, 0) \). Then \( f \) is restricted Morse and its critical values can be calculated in advance, before generating a sample \( F \). There is a single point of local (and absolute) minimum, a single point of local (and absolute) maximum, and two saddle points, whose values are 7.93, 29.93, 22.73 and 15.13 respectively. According to the formula in Sec. 1, this information is enough to tell us that \( M \) is diffeomorphic to \( S^1 \times S^1 \). So we attempt to obtain the information by comparing a scaled histogram \( H \) with the statement of our theorem.

By Furstenberg [1981] Theorem 3.12, any orbit of the difference equation is uniformly distributed over \( M \). Take \( \alpha = \sqrt{2} \), \( x_{(i)} = (1, 0, 1, 0) \), and \( F = \{x_{(i)} : 1 \leq n \leq 100,000 \} \). We plot \( f(x_{(i)}) : 1 \leq n \leq 200 \) against \( n \) as Fig. 1a.

It is difficult to say much about the dynamics just by looking at Fig. 1a. Now we subdivide \( [z_L, z_M] = [7.5, \ldots] \).
30.5] into 100 subintervals. The graph of $H$ is shown in Fig. 1b.

From Fig. 1b, there are two apparent jump discontinuities in $\psi$ at $z = 7.93$ and $z = 29.93$, corresponding by our theorem to the points of local maximum and minimum of $f$. There are two apparent logarithmic singularities of $\psi$ at $z = 22.73$ and $z = 15.13$, corresponding to the saddles of $f$. There seems little doubt that we would correctly conclude from our corollary that $M$ is diffeomorphic to $S^1 \times S^1$.

Example 2. Let $M$ be the torus $S^1 \times S^1 \subset \mathbb{R}^4$ embedded in $\mathbb{R}^3$ by means of

$$(x_1, x_2, x_3, x_4) \mapsto y = \left( x_1 \left( 1 + \frac{x_3}{5} \right), \frac{x_4}{5}, x_2 \left( 1 + \frac{x_3}{5} \right) \right).$$

The embedded torus resembles an inner tube, which looks like an annulus when viewed from the side (see Fig. 2a).

Let $f$ be the height function, given by $f(x_1, x_2, x_3) = y_3$. Then $f$ is restricted Morse, and has a single point of local (and absolute) minimum, a single point of local (and absolute) maximum, and two saddle points, whose values are $-1.2, 1.2, -0.8$ and $0.8$ respectively. This information is enough to tell us that $M$ is diffeomorphic to $S^1 \times S^1$. We attempt to obtain the information by comparing a scaled histogram $H$ with the statement of our theorem. Generate $F$ by distributing points uniformly over the torus. From the graph of $H$ in Fig. 2b, there are two apparent jump discontinuities of $\psi$ at $z = -1.2$ and $z = 1.2$, corresponding to the local minimum and maximum of $f$. There are two apparent logarithmic singularities of $\psi$ at $z = -0.8$ and $z = 0.8$, corresponding to the saddles of $f$. Again we conclude from our corollary that $M$ is diffeomorphic to $S^1 \times S^1$.

In applications we are likely to have examples where the geometry is more complicated, but an advantage of the present example is that we can see very clearly...
what is causing the jumps and logarithmic singularities that are predicted by our theorem. Figure 2a is obtained by plotting the horizontal coordinate against the vertical coordinate \( z \), (which is the value of \( f \)) for an embedded subset of \( F \) of size 1000. It can be seen that the jumps occur at points where the topological type of the space \( f^{-1}(z) \) undergoes a change.

When \( z < -1.2 \) we have \( f^{-1}(z) \) empty, and when \( -1.2 < z < -0.8 \) we have \( f^{-1}(z) \) diffeomorphic to a circle. This change in topological type is indicated by a jump discontinuity in Fig. 2b at \( z = -1.2 \). When \( -0.8 < z < 0.8 \) we have \( f^{-1}(z) \) diffeomorphic to a disjoint union of two circles, and so there is another change of topological type at \( z = -0.8 \). This is indicated in Fig. 2a by a logarithmic singularity at \( z = -0.8 \). The logarithmic singularity at \( z = 0.8 \) and the jump discontinuity at \( z = 1.2 \) have similar interpretations.

It is not always the case that \( \text{jump}_\psi = 2 = \log_\psi \) in the case of \( S^1 \times S^1 \), as the following example shows.

**Example 3.** Take the embedded torus in \( \mathbb{R}^3 \) of Example 1, and bend both top and bottom of the hoop towards its axis of symmetry by composing with

\[
(y_1, y_2, y_3) \mapsto (y_1 + y_2, y_2, y_3)
\]

Let \( f(x) \) be the square of the distance from the image of \( x \) on the bent hoop to the point \((-1.0, 1.0, 1.0)\). This time I have not calculated the critical values of \( f \) and we attempt to determine the information we need from a histogram. Generate \( F \) as before, but increase its size to 300,000 points distributed uniformly over \( S^1 \times S^1 \subset \mathbb{R}^4 \). Taking \( z_L = 0.85, z_M = 13.2, J = 260 \), we obtain the scaled histrogram \( H \) shown in Fig. 3.

It appears that \( \psi \) has 4 points of jump discontinuity, at \( z \approx 1.0, z \approx 4.5, z \approx 8.2, \) and \( z \approx 13.0 \). There appear to be four points of logarithmic singularity of \( \psi \), at \( z \approx 1.8, z \approx 4.7, z \approx 6.5, \) and \( z \approx 7.5 \). If, as we suppose, our theorem is applicable, these features correspond to three points of relative maximum or minimum, and three saddles. So again we correctly conclude that \( M \) is diffeomorphic to a torus.

**Example 4.** Let \( M \) be the unit sphere \( S^2 \) in \( \mathbb{R}^3 \). This is the only case in which \( \psi \) does not necessarily have logarithmic singularities when \( f \) is restricted Morse and the distribution \( \phi \) of \( X \) is \( C^\infty \). Let \( f(x) \) be the height of \( x \in S^2 \subset \mathbb{R}^3 \) above the horizontal plane through the origin. Then \( f \) has a point of local (and absolute) minimum, and a point of local (and absolute) maximum, whose values are \(-1.0\) and \( 1.0 \) respectively. We wish to recover this information from a histogram. To generate \( F \) it is tempting to assign random values to polar coordinates of points in \( S^2 \): this gives a density \( \phi \) of \( X \) that is not \( C^\infty \) at the poles. We therefore generate \( F \) as follows. Choose coordinates of \( x = (x_1, x_2, x_3) \in [-1, 1]^3 \) at random and let \( r \) be the distance from \( x \) to the origin. When \( 0.1 \leq r \leq 1.0 \), replace \( x \) by \( x/r \) to obtain a point on \( S^2 \). We did this 300,000 times to obtain a sample \( F \) of size 157, 144, and set \( z_L = -1.08, z_M = 1.08, J = 100 \). The scaled histogram is shown in Fig. 4.

There appear to be jump discontinuities in \( \psi \) at \( z = -1.0 \) and \( z = 1.0 \), corresponding to the local minimum and maximum values of \( f \), and no other singularities. We correctly conclude that \( M \) is diffeomorphic to \( S^2 \). (A simple calculation shows that, in this case, \( \psi \) is constant on \((-1.0, 1.0)\).)
Example 5. In the situation of Example 4 redefine $f$ by

$$f(x_1, x_2, x_3) = ax_1 + bx_3^2$$

where $a$, $b$ are parameters. Then $f$ is restricted Morse provided $|b| < |a|/2$: Example 4 is the case $(a, b) = (1.0, 0.0)$. Take $(a, b) = (1.0, 1.0)$. Then our theorem no longer applies, but a simple calculation shows that $f$ has a local (and absolute) maximum value of 2.0, a degenerate absolute minimum value of $-0.25$, and a local minimum value of 0.0. Generate $F$ as in Example 4.

The scaled histogram $H$ shown in Fig. 5 has a sharp spike at the degenerate critical value $z = -0.25$ of $f$. On inspection, the spike appears to indicate a jump discontinuity of $\phi$ on the left and a logarithmic singularity on the right. This feature is not predicted by our theorem, but the hypotheses of the theorem are not met. (It is also possible that something like this is happening in Example 3 near $z \approx 4.6$.)

Example 6. In the situation of Example 4, change the density of $X$ on $S^2$ to another $C^\infty$ density $\phi$ by replacing the condition on $r$ by

$$0.1 \leq r \leq [1.0 + \sin^2(x_3)]/2.$$ 

We obtain a sample $F$ of size 27,730 points in this way. Let $f$ be the height function again.

The scaled histogram $H$ in Fig. 6 has jump discontinuities at $z = -1.0$ and at $z = 1.0$, and no other apparent singularities. Although the sample size $#F$ is relatively small, we correctly conclude that $M$ is diffeomorphic to $S^2$.

In Noakes & Mees [1991] we consider a number of other possible applications of the method introduced in the present paper, especially with a view to dynamical systems. That paper contains 10 illustrations which are different to those given here, and readers may wish to look at these as well. Indeed Noakes & Mees [1991] might serve as a useful introduction to the present paper.

4. Variations of Level Curves

The purpose of this section is to prove Lemma 2, which says something about the behavior of a level curve $f^{-1}(y_0)$ of $f$ as $y_0$ varies. The first case is when $y_0$ is a regular value of $f$, namely when $y_0$ is not the image of a critical point of $f$, and then Lemma 2 says that the level curve varies with $y_0$ in such a way that there are no sudden qualitative changes. In the second case, $y_0$ is the image of a critical point $c_i$, and Lemma 2 says that if there are sudden qualitative changes to the level curve then these occur near $c_i$. It requires a certain amount of mathematical technique to prove these assertions, but they are also intuitively easy to accept, and so we can avoid some pain by glossing over this section at first reading. However our assertions may be a little too easy to accept: it is not immediately clear that sudden qualitative changes are likely to occur at all. The fact that sudden qualitative changes do indeed occur near critical points is fundamental to our whole approach.

Choose $y_0 \in \mathbb{R}$. Then $f^{-1}(y_0)$ is either

(first case) a $C^\infty$ one-dimensional submanifold $P$ of $M$, or

(second case) the union of a $C^\infty$ one-dimensional submanifold $P$ of $M$ and a single point $c_i \in C$. 

Fig. 5.
In the second case, $c_i$ may or may not be an accumulation point of $f^{-1}(y_0)$, and we do the following. Let $B(w, \beta)$ be the open disc in $\mathbb{R}^2$ centred on $w \in \mathbb{R}^2$ and of radius $\beta$. Fix $\beta > 0$ and choose a $C^\infty$ embedding $g : B(0, \beta) \to M$ such that $g(0) = c_i$. Suppose further that $c_i$ is the only critical point of $f$ in the closure of $g(B(0, \beta))$, and that $f^{-1}(y_0)$ intersects $\partial g(B(0, \beta))$ transversely if at all. Then $N = M - g(B(0, \beta))$ is a compact $C^\infty$ two-manifold with boundary $\partial N = \partial g(B(0, \beta))$, and $f^{-1}(y_0)$ intersects $\partial N$ transversely if at all. In the first case let $N = M$.

In either case, $Q = f^{-1}(y_0) \cap N$ is a compact $C^\infty$ one-dimensional submanifold of $N$ with (possibly empty) boundary $\partial Q = f^{-1}(y_0) \cap \partial N$. Because $M$ is compact $\partial Q$ is a finite set $\{w_1, w_2, \ldots, w_q\}$.

**Lemma 1.** Choose $z \in \mathbb{R}$. There is an open neighbourhood $W$ of $Q$ in $M$ and a $C^\infty$ vector field $Z$ on $W$ with the properties

1. $d_v f(Z)(w) = z$ for every $w \in W$, and
2. $Z(w) \in T_{g(w)} N$, for every $w \in W \cap \partial N$.

**Proof:** Because $f^{-1}(y_0)$ intersects $\partial N$ transversely if at all, $\partial Q$ contains no critical point of $f|_{\partial N}$, and the inverse function theorem says that, for any $j = 1, 2, \ldots, q$, $w_j$ has an open neighbourhood $U_j(0)$ in $\partial N$ such that $f|_{U_j(0)}$ is a diffeomorphism onto an open interval in $\mathbb{R}$. The requirement that $d_v f(Z)(w) = z$ for every $w \in U_j(0)$ defines a $C^\infty$ vector field $Z^0$ over $\bigcup_{j=1}^q U_j(0)$ uniquely in terms of $z$. For each $j = 1, 2, \ldots, q$, choose an open neighbourhood $U_j$ of $w_j$ in $\partial N$ whose closure is contained in $U_j(0)$.

Because $Q \subset M - C$, Marsden [1974] 7.3 Theorem 3 says that we can choose an open neighbourhood $V^{(0)}$ of any $v \in Q$ in $M$ with the property

3. there is a diffeomorphism $h^{(0)} : V^{(0)} \to W_v$ where $W_v$ is open in $\mathbb{R}^2$ such that $f(w) = pr_1 \circ h^{(0)}(w)$ for every $w \in V^{(0)}$. Here $pr_1$ means projection to the first coordinate.

Choose an open neighbourhood $V_0$ of $v$ in $M$ whose closure is contained in $V^{(0)}$. Let $\{V_{v_1}, V_{v_2}, \ldots, V_{v_k}\}$ be a finite subcover of the open cover $\{V_v : v \in Q\}$ of the compact space $Q$. Write $V_{v_i}$ for $V_{v_i}$, $V^{(0)}_{v_i}$ for $V^{(0)} \cap V_{v_i}$, $h^{(0)}_{v_i}$ for $h^{(0)}|_{V^{(0)}_{v_i}}$, and $h_i$ for $h^{(0)}_{v_i}|_{V_{v_i}}$.

If $u \in (U_1^{(0)} \cup U_2^{(0)} \cup \cdots \cup U_q^{(0)}) \cap V^{(0)}_{v_i}$ then by property (iii) we have $(dh^{(0)}_{v_i})_u(Z^{(0)}(v)) = (z, Y(u))$ where $Y : (U_1^{(0)} \cup U_2^{(0)} \cup \cdots \cup U_q^{(0)}) \cap V^{(0)}_{v_i} \to \mathbb{R}^{m-1}$ is $C^\infty$. Define a $C^\infty$ vector field $Z$ over $(U_1^{(0)} \cup U_2^{(0)} \cup \cdots \cup U_q^{(0)}) \cap V^{(0)}_{v_i}$ by $(dh^{(0)}_{v_i})_u(Z^{(0)}(v)) = (z, Y(v))$ where

$Y(v) : (U_1^{(0)} \cup U_2^{(0)} \cup \cdots \cup U_q^{(0)}) \cap V^{(0)}_{v_i} \to \mathbb{R}^{m-1}$

is a $C^\infty$ extension of $Y|_{(U_1 \cup U_2 \cup \cdots \cup U_q \cap V^{(0)}_{v_i})}$. Repeat this step $t - 1$ times, with $(U_1^{(0)} \cup U_2^{(0)} \cup \cdots \cup U_q^{(0)}) \cap V^{(0)}_{v_i}$, and then $(U_1^{(0)} \cup U_2^{(0)} \cup \cdots \cup U_q^{(0)}) \cup \cdots \cup U_q^{(0)} \cup V^{(0)}_{v_i}$, and then... and finally $(U_1^{(0)} \cup U_2^{(0)} \cup \cdots \cup U_q^{(0)}) \cup \cdots \cup U_q^{(0)} \cup V^{(0)}_{v_i}$ in place of the original $U_1^{(0)} \cup U_2^{(0)} \cup \cdots \cup U_q^{(0)}$. At last we obtain a vector field $Z = Z^0$ on

$W = (U_1^{(0)} \cup U_2^{(0)} \cup \cdots \cup U_q^{(0)}) \cup \cdots \cup U_q^{(0)} \cup V^{(0)}_{v_i}$.

This proves Lemma 1. 

The proof of Lemma 1 actually yields a $C^\infty$ assignment $z \mapsto Z$ of $C^\infty$ vector fields $Z$ on $W$ to numbers $z$; we make the dependence of $Z$ on $z$ explicit by writing $Z_z$ for $Z$. We use Lemma 1 to prove

**Lemma 2.** For some $\epsilon > 0$ there is a $C^\infty$ embedding $\tau : (y_0 - \epsilon, y_0 + \epsilon) \times Q \to f^{-1}(y_0 - \epsilon, y_0 + \epsilon) \cap N$ with the property that $\tau \circ \tau(y, x) = y$ for all $(y, x) \in (y_0 - \epsilon, y_0 + \epsilon) \times Q$.

**Proof.** Let $W$ and $Z$ be as in the statement of Lemma 1, and write $\Phi_Z$ for the flow of $Z$. Because $Q$ is compact $\Phi_Z : Q \times (0, \epsilon) \to M$ is defined for sufficiently small $\epsilon > 0$. Because of property (ii) we have $\Phi_Z(Q \times (0, \epsilon)) \subset N$. By the remark following Lemma 1, there is an $\epsilon$ that will do for any choice of $z$ in $[1 - 1, 1]$.

Define a $C^\infty$ function $\tau : (y_0 - \epsilon, y_0 + \epsilon) \times Q \to N$ by

$\tau(y, x) = \Phi_{Z, \epsilon}(x, y) - y_0$.

Then $\tau(y_0, x) = x$ for all $x \in Q$. Given $(y, x) \in (y_0 - \epsilon, y_0 + \epsilon) \times Q$ write $\gamma(t) = f \circ \Phi_{Z, \epsilon}(x, t(y - y_0))$. Then $\gamma(0) = y_0$ and

$y(t) = \left( (df)_{\Phi_{Z, \epsilon}((y, y_0 - \epsilon))} Z_{y_0} \Phi_{Z, \epsilon}(x, t((y - y_0))) \right) = y - y_0$

by property (i) of $Z_{y_0}$. So $\gamma(1) = y$. But $\gamma(1) = \tau(y, x)$ and so $\tau \circ \tau(y, x) = y$ as stated.

It remains to show that $\tau$ has a $C^\infty$ inverse $\sigma$ defined over $\tau((y_0 - \epsilon, y_0 + \epsilon) \times Q)$. Given $w \in \tau((y_0 - \epsilon, y_0 + \epsilon) \times Q)$ write $\gamma(w) = y$. Let $x = \Phi_{Z, \epsilon}(w, y - y_0)$ and
define \( \sigma(w) = (y, x) \). Then

\[
\tau \circ \sigma(w) = \tau(y, x) = \Phi_{z_{x_{y}}}(x, y - y_0) = \Phi_{z_{x_{y}}}(w_1, y + y_0, y - y_0) = w.
\]

Similarly \( \sigma \circ \tau(y, x) = (y, x) \), and this completes the proof of Lemma 2. ■

Choose \( \epsilon \) as in Lemma 2. Choose \( \zeta \in (0, \epsilon/2) \) so small that \( f(C) \cap (y_0 - 2\zeta, y_0 + 2\zeta) \subseteq \{y_0\} \).

5. Variations of Probabilities

The purpose of this section is to consider what happens to the probability density \( \rho(y) \) of the random variable \( Y \) in a neighbourhood of a point \( y_0 \). When \( y_0 \) is a regular value of \( f \) we apply Lemma 2 of Sec. 4 in a straightforward way to prove that \( \rho \) depends smoothly on \( y_0 \). When \( y_0 \) is the image of a critical point \( c_j \) of \( f \) we choose a region \( N \) that is bounded away from \( c_j \). Then we only prove that the part \( \rho_{y_0} \) of \( \rho \) that is attributable to \( N \) varies smoothly. This is used in Sec. 6 where we also consider the contributions of neighbourhoods of critical points.

Carrying on from Sec. 4, define \( \rho(N) : (y_0 - \zeta, y_0 + \zeta) \times (-\zeta, \zeta) \rightarrow \mathbb{R} \) by

\[
P(y, z) = \int_{\{y, y + z\} \cap N} \phi_y \quad \text{and} \quad \rho_N(y, z) = \int_{\{y, y + z\} \cap N} \phi_y
\]

for \( z \in (0, \zeta) \) and

\[
P(y, z) = \int_{\{y, y + z\} \cap N} \phi_y \quad \text{and} \quad \rho_N(y, z) = \int_{\{y, y + z\} \cap N} \phi_y
\]

for \( z \in (-\zeta, 0) \).

For \( (y, x) \in (y_0 - \epsilon, y_0 + \epsilon) \times Q \) write \( (\tau^* \phi(y, x))_{(y, x)} = v_{(y, x)} dy \) where for each \( y \) the assignment \( x \mapsto v_{(y, x)} \) is an everywhere-positive \( C^\infty \) one-form \( v_y \) on \( \{y\} \times Q \).

**Lemma 3.** \( \rho_N \) is \( C^\infty \). If \( y \notin f(C) \) then \( P \) is \( C^\infty \) and so is \( \rho \) given by

\[
\psi(y) = \partial \partial z P(y, z) = \int_{\{y, y + z\} \cap N} v_y.
\]

**Proof.** For \( z \in (0, \zeta) \) we have

\[
P_N(y, z) = \int_{\{y, y + z\} \cap N} \phi_y = \int_{\{y, y + z\} \times Q} v_y dw = \int_{\{y, y + z\} \times Q} v_y dw.
\]

Therefore \( \partial \partial z P_N(y, z) = \int_{\{y, y + z\} \times Q} v_y dy \) and \( \partial \partial y P_N(y, z) \) are \( C^\infty \) in both \( y \) and \( z \).

In the same way \( P_N \) is \( C^\infty \) in a neighbourhood of any \( (y, z) \) where \( z \in (-\zeta, 0) \), and \( \partial \partial z \partial z P_N(y, z) \) is everywhere-zero it follows that \( P_N \) is \( C^\infty \). If \( y \notin f(C) \) set \( y_0 = y \); this does not affect the definition of \( \psi \). Then \( N = M \), and \( P = P_N \). This proves Lemma 3. ■

Define \( \psi_N : f(M) \rightarrow \mathbb{R} \) by \( \psi_N(y) = \partial \partial z \partial z P_N(y, z) \) \( z = 0 \).

By Lemma 3 \( \psi_N \) extends to a \( C^\infty \) map over an open neighbourhood of \( f(M) \). Now \( \psi \) is a real valued function defined over \( U = f(M) \cap f(C) \); note that \( U \) is open in \( \mathbb{R} \) because \( U \) is open in \( M \) and \( M \subset \mathrm{Int} f(M) \). Extend \( \psi \) to a real valued function defined on the whole of \( \mathbb{R} \) by defining \( \psi(y) = 0 \) for \( y \notin f(M) \). Since \( \mathbb{R} \) is \( f(M) \) is also open we have

**Lemma 4.** \( \psi \) is \( C^\infty \) on \( \mathbb{R} \) \( f(C) \).

6. Probabilities Near Critical Values

In Sec. 5 we prove that \( \rho \) is smooth except in the neighbourhood of a critical value of \( f \). If \( y_0 = f(c_j) \) is a critical value we choose a region \( N \) which is bounded away from \( c_j \) and then \( \psi_{c_j} \) denotes the contribution to \( \rho \) attributable to \( N \). Then \( \psi_{c_j} \) is smooth and it remains to consider the contribution of a neighbourhood of \( c_j \).

Because \( f \) is restricted \text{Morse, in the sense of Sec. 1, a \( C^\infty \) change of local coordinates in \( M \) is sufficient to reduce this contribution to something like that of one of the three examples in Sec. 2. The details are different, because in Sec. 2 we assume a particular form of density on \( \mathbb{R}^2 \), but we are arguing in a similar way. Once this is done our theorem follows.}

Suppose \( y_0 = d_{c_j} = f(c_j) \) where \( c_j \in C \). For \( y \in (y_0 - \zeta, y_0 + \zeta) \) \( y = y_0 + z \) and \( z \in (0, \zeta) \) we have

\[
P_N(y, z) = P(y, z - P_N(y, z) = \int_{\{y, y + z\} \cap N \cap U(y_0, \beta)} \phi_y
\]

where \( \xi \) is the never-zero two-form \( g^*(\phi_y) \) on \( B(y_0, \beta) \).
Since \( f \) is restricted Morse we can choose \( g \) in such a fashion that either

\[
(1 + ) \quad f^g (u, v) = y_0 + u^2 + v^2
\]
or

\[
(1 - ) \quad f^g (u, v) = y_0 - u^2 - v^2
\]
or

\[
(2) \quad f^g (u, v) = y_0 + u^2 - v^2
\]

while still satisfying the requirements of Sec. 4. We treat the three cases separately.

In case \((1 +)\) choose \( y_0 < y < y_0 + \beta^2 \) and let \( z > 0 \) be small. We have

\[
P_B (y, z) = \int_{\{(u, v) \in B(0, \beta); y - y_0 \leq u^2 + v^2 \leq y - y_0 + \beta^2\}} \xi
\]

\[
= \int_{\sqrt{(y - y_0)} \leq r \leq \sqrt{(y - y_0 + \beta^2)}} \int_{-\pi \leq \theta \leq \pi} r \Xi (r, \theta) \, d\theta \, dr
\]

where \( \Xi (r, \theta) \) is \( C^\infty \) and everywhere positive. Then

\[
\psi (y) - \psi_N (y) = -\frac{\partial}{\partial \xi} P_B (y, z) \bigg|_{z = 0} - \psi_N (y)
\]

\[
= \frac{\partial}{\partial z} (P_N (y, z) + P_B (y, z)) \bigg|_{z = 0} - \psi_N (y)
\]

\[
= \frac{1}{2} \int_{-\pi \leq \theta \leq \pi} \Xi (\sqrt{(y - y_0)}, \theta) \, d\theta
\]

Because \( \psi_N \) is \( C^\infty \) on an open neighbourhood of \( y_0 \), we have

\[
\lim_{y \to y_0^+} \psi (y) = \psi_N (y_0) + \pi \Xi (0, 0).
\]

When \( y < y_0 \) we have \( P_B (y, z) = 0 \) and therefore \( \psi (y) = \psi_N (y) \). Therefore

\[
\lim_{y \to y_0^-} \psi (y) = \psi_N (y_0)
\]

and assertion \((1 -)\) of our theorem follows in the case where \( c_i \) is a point of local minimum. The case of a local maximum is proved by arguing similarly in case \((1 -)\).

In case \((2)\) when \( y_0 < y < y_0 + \beta^2 \) and \( z > 0 \) is small, we have

\[
P_B (y, z) = \int_{\{(u, v) \in B(0, \beta); y - y_0 \leq u^2 + v^2 \leq y - y_0 + \beta^2\}} \xi
\]

\[
= \int_{\sqrt{(y - y_0)} \leq r \leq \sqrt{(y - y_0 + \beta^2)}} \int_{-\pi \leq \theta \leq \pi} r \Xi (r, \theta) \, d\theta \, dr
\]

where \( \Xi (r, \theta) \) is \( C^\infty \), bounded in all derivatives, and everywhere positive. Therefore

\[
\psi (y) - \psi_N (y) = -\frac{\partial}{\partial \xi} P_B (y, z) \bigg|_{z = 0}
\]

\[
= \frac{1}{2} \int_{-\pi \leq \theta \leq \pi} \Xi (\sqrt{(y - y_0)}, \theta) \, d\theta
\]

for some \( \theta^{(0)} \). As \( y \to y_0 \) the order of magnitude of this expression is that of

\[
\ln [\sqrt{(\beta^2 + y - y_0)} + \sqrt{(\beta^2 - y + y_0)}] - \ln (y - y_0) \approx -\ln (y - y_0).
\]

A similar argument in the case where \( y < y_0 \) proves that

\[
\psi (y) - \psi_N (y) \approx -\ln (y_0 - y).
\]

\textbf{Acknowledgements}

I am most grateful to Tim Brown for some stimulating and helpful conversations on density estimation, and for kindly lending me his copies of Silverman [1986], Silverman [1982], and Jones & Lotwick [1984]. It
would be very interesting to examine more closely some of the statistical questions arising from these conversations. I want to thank Alistair Mees for his interest and encouragement, and for valuable discussions about extending the methods of the present paper to other problems in dynamical systems. I am most grateful to Leon Chua for inviting me to submit this account of a seminar on the happy occasion of a visit to Perth by Kazu Aihara. I also thank Professor Chua and an anonymous referee for their helpful comments towards improving the exposition in this paper.

References