Takens' embedding theorem for infinite-dimensional dynamical systems

James C. Robinson

Mathematics Institute, University of Warwick, Coventry, CV4 7AL, U.K.

E-mail: `jcr@maths.warwick.ac.uk`

Abstract. Takens' time delay embedding theorem is shown to hold for finite-dimensional subsets of infinite-dimensional spaces, thereby generalising previous results which were only valid for subsets of finite-dimensional spaces:

Let X be a subset of a Hilbert space with upper box-counting dimension d and 'thickness exponent' τ, which is invariant under a Lipschitz map Φ. Suppose that $\mathcal{E}$, the set of all fixed points of Φ, is not too large, in particular that $d_{\rm f}(\mathcal{E}) < 1/2$. Then for every $k > (2+\tau)d$ such that Φ has no periodic orbits of period $2, \ldots, k$, a prevalent set of Lipschitz observation functions $f : H \to \mathbb{R}$ makes the k-fold time delay map

$$u \mapsto [h(u), h(\Phi(u)), h(\Phi^{k-1}(u))]$$

one-to-one between X and its image. The same result is true if X is a subset of a Banach space provided that $k > 2(1+\tau)d$.

The result follows from a version of Takens' theorem for Hölder continuous maps adapted from Sauer, Yorke, & Casdagli (*J. Stat. Phys.* (1991) **65** 529–547), and makes use of an embedding theorem for finite-dimensional sets due to Hunt & Kaloshin (*Nonlinearity* (1999) **12** 1263–1275).

Submitted to: *Nonlinearity*

1. Introduction

In 1981 Takens published a now celebrated result that allows the dynamics of smooth finite-dimensional systems evolving on smooth manifolds to be reconstructed from an appropriate time series. This was improved by Sauer, Yorke, & Casdagli (1991), who removed the restriction that the dynamical system had to evolve on a manifold, allowing more general fractals, such as those that arise as globally attracting sets in many applications.

However, the treatment has until now been restricted to finite-dimensional systems, excluding the partial differential equations which model a wide variety of physical phenomena, in particular problems in fluid dynamics. It is the purpose of this paper to extend the theory to cover the case when the attractor is a finite-dimensional subset of an infinite-dimensional phase space. The argument is simple, and relies heavily on both the work of Sauer et al. and on a more recent paper of Hunt & Kaloshin (1999) concerning the embedding of finite-dimensional sets into finite-dimensional spaces: indeed, the problem addressed in this paper is highlighted there as an interesting open question.

Suppose that the underlying physical model generates a dynamical system on an infinite-dimensional Hilbert space H, so that the solution at time t through the initial condition u_0 is given by

$$u(t; u_0) = S(t)u_0,$$

with $S(t)$ a semigroup of operators from H into H satisfying the properties

$$S(0) = \mathrm{id}, \qquad S(t)S(s) = S(t+s), \qquad \text{and} \quad S(t)u_0 \quad \text{continuous in } t \text{ and } u_0.$$

Such a dynamical system is generated by many interesting partial differential equations, including the two-dimensional Navier-Stokes equations (see Temam (1988) or Robinson (2001) for details).

For many dissipative equations, it is possible to show that this dynamical system has a *global attractor* $\mathcal{A}$: a compact, positively invariant set which attracts (as $t \to \infty$) the orbits of all bounded sets. The "asymptotic behaviour" of the system can then be regarded as the dynamics of $S(t)$ restricted to $\mathcal{A}$ (cf. Hale, 1988; Robinson, 2001; Temam, 1988). In many cases (see Temam (1988) for numerous examples) these attractors can be shown to be finite-dimensional subsets of the ambient infinite-dimensional phase space, and this paper is concerned with this situation.

Setting $\Phi = S(T)$ for some T, it is sufficient to consider only the dynamics of iterated maps. Within this framework it is shown (under some technical assumptions) that if k is large enough then a prevalent set (made precise in section 2) of Lipschitz mappings $h : H \to \mathbb{R}$ give rise to a time-delay map

$$u \mapsto [h(u), h(\Phi(u)), \dots, h(\Phi^{k-1}(u))]$$

that is $1-1$ on the attractor. [The expression "ϕ is $1-1$ on X" will be used throughout to mean that ϕ is $1-1$ between X and its image.]

2. Prevalence

In line with the treatment in Sauer et al. (1991) and in Hunt & Kaloshin (1999), the theorem here is expressed in terms of 'prevalence'. This concept, which generalises the notion of 'almost every' from finite to infinite-dimensional spaces, was introduced by Hunt, Sauer, & Yorke (1992); see their paper for a detailed discussion.

Definition 2.1 *A Borel subset S of a normed linear space V is* prevalent *if there is a finite-dimensional subspace E of V ('the probe space') such that for each $v \in V$, $v+e$ belongs to S for (Lebesgue) almost every $e \in E$. (In particular, if S is prevalent then S is dense in V.)*

If V is finite-dimensional then this corresponds (via the Fubini theorem) to S being a set whose complement has zero measure.

3. Embedding finite-dimensional sets in $\mathbb{R}^N$

That general finite-dimensional sets can be embedded into a Euclidean space of high enough dimension is a result first due to Mañé (1981). The argument here makes use of a powerful extension of this result due to Hunt & Kaloshin (1999) which gives some information on the smoothness of the parametrisation of the set that is obtained from this embedding. The statement of the result involves the fractal dimension and the "thickness" of the set.

The fractal ('upper box-counting') dimension of a set X, measured in a Banach space $\mathcal{B}$, $d_{\rm f}(X,\mathcal{B})$, is defined as follows. Let $N_{\mathcal{B}}(X,\epsilon)$ be the minimum number of balls of radius ϵ (in the norm of $\mathcal{B}$) necessary to cover the set X. Then

$$d_{\rm f}(X;\mathcal{B}) = \limsup_{\epsilon\to 0} \frac{\log N_{\mathcal{B}}(X,\epsilon)}{-\log \epsilon}.$$

This expression essentially captures the exponent d from the relationship $N_{\mathcal{B}}(X,\epsilon) \sim \epsilon^{-d}$. For more on the 'fractal' dimension see Eden et al. (1994), Falconer (1990), or Robinson (2001).

If X is a subspace of a Banach space $\mathcal{B}$, then the thickness exponent of X in $\mathcal{B}$, $\tau(X;\mathcal{B})$, is a measure of how well X can be approximated by linear subspaces of $\mathcal{B}$.

Denote by $\varepsilon_{\mathcal{B}}(X, n)$ the minimum distance between X and any n-dimensional linear subspace of $\mathcal{B}$. Then

$$\tau(X; \mathcal{B}) = \lim_{n\to\infty} \frac{-\log n}{\log \varepsilon_{\mathcal{B}}(X, n)}, \tag{1}$$

which says that if $\varepsilon_{\mathcal{B}}(X, n) \sim n^{-1/\tau}$ then τ is the thickness exponent of X. (Although less elegant, this form of the definition is perhaps more practical than Hunt & Kaloshin's original; the equivalence of the two definitions is shown in Lemma 2.1 in Kukavica & Robinson, 2004).

Theorem 3.1 *(Hunt & Kaloshin, Theorem 3.6) Let H be a Hilbert space and $X \subset H$ be a compact set with fractal dimension d and thickness exponent τ (measured in H). Let $N > 2d$ be an integer, and let α be a real number with*

$$0 < \alpha < \frac{N - 2d}{N(1 + \tau/2)}. \tag{2}$$

Then for a prevalent set of bounded linear functions $L : H \to \mathbb{R}^N$ there exists $C > 0$ such that for all $x, y \in X$,

$$C|Lx - Ly|^{\alpha} \geq |x - y|. \tag{3}$$

The same result is true if H is a Banach space, but the right-hand side of (2) must be replaced by $(N - 2d)/N(1 + \tau)$.

[The density of Hölder continuous parametrisations of finite-dimensional sets was first shown by Foias & Olson (1996). As well as improving 'density' to 'prevalence', Hunt & Kaloshin also provided the explicit bound on the Hölder exponent in (2).]

Hunt & Kaloshin also give an example that shows, in general, that the upper limit on α of $2/(2 + \tau)$ is sharp, no matter how large the embedding dimension. Since this upper limit becomes one when $\tau = 0$, it is interesting to have a condition guaranteeing that the thickness is zero. One such condition is provided by the following result‡ due to Friz & Robinson (1999).

Proposition 3.2 *Let Ω be a bounded subset of $\mathbb{R}^m$. Suppose that X is a subset of $L^2(\Omega)$ that is uniformly bounded in $H^s(\Omega)$. Then $\tau(X; L^2(\Omega)) \leq m/s$. In particular if X consists of 'smooth functions', i.e. is uniformly bounded in $H^s(\Omega)$ for every $s \in \mathbb{N}$, then $\tau(X; L^2(\Omega)) = 0$.*

‡ There is a small hole in the proof in this paper, since a function in $H^s(\Omega)$ is not necessarily in $D(A^{s/2})$, where A is the Laplacian on Ω with Dirichlet boundary conditions. This can be corrected by first extending each $u \in X$ to a function in $H^s(\Omega')$ that has compact support, for some $\Omega' \supset \Omega$, and then considering the Laplacian on Ω'.

4. A Hölder finite-dimensional Takens theorem

This section gives a statement of a finite-dimensional version of Takens' theorem that allows for maps that are only Hölder continuous.

Theorem 4.1 *(Version of Theorem 4.13 from Sauer et al., allowing for Hölder continuous maps; statement after Hunt & Kaloshin's Theorem 4.1.) Let X be a compact subset of $\mathbb{R}^N$ with $d_{\rm f}(X) = d$. Let $g : X \to X$ be a map such that g^r is a θ-Hölder function for any $r \in \mathbb{N}$. Let $k > 2d/\theta$ and assume that*

(i) the set $\mathcal{E}$ of points $x \in X$ such that $g(x) = x$ satisfies $d_{\rm f}(\mathcal{E}) < 1/2$; and
(ii) X contains no periodic orbits of g of period $2, \ldots, k$.

Let $h_1, \ldots, h_m$ be a basis for the polynomials in N variables of degree at most $2k$, and given any θ-Hölder function $h_0 : \mathbb{R}^N \to \mathbb{R}$ define

$$h_\alpha = h_0 + \sum_{j=1}^{m} \alpha_j h_j.$$

Then the k-fold observation function

$$F^k(h_\alpha, g) = [h_\alpha(x), h_\alpha(g(x)), \ldots, h(g_\alpha^{k-1}(x))]^T$$

is one-to-one on X for almost every $\alpha \in \mathbb{R}^m$.

[Note that (i) the condition that iterates of g be Hölder is in fact only required for $g, \ldots, g^{k-1}$; (ii) the condition on periodic orbits has been strengthened from that in Sauer et al. to that in Hunt & Kaloshin to avoid the problems with linearisation about periodic orbits that would arise from the fact that g is only Hölder continuous; (iii) the standard result requiring $k > 2d$ is recovered if $\theta = 1$.]

Proof. The proof follows that in Sauer et al., apart from minor adjustments in the proof of Lemma 4.4/4.5 where the functions $G_0, \ldots, G_t$ are only taken to be θ-Hölder; the image of any ϵ-ball under G_α is then contained in a ball of radius $C\epsilon^\theta$, and $G_\alpha^{-1}(0)$ is empty for almost every α provided that $r > d/\theta$.

The portion of the argument in the proof of Theorem 4.13 that deals with points lying on distinct periodic orbits requires that the set of elements in $\mathcal{E} \times \mathcal{E}$ has (lower) box-counting dimension less than one. This is certainly true if $d_{\rm f}(\mathcal{E}) < 1/2$. □

5. An infinite-dimensional version of Takens' Theorem

Theorems 3.1 and 4.1 are now combined to give a version of Takens' theorem valid in infinite-dimensional spaces.

Theorem 5.1 *Let $\mathcal{A}$ be a compact set whose fractal dimension satisfies $d_{\rm f}(\mathcal{A}) < d$, $d \in \mathbb{N}$, and which has thickness τ. Choose $k > (2+\tau)d$, and suppose further that $\mathcal{A}$ is an invariant set for a Lipschitz map $\Phi : H \to H$, such that*

(i) the set $\mathcal{E}$ of points in $\mathcal{A}$ such that $\Phi(x) = x$ satisfies $d_{\rm f}(\mathcal{E}) < 1/2$, and
(ii) $\mathcal{A}$ contains no periodic orbits of Φ of period $2, \ldots, k$.

Then a prevalent set of Lipschitz maps $f : H \to \mathbb{R}$ make the k-fold delay embedding $F^k(f, \Phi) : H \to \mathbb{R}^k$ one-to-one on $\mathcal{A}$.

Proof. Given $k > (2+\tau)d$, choose N large enough that

$$k > \frac{N(2+\tau)}{N-2d}d,$$

and then $\alpha < (N-2d)/[N(1+\tau/2)]$ such that $k > (2+\alpha)d$.

Use Theorem 3.1 to find a bounded linear function $L : H \to \mathbb{R}^N$ that is one-to-one on $\mathcal{A}$ and satisfies

$$C|Lx - Ly|^\alpha \geq |x - y| \qquad \text{for all} \qquad x, y \in \mathcal{A}.$$

The set $X = L\mathcal{A} \subset \mathbb{R}^N$ is an invariant set for the induced mapping $g : \mathbb{R}^N \to \mathbb{R}^N$ defined by

$$g(\xi) = L\Phi(L^{-1}\xi).$$

This mapping g is clearly α-Hölder, and since

$$g^n(\xi) = L\Phi^n(L^{-1}\xi)$$

the iterates of g are also α-Hölder.

Observe that if x is a fixed point of Φ^j then $\xi = Lx$ is a fixed point of g^j, and vice versa. It follows that g has no periodic orbits of period $2, \ldots, k$, and that the set of fixed points of g is given by $L\mathcal{E}$. Since L is Lipschitz and the fractal dimension does not increase under the action of Lipschitz maps (see e.g. Robinson (2001)), $d_{\rm f}(L\mathcal{E}) < 1/2$.

Now given a Lipschitz map $f_0 : H \to \mathbb{R}$, define the α-Hölder map $h_0 : \mathbb{R}^N \to \mathbb{R}$ by

$$h_0(\xi) = f_0(L^{-1}\xi) \qquad \text{for all} \qquad \xi \in X.$$

With $\{h_j\}_{j=1}^m$ a basis for the polynomials in N variables of degree at most $2k$, all the conditions of Theorem 4.1 are satisfied, and hence for almost every $\alpha \in \mathbb{R}^m$, the k-fold time delay map on $\mathbb{R}^N$ given by $F^k(h_\alpha, g)$ is one-to-one on X.

Since

$$F^k(h_\alpha, g)(\xi) = [h_\alpha(\xi), h_\alpha(g(\xi)), \ldots, h_\alpha(g^{k-1}(\xi))],$$

and points in $\mathcal{A}$ and X are in one-to-one correspondence via L ($\xi = Lx$, $x = L^{-1}(\xi)$) this is equivalent to the k-fold delay embedding on H given by

$$F^k(f_\alpha, \Phi)(x) = [f_\alpha(x), f_\alpha(\Phi(x)), \ldots, f_\alpha(\Phi^{k-1}(x))],$$

where $\xi = Lx$ and

$$\begin{aligned} f_\alpha &= h_0 \circ L + \sum_{j=1}^{M} \alpha_j (h_j \circ L) \\ &= f_0 + \sum_{j=1}^{M} \alpha_j f_j, \end{aligned}$$

where f_j form a basis for the linear space of polynomials on LH of degree at most $2k$.

It follows that a prevalent set of Lipschitz f make the map $F^k(f, \Phi)$ one-to-one on $\mathcal{A}$. $\square$

Note that the condition on the number of delay coordinates required increases with the thickness of the set $\mathcal{A}$. In the case when $\mathcal{A}$ has zero thickness ($\tau = 0$), this reduces to the $k > 2d$ familiar from the deterministic theory.

It is worth remarking that Yorke (1969) has shown that any periodic orbit of the ordinary differential equation $\dot{x} = F(x)$ must have period at least $2\pi/L$, where L is the Lipschitz constant of F. In the finite-dimensional case this enables the conditions of Theorem 5.1 to be satisfied by taking $g = S(T)$ for T small enough.

A similar result is valid for those infinite-dimensional systems that can be written as semilinear evolution equations (Robinson & Vidal-López, 2004): Let H a Hilbert space H, with norm $|\cdot|$ and inner product $(\cdot, \cdot)$, and let A be an unbounded positive linear operator with compact inverse that acts on H. This means, in particular, that A has a set of orthonormal eigenfunctions $\{w_j\}_{j=1}^{\infty}$ with corresponding positive eigenvalues λ_j, $Aw_j = \lambda_j w_j$, which form a basis for H. Denote by $D(A^\alpha)$ the domain in H of the fractional power A^α, which in this setting has the simple characterisation

$$D(A^\alpha) = \{\sum_{j=1}^{\infty} c_j w_j : \sum_{j=1}^{\infty} \lambda_j^{2\alpha} |c_j|^2 < \infty\}.$$

Following Henry (1981) consider the semilinear evolution equation

$$\mathrm{d}u/\mathrm{d}t = -Au + f(u), \tag{4}$$

where $f(u)$ is globally Lipschitz from $D(A^\alpha)$ into H for some $0 \le \alpha \le 1/2$. Then for each α with $0 \le \alpha \le 1/2$ there exists a constant K_α such that if

$$|f(u) - f(v)| \le L|A^\alpha(u - v)| \qquad \text{for all} \qquad u, v \in D(A^\alpha),$$

any periodic orbit of (4) must have period at least $K_\alpha L^{-1/(1-\alpha)}$. For such examples it follows that condition (ii) of Theorem 5.1 can be satisfied by choosing $\Phi = S(T)$ for any T sufficiently small.

6. Conclusion

Theorem 5.1 generalises Takens' embedding theorem to the infinite-dimensional case, the statement of the result being in a similar form to the finite-dimensional version given in Hunt & Kaloshin (1999).

A related result, originally proved in the periodic case by Friz & Robinson (2001) and recently generalised by Kukavica & Robinson (2004), shows that a sufficiently large number of point observations are sufficient to distinguish between elements of a finite-dimensional set consisting of *analytic* functions (this can be weakened slightly). If $k \geq 16 d_f(\mathcal{A}) + 1$ then almost every set $\mathbf{x} = (x_1, \ldots, x_k)$ of k points in Ω makes the map

$$u \mapsto (u(x_1), \ldots, u(x_k))$$

one-to-one between X and its image.

What would be more desirable in spatially extended systems such as those modelled by partial differential equations would be to construct a one-to-one time series by sampling at a single spatial point. However, this simple form of result cannot be true in general: Consider as in Kukavica & Robinson (2004) the complex Ginzburg-Landau equation (CGLE)

$$u_t - (1 + i\nu) u_{xx} + (1 + i\mu)|u|^2 u - au = 0 \tag{5}$$

with periodic boundary conditions on $\Omega = [0, 1]$. If $a > 4\pi^2$ then such a result is not possible. Indeed, given any $x_0 \in \mathbb{R}$, the two explicit solutions

$$u_j(x,t) = \sqrt{a - 4\pi^2} \exp\Big(2\pi \mathrm{i}(-1)^j (x - x_0) - 4\pi^2 \mathrm{i}\nu t - a\mu \mathrm{i}t + 4\pi^2 \mu \mathrm{i}t\Big)$$

for $j = 1, 2$, which are both contained in the attractor $\mathcal{A}$, coincide at x_0 for all t, while they are clearly distinct. Of course, this does not contradict theorem 4.1, since the set of those observations consisting of point values form a finite-dimensional subset of the Lipschitz observation functions from L^2 into $\mathbb{R}$.

Nevertheless, for this example Kukavica & Robinson (2004) have shown that repeated observations at *two* sufficiently close spatial points do serve to distinguish solutions. Note, however, that it cannot be guaranteed that these time points are equally spaced.

Theorem 6.1 *There exists a $\delta_0 > 0$ such that the following holds: Let x_1 and x_2 be two points with $|x_1 - x_2| \leq \delta_0$, choose $T_0 > 0$, and let $k \geq 16d_{\rm f}(\mathcal{A}) + 1$. Then for almost every set of k times $\mathbf{t} = (t_1, t_2, \dots, t_k)$ where $t_1, \dots, t_k \in [0, T_0]$ the mapping $E_{\mathbf{t}}: \mathcal{A} \to \mathbb{R}^{2k}$ defined by*

$$E_{\mathbf{t}}(u) = \Big([S(t_1)u](x_1), \dots, [S(t_k)u](x_1), [S(t_1)u](x_2), \dots, [S(t_k)u](x_2)\Big)$$

is one-to-one between $\mathcal{A}$ and its image.

It is shown in the same paper that repeated observations at a single point sufficiently close to the boundary does give a one-to-one mapping for the CGLE with Dirichlet boundary conditions; and that observations at four points that are sufficiently close will work for the Kuramoto-Sivashinksy equation.

It is an outstanding problem to prove a version of Takens' theorem based on measurements at a small number of spatial points repeated at equal time intervals.

Acknowledgments

JCR is currently a Royal Society University Research Fellow, and would like to thank the Society for all their support.

References

Eden A, Foias C, Nicolaenko B, and Temam R 1994 *Exponential attractors for dissipative evolution equations* (Chichester: RAM, Wiley)

Falconer K J 1990 *Fractal Geometry* (Wiley, Chichester).

Foias C and Olson E J 1996 Finite fractal dimensions and Hölder-Lipschitz parametrization *Indiana Univ. Math. J.* **45** 603–616

Friz P K and Robinson J C 1999 Smooth attractors have zero 'thickness' *Journal of Mathematical Analysis and Applications* **240** 37–46

Friz P K and Robinson J C 2001 Parametrising the attractor of the two-dimensional Navier-Stokes equations with a finite number of nodal values *Physica D* **148** 201–220

Hale J K 1988 *Asymptotic behaviour of dissipative systems* (Providence: Math. Surveys and Monographs Vol. 25, Amer. Math. Soc.)

Henry D 1981 *Geometric theory of semilinear parabolic equations* (Springer LNM 840, Springer Verlag, Berlin).

Hunt B R, Sauer T, and Yorke J A 1992 Prevalence: a translation-invariant almost every for infinite dimensional spaces *Bull. Amer. Math. Soc.* **27** 217–238; 1993 Prevalence: an addendum *Bull. Amer. Math. Soc.* **28** 306–307

Hunt B R and Kaloshin V Y 1999 Regularity of embeddings of infinite-dimensional fractal sets into finite-dimensional spaces *Nonlinearity* **12** 1263–1275.

Kukavica I and Robinson J C 2004 Distinguishing smooth functions by a finite number of point values, and a version of the Takens embedding theorem *Physica D* **196**, 45–66

Mañé R 1981 On the dimension of the compact invariant sets of certain nonlinear maps *Springer Lecture Notes in Math.* **898** 230–242

Robinson J C 1999 Global attractors: topology and finite-dimensional dynamics *J. Dyn. Diff. Eq.* **11** 557–581

Robinson J C 2001 *Infinite-dimensional dynamical systems* (Cambridge University Press, Cambridge)

Robinson J C and Vidal-López A 2004 Minimal periods of semilinear evolution equations with Lipschitz nonlinearity, Submitted

Sauer T, Yorke J A, and Casdagli M 1993 Embedology *J. Stat. Phys.* **71** 529–547

Takens F 1981 Detecting strange attractors in turbulence *Springer Lecture Notes in Mathematics* **898** 366–381

Temam R 1988 *Infinite dimensional dynamical systems in mechanics and physics* (Berlin: Springer AMS 68)

Yorke J A 1969 Periods of periodic solutions and the Lipschitz constant *Proceedings of the American Mathematical Society* **22** 509–512