BULLETIN BRAZILIAN MATHEMATICAL SOCIETY

Bull Braz Math Soc, New Series 33(2), 231-262

The reconstruction theorem for endomorphisms

Floris Takens

— Dedicated to IMPA on the occasion of its 50^{th} *anniversary*

Abstract. The reconstruction theorem deals with dynamical systems which are given by a map $\varphi : M \to M$ together with a read out function $f : M \to \mathbb{R}$. Restricting to the cases where φ is a diffeomorphism, it states that for generic (φ, f) there is a bijection between elements $x \in M$ and corresponding sequences $(f(x), f(\varphi(x)), \ldots, f(\varphi^{k-1}(x)))$ of k successive observations, at least for k sufficiently big. This statement turns out to be wrong in cases where φ is an endomorphism.

In the present paper we derive a version of this theorem for endomorphisms (and which is equivalent to the original theorem in the case of diffeomorphisms). It justifies, also for dynamical systems given by endomorphisms, the algorithms for estimating dimensions and entropies of attractors from obervations.

Keywords: reconstruction, read out function, generic, dimension, entropy.

Mathematical subject classification: 37C20, 37A35, 37C45, 37C40, 62M10.

1 Introduction

In this paper we discuss the analogue of the reconstruction theorem, see [T1], [A], and [SYC], for dynamical systems which are given by an endomorphism. We first give the statement of the reconstruction theorem for diffeomorphisms and then discuss the situation for endomorphisms.

The original reconstruction theorem deals with dynamical systems, given by a diffeomorphism $\varphi : M \to M$ on a compact manifold M together with a function $f : M \to \mathbb{R}$. Both φ and f are supposed to be at least C^1. The diffeomorphism φ determines the time evolution, or dynamics, and the function f is interpreted here as a *read out function*. This setup is supposed to represent the situation of a dynamical system where one has only partial information about the states as a function of time: if the system is in the state $x \in M$, one observes, or measures, only the value of $f(x)$. So an evolution $\{x_n = \varphi^n(x_0)\}$ leads

Received 20 June 2002.

to the observations $\{y_n = f(x_n)\}$; such a sequence of observations is called a *time series*. The reconstruction theorem deals with the question whether one can reconstruct from such observations information about the evolution of the dynamical system itself.

In fact, the original reconstruction theorem also dealt with systems given by a vector field (continuous time systems). Both in the case of systems given by diffeomorphisms and by vector fields, the past and the future can be deduced from the present. Systems given by endomorphisms are much like those given by diffeomorphisms except that in the case of endomorphisms one cannot deduce in general the past from the present. There is no such analogue for vector fields where the present does not determine the past. Since we are primarily interested here in systems given by endomorphisms we disregard systems given by vector fields because they are not useful as analogues.

For the formulation of the theorem we need some definitions. Given a diffeomorphism (or endomorphism) $\varphi : M \to M$ and a read out function f as above, we define for any k the *reconstruction map* $\mathrm{Rec}_k : M \to \mathbb{R}^k$ by

$$\mathrm{Rec}_k(x) = (f(x), f(\varphi(x)), \dots, f(\varphi^{k-1}(x))).$$

The image of M under this map is denoted by $\mathcal{X}_k$. We call vectors of the form $(f(x), f(\varphi(x)), \dots, f(\varphi^{k-1}(x))) \in \mathbb{R}^k$ *(k-dimensional) reconstruction vectors*.

If the dimension k is clear from the context we may write Rec or $\mathcal{X}$ instead of Rec_k and $\mathcal{X}_k$; if it is necessary to specify the dynamical system (φ, f) which is used in defining these objects, we write $\mathrm{Rec}^{(\varphi,k)}$ and $\mathcal{X}^{(\varphi,f)}$.

Theorem 1. *There is an open and dense subset $\mathcal{U} \subset \mathrm{Diff}^1(M) \times C^1(M)$, the product of the space of C^1-diffeomorphisms on M and the space of C^1-functions on M, such that, whenever $(\varphi, f) \in \mathcal{U}$ and $k > 2m$, then Rec_k is an embedding of M into $\mathbb{R}^k$, implying that $\mathcal{X}_k$ is a submanifold of $\mathbb{R}^k$ which is diffeomorphic to the state space manifold M.* □

This means that for generic (φ, f), i.e. belonging to the open and dense subset $\mathcal{U}$, and $k > 2m$, each state x of the system is uniquely determined by the k 'measurements' which one obtains if the systems follows 'its dynamics' starting at the state x.

In the case where φ is not a diffeomorphism but an endomorphism the above theorem is no longer true: there are persistent examples where Rec_k is not an embedding. We mean here persistent in the sense that there are no C^1-small perturbations of the given system (φ, f) making Rec_k an embedding. Such examples will be discussed in section 2. What one still can prove is that, under generic assumptions, and for $k > 2m$, a sequence of k successive measurements does determine the state of the system at the *end* of the sequence of measurements;

in general, for each of these 'final states' there are however several corresponding reconstruction vectors.

Theorem 2. *There is an open and dense subset $\mathcal{U} \subset \mathrm{End}^1(M) \times C^1(M)$, where $\mathrm{End}^1(M)$ denotes the space of C^1-endomorphisms on M, such that, whenever $(\varphi, f) \in \mathcal{U}$ and $k > 2m$, there is a map $\pi_k : \mathcal{X}_k \to M$ with $\pi_k \,\mathrm{Rec}_k = \varphi^{k-1}$.*

Moreover the map π_k has bounded expansion, or is Lipschitz, meaning that for all $p \neq q \in \mathcal{X}_k$ the the ratio of the distances

$$\frac{d(\pi_k(p), \pi_k(q))}{\| \; p - q \; \|},$$

is bounded by some constant C which is independent of p and q; d denotes the distance on M with respect to some Riemannian metric. □

Remark. We should point out that without extra work we prove a somewhat stronger statement: the map π_k is differentiable, in the sense that it admits a differentiable extension to a neighbourhood of $\mathcal{X}_k$ in $\mathbb{R}^k$. This in spite of the fact that $\mathcal{X}_k$ is in general not a manifold, so that it has no 'differentiable structure'. For the applications it is important that π_k is Lipschitz; the fact that it is differentiable is, as far as I know, of no use. □

Though this theorem is much weaker than in the case of diffeomorphisms in the sense that neither $\mathrm{Rec}_k : M \to \mathcal{X}_k$ nor π_k needs to be a diffeomorphism, for many practical purposes the result is just as good as in the case of diffeomorphisms. We shall discuss this in section 6, but we indicate here already the applications which we have in mind:

- under the generic assumptions, the set $\mathcal{X}_k$, for $k > 2m + 1$, completely determines the deterministic structure of the time series produced by the dynamical system, in the sense that from any segment of at least $k - 1$ successive values of such a time series all future values can be deduced, using the shape of $\mathcal{X}_k$;

- the estimation of dimension and entropy from observed data, as discussed in [GPa], [GPb], [T3] and [KS] is also justified for systems where the dynamics is given by an endomorphism (provided the generic conditions are satisfied).

The paper is organised as follows. In section 2 we discuss various examples, essentially all the local singularities and self intersections which occur in reconstruction maps of generic dynamical systems with state space manifolds of dymension 1 or 2. These examples show indeed that in the case of dynamical systems given by an endomorphism self intersections and non-immersion points occur in a persistent way so that in general $\mathcal{X}_k$ is not a manifold. In section 3 we formulate the *reconstruction condition*, which has to be satisfied in the case of endomorphisms for the conclusion of theorem 2 to be valid. In section 4 we prove that this reconstruction condition is indeed generically satisfied if $k > 2m$. In section 5 we prove that the reconstruction condition implies the existence of the map π_k with the properties as announced in the above theorem 2. Finally, in section 6, we discuss the applications refered to above.

2 Examples

In this section we give examples of the local structure of reconstruction maps of generic dynamical systems. We restrict to the cases where the state space manifold M is low dimensional (in fact only the dimensions 1 and 2 are considered). Also we shall assume that there are no fixed points and periodic orbits (of low period) and that the dynamical systems are sufficiently differentiable (this may mean more than just C^1). Within all these restriction we describe all local phenomena which occur in reconstruction maps of generic dynamical systems. The examples are mainly intended to explain the rather complicated reconstruction condition in section 3.

2.1 Generic reconstructions of 1-dimensional systems

For a generic dynamical system (φ, f) on a 1-dymensional manifold M, the derivative of f is only zero at isolated points. At the points where the derivative of f is non-zero, the reconstruction map Rec_k, for $k \geq 1$, is an immersion. Now we consider such an exceptional point $p \in M$ where the derivative of f is zero. Since we assume that the dynamical system (φ, f) is generic we may assume that at these exceptional points there are no other exceptional things. Otherwise we would have a situation with co-dimension at least 2 and this should not happen in the 1-dimensional manifold M (for the use of the notion of co-dimension in relation with genericity arguments, see section 4). In particular, we may assume that $\mathrm{d}\varphi(p)$ is non-zero, that $\varphi(p) \neq p$, and that the derivative of f in $\varphi(p)$ is non-zero. It is easy to see that this implies that Rec_k, for $k \geq 2$ is an immersion at p. So as a first result we have:

Lemma 1. *For generic dynamical systems on a 1-dimensional manifold M the reconstruction map* Rec_k*, for $k \geq 2$, is an immersion (in each point of M).* $\square$

Next we consider possible self-intersections. For this we have to consider pairs of points $(p, q) \in M \times M$ with $p \neq q$. For such points to have the same image under a reconstruction map we need at least that $f(p) = f(q)$. This is a co-dimension 1 condition, so in $M \times M$ there will be generically a

1-dimensional submanifold, i.e. a curve, where this condition is satisfied. For such a pair to have the same image under Rec_2, we must have that $\varphi(p) = \varphi(q)$, a co-dimension 1 condition, or $\varphi(p) \neq \varphi(q)$ and $f(\varphi(p)) = f(\varphi(q))$, which is also a co-dimension 1 condition. If the former of these conditions is satisfied, then we have that $\mathrm{Rec}_k(p) = \mathrm{Rec}_k(q)$ for all k; generically this will happen at isolated points of the 1-dimensional manifold in $M \times M$ defined by the condition $f(p) = f(q)$. If the latter of these conditions is satisfied, then we have $\mathrm{Rec}_k(p) = \mathrm{Rec}_k(q)$ for $k = 1, 2$; in order to have the same for $k = 3$ we need a further condition to be satisfied, namely $\varphi^2(p) = \varphi^2(q)$ or $\varphi^2(p) \neq \varphi^2(q)$ but $f(\varphi^2(p)) = f(\varphi^2(q))$. Again both these conditions imply one more co-dimension. So this would be a co-dimension 3 situation. Generically this does not happen on the 2-dimensional manifold $M \times M$. So the only way in which self-intersections for Rec_k, $k \geq 3$ can occur for generic 1-dimensional systems is at pairs p, q with $p \neq q$, $f(p) = f(q)$ and $\varphi(p) = \varphi(q)$. It is not hard to show by similar arguments that generically there are no triple points, i.e. no three points $p \neq q \neq r \neq p$ such that $\mathrm{Rec}_k(p) = \mathrm{Rec}_k(q) = \mathrm{Rec}_k(r)$ with $k \geq 2$.

Lemma 2. *For generic dynamical systems on a 1-dimensional manifold M the reconstruction map Rec_k, $k \geq 3$ may have double points but not triple points. Moreover, if $\mathrm{Rec}_k(p) = \mathrm{Rec}_k(q)$, with $p \neq q$ and $k \geq 3$, then $\varphi(p) = \varphi(q)$.* □

So we see here the first instance of a reconstruction map of a generic dynamical system which is not generic as map: generic maps from a 1-dimensional manifold in $\mathbb{R}^3$ are embeddings.

2.2 Generic reconstruction maps of 2-dimensional systems

A similar analysis can be carried out in the case the dimension of M is 2, the details are however more extensive. The first difference is that, even for generic dynamical systems, the reconstruction maps are in general no longer immersions. This can be seen as follows. For a point $p \in M$ it is a co-dimension 1 condition that the derivative $\mathrm{d}\varphi$ is not invertible in p. So this happens generically along so-called fold curves. Along these fold curves it is another co-dimension 1 condition that the kernels of $\mathrm{d}\varphi$ and $\mathrm{d}f$ coincide. So this will occur (persistently) in isolated points for generic 2-dimensional dynamical systems. In such a point a reconstruction map cannot be an immersion: if v is a non-zero vector in the common kernels of $\mathrm{d}\varphi$ and $\mathrm{d}f$, then this vector will be mapped to zero by the derivative of any reconstruction map. This is in fact the only way in which the reconstruction maps Rec_k, with $k \geq 4$ can fail to be immersions. If the dynamical system is sufficiently differentiable and if the higher order terms of φ and f are generic, then the resulting singularity in Rex_3 forms a Whitney umbrella, see [W4]. This singularity remains in Rec_k for $k > 3$. So we have:

Lemma 3. *For a generic dynamical system (φ, f) on a 2-dimensional manifold M the reconstruction map Rec_k, with $k \geq 4$ can have isolated points where it is not an immersion; in these points it has a singularity in the form of a Whitney umbrella.* □

With arguments similar to those used in the 1-dimensional case, we obtain for generic 2-dimensional systems that Rec_k, with $k \geq 5$ may have curves of double points and also isolated double points. The former ones correspond to those pairs of points $p \neq q$ for which $\varphi(p) = \varphi(q)$, the latter correspond those pairs $p \neq q$ for which $\varphi(p) \neq \varphi(q)$, but $\varphi^2(p) = \varphi^2(q)$. Also there will be isolated tripple points; in fact they will be the intersection of 3 lines of double points of the first kind. However there will never be four different points which are mapped to the same reconstruction vector.

3 The reconstruction condition for endomorphisms

In this section we formulate the condition which has to be satisfied by (φ, f) in order to belong to the set $\mathcal{U}$ in theorem 2. For this we need some definitions.

Definition (meeting number). Let (φ, f) be a dynamical system on M and let $(x, \tilde{x})$ be an element of $M \times M$. The meeting number $j(x, \tilde{x})$ of x and $\tilde{x}$ is the smallest integer such that $\varphi^{j(x,\tilde{x})}(x) = \varphi^{j(x,\tilde{x})}(\tilde{x})$. If no such number exists, then the meeting number is ∞. This meeting number depends on φ; if this needs to be expressed in the notation, we write $j^{\varphi}(x, \tilde{x})$. Note that $j(x, \tilde{x}) = 0$ if and only if $x = \tilde{x}$. □

Definition (s-embedding). Let (φ, f) be a dynamical system on a manifold M. We say that a point $x \in M$ is s-embedding if the co-vectors $\mathrm{d}f(x)$, $\mathrm{d}(f\varphi)(x), \ldots, \mathrm{d}(f\varphi^{s-1})(x)$ contain a co-basis of $T_x(M)$. Note that this condition is equivalent with the condition that Rec_s, restricted to some neighbourhood of x, is an embedding into $\mathbb{R}^s$.

Note that if x is s-embedding for (φ, f), then this is still so for $\tilde{x}$ sufficiently close to x and $(\tilde{\varphi}, \tilde{f})$ C^1 sufficiently close to (φ, f); this is due to the fact that 'forming a basis' is persistent under small perturbations. □

Definition (property $\mathcal{P}$). Let (φ, f) be a dynamical system on a manifold M, and let k be an integer. We say that a pair of points $(x, \tilde{x}) \in M \times M$ has the property $\mathcal{P}_k$ if:

- $\mathrm{Rec}_k(x) \neq \mathrm{Rec}_k(\tilde{x})$, or
- for some $j(x, \tilde{x}) \leq j < k$, $\varphi^j(x) = \varphi^j(\tilde{x})$ is $(k - j)$-embedding.

We note that in this definition we do not exclude the case $x = \tilde{x}$. Also observe that if the pair $x, \tilde{x}$ has the property $\mathcal{P}_k$ then it has the property $\mathcal{P}_l$ for $l \geq k$.

When the value of k is irrelevant, or clear from the context, we write $\mathcal{P}$ instead of $\mathcal{P}_k$. □

Definition (reconstruction condition). Let (φ, f) be a dynamical system on a manifold M and let k be some integer. We say that (φ, f) satisfies the k-reconstruction condition if each pair of points $(x, \tilde{x}) \in M \times M$ has the property $\mathcal{P}_k$. □

There is an equivalent form of the reconstruction condition which will be useful in section 5.

Lemma. *A dynamical system (φ, f) on a manifold M satisfies the k-reconstruction condition if and only if for each k-dimensional reconstruction vector ξ there are an integer $j(\xi)$ and a point $p(\xi) \in M$ such that, whenever* $\mathrm{Rec}_k(x) = \xi$, *then*

- $\varphi^{j(\xi)}(x) = p(\xi)$;
- $p(\xi)$ *is* $(k - j(\xi))$*-embedding.* □

Proof. First we assume that (φ, f) satisfies the k-reconstruction condition. For $x \in M$ the property $\mathcal{P}_k$ then holds for the pair (x, x). This means that there is some $j < k$ such that $\varphi^j(x)$ is $(k - j)$-embedding. We define $j(x)$ to be the biggest such integer. Then, if, for $x \neq \tilde{x}$, we have $\mathrm{Rec}_k(x) = \mathrm{Rec}_k(\tilde{x})$, it follows from the property $\mathcal{P}_k$ that there is some j such that $\varphi^j(x) = \varphi^j(\tilde{x})$ and such that $\varphi^j(x)$ is $(k - j)$-embedding. This however implies that we have $j(x) = j(\tilde{x})$ (since we required $j(x)$ respectively $j(\tilde{x})$ to be the *biggest* integer such that $\varphi^{j(x)}(x)$ respectively $\varphi^{j(\tilde{x})}(\tilde{x})$ is $(k - j(x))$, respectively $(k - j(\tilde{x}))$-embedding). In other words we can define for a k-dimensional reconstruction vector ξ the integer $j(\xi)$ as $j(x)$ for any x with $\mathrm{Rec}_k(x) = \xi$. Also we can define $p(\xi)$ as $\varphi^{j(x)}(x)$ whenever $\mathrm{Rec}_k(x) = \xi$. With these definitions it is clear that $j(\xi)$ and $p(\xi)$ have the properties as announced in the lemma.

Next we assume that we have for each reconstruction vector ξ an integer $j(\xi)$ and a point $p(\xi) \in M$ with the properties as formulated in the lemma. Let $(x, \tilde{x})$ be any pair of points in $M \times M$. If $\mathrm{Rec}_k(x) \neq \mathrm{Rec}_k(\tilde{x})$ then the pair $(x, \tilde{x})$ has the property $\mathcal{P}_k$. If $\mathrm{Rec}_k(x) = \mathrm{Rec}_k(\tilde{x}) = \xi$, then it follows from the various definitions that $j(x, \tilde{x}) \leq j(\xi)$ and that the pair $(x, \tilde{x})$ has the property $\mathcal{P}_k$ with $j = j(\xi)$. □

4 The proof of the genericity

Our first step, in section 4.1, is to prove that for a pair $(x, \tilde{x}) \in M \times M$ the property $\mathcal{P}_k$ is persistent, both under C^1-small perturbations of φ and f and under small perturbations of x and $\tilde{x}$. The rest of the proof of genericity is based on transversality. In section 4.2 we recall the basic ideas of this method as we will use it. In section 4.3 we prove the genericity result for pairs $(x, \tilde{x})$ where x or $\tilde{x}$ is close to a periodic point of φ with 'low' period or to 'some of the pre-images' of such periodic points with low period. In section 4.4 we finally complete the genericity proof.

4.1 Persistence of the property $\mathcal{P}$

We assume we are given a dynamical system (φ, f), and an integer k such that the property $\mathcal{P}_k$ holds for $(x, \tilde{x}) \in M \times M$. We shall prove that this property remains if the points x and $\tilde{x}$ are slightly perturbed and also if φ and f are slightly perturbed in the C^1-sense. In order to prove this, we assume that we have sequences x_n, $\tilde{x}_n$, φ_n, and f_n converging to x, $\tilde{x}$, φ, and f respectively and derive a contradiction from the *assumption* that for each n the property $\mathcal{P}_k$ does not hold for (φ_n, f_n) at $(x_n, \tilde{x}_n)$.

If for $(x, \tilde{x})$ the first alternative holds, i.e., if for some $0 \leq j < k$ we have $f(\varphi^j(x)) \neq f(\varphi^j(\tilde{x}))$ then we also have, for n sufficiently large, that $f_n(\varphi_n^j(x_n)) \neq f_n(\varphi_n^j(\tilde{x}_n))$ which implies property $\mathcal{P}_k$ for n sufficiently large: a contradiction with our assumption. This means that from now on we may assume that for $(x, \tilde{x})$ the second alternative holds, and even that for each $0 \leq j < J = j(x, \tilde{x})$ we have $f(\varphi^j(x)) = f(\varphi^j(\tilde{x}))$ ($j(x, \tilde{x})$ is the meeting number as defined in section 3).

If necessary by restricting to a subsequence, we may assume that the values of $j^{\varphi_n}(x_n, \tilde{x}_n)$ are independent of n. We denote this value by J'. Clearly $J \leq J'$. If $J = J'$ we obtain a contradiction because the property of being s-embedding is persistent under small perturbations, see section 3.

For $J \leq j < J'$ the points $\varphi_n^j(x_n)$ and $\varphi_n^j(\tilde{x}_n)$ are very close, but not equal. We can interpret the difference between these points as 'infinitesimal vectors'. This can be made precise by taking local coordinates in a neighbourhood of each of the points $\varphi^j(x) = \varphi^j(\tilde{x})$. With respect to such coordinates, we have unit vectors $v_{n,j}$ and positive real numbers $\varepsilon_{n,j}$ (converging to zero for $n \to \infty$) such that $\varphi_n^j(x_n) - \varphi_n^j(\tilde{x}_n) = \varepsilon_{n,j} v_{n,j}$. If necessary by restricting again to a subsequence, we have that for each j (still with $J \leq j < J'$) the limit $v_j = \lim_{n\to\infty} v_{n,j}$ exists. These limit vectors belong to $T_{\varphi^j(x)}(M)$ and are independent, up to scalar multiplication, of the coordinates which we used. One easily verifies that, for some real λ_j, $\mathrm{d}\varphi(v_j) = \lambda_j v_{j+1}$ for $J \leq j < J' - 1$ and that $\mathrm{d}\varphi(v_{J'-1}) = 0$. Furthermore, for each j such that $J \leq j < J'$, we have that $\mathrm{d}f(v_j) = 0$, because otherwise, for n sufficiently large, we would have $f_n(\varphi_n^j(x_n)) \neq f_n(\varphi_n^j(\tilde{x}_n))$ which would imply property $\mathcal{P}_k$ to hold for such n which is again against our assumption. This means that each of the vectors v_j is in the kernel of $\mathrm{d}(f\varphi^{j'})(\varphi^j(x))$ for all $j' \geq 0$.

We return to the fact that the second alternative in the condition $\mathcal{P}_k$ holds for (φ, f) at $(x, \tilde{x})$. This means that there is an integer $\bar{j} \geq J$ such that $\varphi^{\bar{j}}(x)$ is $(k - \bar{j})$-embedding. We concluded above that for every j with $J \leq j < J'$ there is a non-zero vector in $T_{\varphi^j(x)}(M)$, namely v_j, which is in the kernel of $\mathrm{d}(f\varphi^{j'})(\varphi^j(x))$ for all $j' \geq 0$. This means that we must have $\bar{j} \geq J'$. Now it follows that, for n sufficiently large, also $\varphi_n^{\bar{j}}(x_n)$ is $(k - \bar{j})$-embedding for (φ_n, f_n), meaning that the property $\mathcal{P}_k$ holds for such (φ_n, f_n) at $(x_n, \tilde{x}_n)$. This is the final contradiction which completes the proof of the persistence of the property $\mathcal{P}_k$.

4.2 Transversality

For a good exposition of the theory of transversality we refer to [B] and [H]. We recall here the definition of transversality, prove the transversality theorem for a simple situation, and indicate the role of the *perturbing families* which are necessary in the proof of our transversality results.

Definition. Let V be a submanifold of a manifold W and let f be a C^1-map from another manifold N into W. We say that f is transversal with respect to V if we have for each $x \in N$ either $f(x) \notin V$ or $\mathrm{d}f(T_x(N)) + T_{f(x)}(V) = T_{f(x)}(W)$. □

Remarks.

1. Note that if the derivative of f is, for each $x \in N$, a surjective map from $T_x(N)$ to $T_{f(x)}(W)$, then f is transversal with respect to any submanifold of W.

2. Note that if the dimension of N is smaller than the co-dimension of V in W, i.e. smaller than $\dim(W) - \dim(V)$, then f is transversal with respect to V if and only if $f(N)$ is disjoint from V.

3. It is easy to see that if V is a topologically closed submanifold and if N is compact, then the set of C^k-maps, $k \geq 1$, from N to W which are transversal with respect to V is open in the C^k-topology; in the case where moreover the dimension of N is smaller than the co-dimension of V in W, as in the above remark, transversality is even open in the C^0-topology.

4. If f is transversal with respect to V, then $\tilde{V} = f^{-1}(V)$ is a submanifold of N and, if $\tilde{V}$ is nonempty, the co-dimension $\dim(N) - \dim(\tilde{V})$ of $\tilde{V}$ in N equals the co-dimension $\dim(W) - \dim(V)$ of V in W.

Theorem (Thom's transversality lemma [T6], [H]). *Let V be a submanifold of W which is topologically closed and let N be a compact manifold. Then the set of C^k-maps, $k \geq 1$, which are transversal with respect to V is open and dense in the C^k-topology.* □

Sketch of the proof for the case that W is a vector space. As we observed before, transversality is open in the C^1-topology. Since C^l- and even C^∞-maps are dense in the C^1-maps, it is enough to prove only density of the transversal maps in some C^l-topology with $l \geq 1$.

Here we only indicate the proof for the case where $W = \mathbb{R}^L$ for some L. Let $f : N \to \mathbb{R}^L$ be a C^l-map. We want to show that we can approximate f by a map which is transversal with respect to V. For this we take a parametrised family $f_\mu : N \to \mathbb{R}^L$, where the parameter μ has values in an open neighbourhood P of the origin in some vector space, such that

- $f_0 = f$;
- the map $F : P \times N \to \mathbb{R}^L$, defined by $F(\mu, x) = f_\mu(x)$ is C^l and transversal with respect to V.

In this case one can take $P = \mathbb{R}^L$ and $f_\mu(x) = f(x) + \mu$. It is easy to see that with this definition the above two conditions are satisfied. Note that we don't have to specify $V \subset \mathbb{R}^L$ since the derivative of F is surjective everywhere.

Next we consider the submanifold $\tilde{V} = F^{-1}(V)$ and its projection π on the parameter space P. According to Sard's theorem, e.g. see [B] or [H], the set C_π of critical values of π, i.e. the set of parameter values, which is defined by $C_\pi = \{\mu \in P \mid \exists(\mu, x) \in \tilde{V}$ such that $d\pi_{(\mu,x)} : T_{(\mu,x)}(\tilde{V}) \to T_\mu(P)$ is *not* surjective$\}$, has Lebesgue measure zero. In fact, for Sard's theorem to apply we need π and hence f to be sufficiently differentiable (the required differentiability l is the maximum of 1 and $(\dim(\tilde{V}) - L + 1)$). As we observed in the beginning of the proof this is no problem: we could assume f to be as differentiable as needed. Finally it is not hard to verify that a parameter value μ is non-critical if and only if f_μ is transversal with respect to V. This means that arbitrarily close to 0 there are parameter values for which f_μ is transversal with respect to V. □

Remarks.

1. In the statement of the transversality theorem, and in its proof, we can take N to be a non-compact manifold with a compact subset K. Then the conclusion is that for an open and dense set of maps from N to W transversality holds on K. We say that $f : N \to W$ is transversal with respect to V on K if for any $x \in K$ we have either $f(x) \notin V$ or $\mathrm{d}f(T_x(N)) + T_{f(x)}(V) = T_{f(x)}(W)$. The conclusion holds both for the strong and the weak topology on the space of C^k-maps from N to W, $k \geq 1$. Also, if the manifold V is not closed as a subset of W, but contains a subset L which is closed in W the conclusion of the theorem holds for transversality with respect to V restricted to L. We say that $f : N \to W$ is transversal with respect to V restricted to L if for each $x \in N$ we have either $f(x) \notin L$ or $\mathrm{d}f(T_x(N)) + T_{f(x)}(V) = T_{f(x)}(W)$.

2. We call the parametrised family f_μ in the above proof a *perturbing family*. If one wants to prove the above theorem for an arbitrary manifold W, the only part of the proof which needs adaptation is the construction of the perturbing family.

Perturbing families for dynamical systems. In the proof that the reconstruction condition is generic (for appropriate k), we shall use several times an argument of the following type:

We have some construction C which assigns to each endomorphism $\varphi : M \to M$ and read out function $f : M \to \mathbb{R}$ a map $C(\varphi, f) : C_M \to C_M{}'$, where C_M and $C_M{}'$ are smooth manifolds and where $C_M{}'$ has a submanifold S_M. We then need to show that for generic (φ, f), $C(\varphi, f)$ is transversal with respect to S_M. So what we will have to show then is that we can perturb $C(\varphi, f)$ in a sufficiently general way by perturbing φ and f. More precisely, we need a parametrised family (φ_μ, f_μ), μ in some parameter space, such that $\mu \mapsto C(\varphi_\mu, f_\mu)$ is a perturbing family as introduced in the proof of the transversality theorem. It happens that there is one parametrised family (φ_μ, f_μ) (still depending on k) which is sufficiently rich to generate the required perturbing families for all the constructions which we will need. It is this parametrised family (φ_μ, f_μ) which we will now construct. It will be denoted by *the perturbing family of* (φ, f); note that we use here the term ‘perturbing family’ in a meaning which is somewhat different from the original one, namely a family of perturbations of (φ, f) which *induces* a perturbing family in the original sense of maps like $C(\varphi, f)$.

In order to construct this perturbing families for φ and f, we need to fix the value k in the reconstruction condition and to identify the state space manifold M with a submanifold of $\mathbb{R}^L$, which certainly can be done if $L > 2\dim(M)$, see [H]. We denote by p a smooth projection of a (small) neighbourhood U of M in $\mathbb{R}^L$ to M so that $p|M$ is the identity in M. We then define E as the linear space of polynomials of degree at most $(4k-1)$ on $\mathbb{R}^L$. This degree is chosen to ensure that for each collection of $2k$ pairwise different points $X^1, \dots, X^{2k} \in \mathbb{R}^L$, real numbers $\alpha^1, \dots, \alpha^{2k}$ and co-vectors $\beta^1, \dots, \beta^{2k} \in (\mathbb{R}^L)^\star$ there is an element $g \in E$ such that $g(X^i) = \alpha^i$ and $\mathrm{d}g(X^i) = \beta^i$ for all $i = 1, \dots, 2k$. In order to see that this degree $(4k-1)$ suffices, we note the following:

There is a polynomial p of degree $(4k-2)$ on $\mathbb{R}^L$ which is zero, and has zero derivative, in the points $X^1, \dots, X^{2k-1}$ and which is nonzero everywhere else, e.g. take $p(x) = \prod_{i=1}^{2k-1} \parallel X^i - x \parallel^2$. Then, multiplying this polynomial with a polynomial of degree one, we kan make the value and the first derivative in X^{2k} whatever we want, without changing the values or derivatives in the other points $X^1, \dots, X^{2k-1}$.

As a parameter space P for our perturbations of φ and f we take the $(L+1)$-fold power E^{L+1} of E, or at least a neighbourhood of the origin 0 in that vector space. For $\mu = (g_0, \dots, g_L) \in P$ we define

- $f_\mu = f + g_0|_M$;
- $\varphi_\mu = p(\varphi + (g_1, \dots, g_L))$, where $(g_1, \dots, g_L)$ should be interpreted as the restriction, of the polynomial map $(g_1, \dots, g_L)$ on $\mathbb{R}^L$, to M.

The neighbourhood P of the origin in E^{L+1} should be so small that the image of $\varphi + (g_1, \dots, g_L)$ is contained in U, the neighbourhood on which the projection p on M is defined, whenever $(g_0, \dots, g_L) \in P$.

The use of this perturbing family is further explained in connection with the jet extensions to be discussed below.

Jet extensions. For any smooth map $g : V \to W$, the 1-jet of this map in a point $y \in V$ consists of the pair $y, g(y)$ together with the derivative $\mathrm{d}g(y)$ of g at the point y. For the dynamical systems which we are considering here, consisting of a smooth endomorphism φ and a smooth read out function f, both defined on M, the 1-jet at a point $x \in M$ consists of x, $\varphi(x)$ and $f(x)$, together with the derivatives $\mathrm{d}\varphi(x)$ and $\mathrm{d}f(x)$. We call the map, assigning to each $x \in M$ the 1-jet of (φ, f), the *1-jet extension of* (φ, f); it is denoted by $J^1(\varphi, f)$ and it is a map whose degree of differentiability is one less than that of (φ, f); its range can obviously be given the structure of a smooth manifold. This range is a vector bundle over $M \times M \times \mathbb{R}$, whose fibre over (x_1, x_2, s) is the product of the set of linear maps $\mathrm{Lin}(T_{x_1}(M), T_{x_2}(M))$, from the tangent space at x_1 to the tangent space at x_2, and $T^\star_{x_1}(M)$; in this representation the 1-jet of (φ, f) at x corresponds to the 'base point' $(x, \varphi(x), f(x))$, while the element of the fibre determines the derivatives of φ and f respectively at x. The space of all these 1-jets is denoted by $J^1(M)$.

Apart from 1-jets, we also consider *multi-1-jets*. First we define $M^{\sim K}$ as the set of K-tuples $(X^1, \dots, X^K)$ of points in M which are pairwise different. The

K-1-jet of (φ, f) at $(X^1, \dots, X^K)$ is the sequence of 1-jets of (φ, f) at the successive points $X^1, \dots, X^K$; the map assigning to each such $(X^1, \dots, X^K)$ its K-1-jet is the K-1-jet extension of (φ, f) and is denoted by $J^{K,1}(M)$, its range by $J^1(M)^{\sim K}$. The proposition below formulates the main property of our perturbing family of (φ, f) in terms of multi-1-jets. Its proof is straightforward.

Proposition. *For the parametrised family (φ_μ, f_μ), as constructed above, the map which assigns to each element $(\mu, X^1, \cdots, X^{2k})$ of $P \times M^{\sim 2k}$ the $2k$-1-jet of (φ_μ, f_μ) at $(X^1, \dots, X^{2k})$ has a derivative which is everywhere surjective, and hence is transversal with respect to any submanifold of $J^1(M)^{\sim 2k}$.* □

Remark. In order to prove that the k-reconstruction condition is generic (for appropriate values of k) we will need to apply the transversality argument to $2k$-1-jet extensions of (φ, f), or to maps which are derived from such extensions. Since in the transversality theorem one assumes maps to be at least C^1, one would expect that (φ, f) should belong to some open and dense subset in the C^2-topology. The reason that the C^1-topology is sufficient is due to the fact that, wherever 1-jet extensions are involved, we will only use the transversality argument in situations where transversality means 'no intersection' in which case transversality is even open in the C^0-topology. Another argument, showing that the C^1-topology is sufficient, is based on the fact that the property $\mathcal{P}_k$ is persistent under C^1-small perturbations.

4.2.1 Transversality with respect to a (semi-)algebraic subset

The Thom transversality theorem concerning maps from N to W can be generalised to the case where the subset V of W is no longer a submanifold but an algebraic, or even a closed semi-algebraic, subset (an algebraic subset is given by algebraic equalities, in the definition of a semi-algebraic subset also (algebraic) inequalities may occur). The reason that such a generalisation is possible is based on the fact that closed (semi-)algebraic subsets admit Whitney stratifications; for the proof we refer to [L], see also [T9] or [GWPL]. A Whitney stratification of a closed set is a decomposition of the set into a finite number of manifolds, or strata, such that for each stratum S, its topological boundary $\bar{S} \setminus S$, is contained in the union of strata of lower dimension, and such that certain compatibility conditions are satisfied. These compatibility conditions imply that for any point x of a stratified set, belonging to a stratum S and any sequence x_i converging to x and belonging to a stratum S' and such that the tangent spaces $T_{x_i}(S')$ converge to a limit, the tangent space $T_x(S)$ is contained in that limit. Moreover, in a stratified set V, the union of all strata of dimension smaller than or equal to i, also called the i-dimensional skeleton of V and denoted by V^i, is again a closed stratified set.

It follows from this compatibility condition that, whenever the manifold N is compact, the maps from N to W which are transversal with respect to all the strata of some closed (semi-)algebraic subset V, or, more generally a closed stratified subset V, of W form an open subset in the C^1-topology. The density of transversal maps is obtained by induction on the dimension of the skeletons:

Transversality with respect to V^0 is certainly dense because V^0 is a topologically closed submanifold. Maps which are transversal with respect to all strata in V^i are, in a sufficiently small neighbourhood of V^i, also transversal with respect to the strata in V^{i+1} (due to the compatibility conditions). The remaining part of V^{i+1} is a closed subset of a smooth manifold, so there transversality follows as usual. In this way we obtain transversality with respect skeletons of increasing dimensions and finally with respect to all of V.

4.2.2 Reconstruction maps of generic dynamical systems

Here we show how the method of perturbing families can be used for our reconstruction problem. In fact, what we do here is to prove, as an example, a part of our general genericity theorem.

Proposition. *Let $\varphi : M \to M$ be an endomorphism, let $k > 2\dim(M)$ be an integer and let $K \subset M \times M$ be a compact subset such that for any pair $(x, \tilde{x}) \in K$ we have:*

- *the meeting number $j^\varphi(x, \tilde{x})$ is at least k;*
- *the first $2k - 1$ iterates of x, namely $x, \varphi(x), \dots, \varphi^{2k-1}(x)$ are pairwise different;*
- *the first $2k - 1$ iterates of $\tilde{x}$, namely $\tilde{x}, \varphi(\tilde{x}), \dots, \varphi^{2k-2}(\tilde{x})$ are pairwise different.*

Then there is an open and dense subset in $C^1(M)$, the space of C^1-functions on M, such that for f in that subset and $(x, \tilde{x}) \in K$ we have $\mathrm{Rec}_k^{(\varphi,f)}(x) \neq \mathrm{Rec}_k^{(\varphi,f)}(\tilde{x})$. □

Proof. Let E be a vector space of functions on M such that, whenever $X^1, \dots, X^{2k}$ are pairwise different points on M and $\alpha^1, \dots, \alpha^{2k} \in \mathbb{R}$, there is a function $g \in E$ such that $g(X^i) = \alpha^i$ for $i = 1, \dots, 2k$. For the construction of those (finite dimensional) vector spaces we refer back to the discussion of the pertubing families in section 4.2. For a function f_0 on M we consider the perutbing family f_μ, with $\mu \in E$ given by:

$$f_\mu(x) = f_0(x) + \mu(x).$$

To each μ we associate the map $R_\mu : M \times M \to \mathbb{R}^k$, defined by

$$R_\mu(x, \tilde{x}) = \mathrm{Rec}_k^{(\varphi, f_\mu)}(x) - \mathrm{Rec}_k^{(\varphi, f_\mu)}(\tilde{x}),$$

and the corresponding map $R : E \times M \times M \to \mathbb{R}^k$ by

$$R(\mu, (x, \tilde{x})) = R_\mu(x, \tilde{x}).$$

Our proposition follows (by the transversality argument) if the map R has, in each point of $E \times K$, a surjective derivative.

For a pair $(x, \tilde{x})$, such that all the $2k$ points $\varphi^i(x)$ and $\varphi^i(\tilde{x})$, for $i = 0, \dots, k-1$, are pairwise different it is obvious that the derivative of R is surjective in each point $(\mu, (x, \tilde{x}))$. For a pair $(x, \tilde{x}) \in K$, the only way in which parts of the two orbit segments $x, \dots, \varphi^{k-1}(x)$ and $\tilde{x}, \dots, \varphi^{k-1}(\tilde{x})$ may coincide is

- for some $0 \leq i < j$, $\varphi^{i+s}(x) = \varphi^{j+s}(\tilde{x})$ for $s \geq 0$ and no other points coincide;
- the same as above with the roles of x and $\tilde{x}$ interchanged.

(Here we made heavily use of the second and third condition in the proposition which for example exclude the possibility that $\tilde{x} = \varphi^{k-1}(x)$ and at the same time $x = \varphi^{k-1}(\tilde{x})$.)

It is however not hard to verify that also in these cases, one can find for each $(\alpha^0, \dots, \alpha^{k-1}) \in \mathbb{R}^k$ an element $\mu \in E$ such that for each $i = 0, \dots, k-1$, we have $\mu(\varphi^i(x)) - \mu(\varphi^i(\tilde{x})) = \alpha^i$. This implies that in each point of $E \times K$ the derivative of R is surjective. This completes the proof. $\square$

4.3 The property $\mathcal{P}$ at (pre-)periodic points of low period

We now proceed to the actual genericity proof, or rather to the density proof since we have already the openness of the property $\mathcal{P}$. In this proof the periodic points of low period (and a number of their inverse images) need a special treatment. That is the subject of the present section.

As before we consider dynamical systems, given by an endomorphism φ and a function f on a closed manifold M of dimension m. Since we only have to establish the density of the reconstruction condition, and since the C^∞-mappings are dense in the C^1-mappings, we may assume φ and f to be as differentiable as needed.

For the proof of the main theorem we only need to consider the property $\mathcal{P}_k$ with $k > 2m$ (so that we may even restrict to $k = 2m+1$); in the present section we can even take $k > m$ (for the most part even $k \geq m$); for the time being, we will restrict to $k = m$ and refer to the property $\mathcal{P}_m$ as the property $\mathcal{P}$.

Fixed points. For a dynamical system (φ, f) we consider its 1-jet extension $J^1(\varphi, f)$, which maps each $x \in M$ to $(x, \varphi(x), f(x), \mathrm{d}\varphi(x), \mathrm{d}f(x))$. Generically this extension is transversal to the fixed point submanifold V, defined by $\{x = \varphi(x)\}$. Since V has co-dimension m in $J^1(M)$, the fixed points are isolated for generic φ.

Next we define the smaller subset $\tilde{V} \subset V$ by the conditions:

- $x = \varphi(x)$ and
- $(\mathrm{d}f(x), \dots, \mathrm{d}(f\varphi^{m-1})(x))$ is linearly dependent in $T_x^\star(M)$, or $\mathrm{d}\varphi(x)$ has an eigenvalue which is 0 or an l^{th} root of unity with $l \leq 4m$.

Since the second condition is algebraic (at least one of a finite number of determinants has to be zero), $\tilde{V}$ admits a stratification. We now show that $\tilde{V}$ is a proper subset of V and hence consists of strata whose co-dimension in $J^1(M)$ is at least $m+1$. For this we first show that there are a linear map $A : \mathbb{R}^m \to \mathbb{R}^m$ and a linear function $F : \mathbb{R}^m \to \mathbb{R}$, i.e. an element of $(\mathbb{R}^m)^\star$, such that $(F, FA = A^\star(F), \dots, FA^{m-1} = (A^\star)^{m-1}(F))$ is a basis of $(\mathbb{R}^m)^\star$ and such that A has no eigenvalue which is 0 or which is an l^{th} root of unity with $l \leq 4m$. One can take for example for A a linear map whose (real) matrix is diagonal with no multiple eigenvalues and no eigenvalues equal to 0 or ± 1 and F given by $F(u_1, \cdots, u_m) = u_1 + \cdots + u_m$.

In the vector space $\mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^m) \times (\mathbb{R}^m)^\star$ the set of those elements (A, F), for which $(F, FA, \dots FA^{m-1})$ is linearly dependent or for which A has an eigenvalue which is 0 or an l^{th} root of unity with $l \leq 4m$, is an algebraic set with non-empty complement (by the above argument). Hence it has no interior points. So it is stratified with strata of co-dimension at least 1 in $\mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^m) \times (\mathbb{R}^m)^\star$. This implies that $\tilde{V}$ is stratified and that its strata have at least co-dimension $m+1$ in $J^1(M)$.

So for generic (φ, f) the image of the 1-jet extension is disjoint from $\tilde{V}$ and transversal to V. For dynamical systems which are generic in this sense we have:

- the fixed points are isolated (hence there is only a finite number of them) and the derivative $\mathrm{d}\varphi$ in such a fixed point has no eigenvalue which is 0 or an l^{th} root of unity with $l \leq 4m$ (so that in a small neighbourhood of these fixed points there are no periodic points with period at least 2 and at most $4m$);
- whenever x is a fixed point, x is m-embedding and hence the property $\mathcal{P}$ holds at (x, x).

We shall explain later why we also excluded the eigenvalue 0, see the discussion on pre-periodic points in this section.

To the above generic conditions we add another condition which is clearly generic: there should not be two different fixed points where the value of f is the same. If this condition holds, then:

– whenever x and $\tilde{x}$ are fixed points, the property $\mathcal{P}$ holds at $(x, \tilde{x})$.

Due to the persistence of the property $\mathcal{P}$, and the other generic conditions, we have for these generic (φ, f): there are a neighbourhood $\mathcal{U}$ of (φ, f) in the C^1-topology and a neighbourhood U of the set of fixed points of φ such that: whenever $(\tilde{\varphi}, \tilde{f})$ is in $\mathcal{U}$ and $x, \tilde{x} \in U$, then:

– the property $\mathcal{P}$ holds for $(\tilde{\varphi}, \tilde{f})$ at $(x, \tilde{x})$;

– if x is a fixed point of $\tilde{\varphi}$, then $x \in U$ and $\mathrm{d}\tilde{\varphi}(x)$ has no eigenvalue which is 0 or an l^{th} root of unity with $l \leq 4m$;

– the neighbourhood U contains no periodic points of $\tilde{\varphi}$ with period at least 2 and at most $4m$;

– if $x \neq \tilde{x}$ are both fixed points of $\tilde{\varphi}$, then $f(x) \neq f(\tilde{x})$.

Periodic points. We will need the analogue of the above result also for all periodic points of period at most $4m$. We will treat the case of points of period 2 in detail and in such a way that it is clear how to proceed by induction.

We assume that the dynamical system (φ, f) is generic in the above sense so that there are neighbourhoods $\mathcal{U}$ and U with the above mentioned properties. We may assume that all the perturbations of (φ, f), which we describe below, remain within $\mathcal{U}$. Since in U there are no points of period 2, we can restrict our transversality arguments to the compact complement K of U in M.

With a *first* arbitrarily small perturbation of φ we obtain that on K, the complement of U, the map $M \ni x \mapsto (x, \varphi^2(x)) \in M \times M$ is transversal with respect to the diagonal $\Delta \subset M \times M$. This implies that the points of period 2 are isolated.

As in the case of fixed points, with a *second*, arbitrarily small, perturbation we obtain that:

– for each $x \in K$ with $\varphi^2(x) = x$, the co-vectors $\mathrm{d}f(x), \ldots, \mathrm{d}(f\varphi^{m-1})(x)$ form a basis of $T_x^\star(M)$, so that x is m-embedding and $\mathrm{d}\varphi^2(x)$ has no eigenvalue which is 0 or an l^{th} root of unity with $l \leq 4m$.

Finally, with a *third* arbitrarily small perturbation we can arrange that for any two points $x \neq \tilde{x}$, belonging to the set of points with period at most 2 (including the fixed points), we have $f(x) \neq f(\tilde{x})$.

For dynamical systems (φ, f) which have generic fixed points and points of period 2 in the above sense, is follows that whenever $x, \tilde{x}$ belong to the set of points of period at most 2 (including fixed points), then the property $\mathcal{P}$ holds at $(x, \tilde{x})$. Due to persistence we have for such generic (φ, f):

There are a neighbourhood $\mathcal{U}'$ of (φ, f) in the C^1-topology and a neighbourhood U' of the set of points with period at most 2 (including the fixed points) such that for $x, \tilde{x} \in U'$ and $(\tilde{\varphi}, \tilde{f}) \in \mathcal{U}'$:

- the property $\mathcal{P}$ holds for $(\tilde{\varphi}, \tilde{f})$ at $(x, \tilde{x})$;
- in the neighbourhood U' there are no points of period greater than 2 and at most $4m$;
- if x is a fixed point or a point of period 2 of $\tilde{\varphi}$, then $x \in U'$ and $\mathrm{d}\tilde{\varphi}(x)$, respectively $\mathrm{d}\tilde{\varphi}^2(x)$, has no eigenvalue which is 0 or an l^{th} root of unity with $l \leq 4m$;
- if $x \neq \tilde{x}$ and both x and $\tilde{x}$ have period at most 2, then $f(x) \neq f(\tilde{x})$.

It is now clear how one can proceed with the successive periods 3, 4 etcetera up to period $4m$. The final result will be stated explicitely after we have considered also the pre-periodic points.

Pre-periodic points. We assume that (φ, f) is a dynamical system which is generic in the above sense (for all periodic orbits of periods up to $4m$). Let P denote the set of all these points of low period, i.e. of period at most $4m$. We now consider the set of *first pre-low-periodic points* P^1 which is defined as $P^1 = \varphi^{-1}(P) \setminus P$. We note that, due to the fact that in the case of periodic points of low period we avoided derivatives with eigenvalue 0, P^1 is bounded away from P (in fact it was for this reason that we excluded the eigenvalue 0). From this it easily follows that it is a generic property for φ that the points of P^1 are isolated and that the derivative $\mathrm{d}\varphi$ in each point of P^1 is invertible: this is equivalent with the property that φ, restricted to the complement of a small neighbourhood of P, is transversal with respect to P.

The next property for these first pre-low-periodic points which we need to be generic is that for each $x \in P^1$ we have that

$$\mathrm{d}f(x), \ldots, \mathrm{d}(f\varphi^{m-1})(x)$$

is a basis of $T_x^\star(M)$, i.e. that x is m-embedding. In order to see that also this is generic, we observe that for each $x \in P^1$ we have $\varphi(x) \in P$; due to the generic conditions, holding for P, we have that

$$\mathrm{d}f(\varphi(x)), \ldots, \mathrm{d}(f\varphi^{m-1})(\varphi(x))$$

is a basis of $T^\star_{\varphi(x)}(M)$ and hence

$$\mathrm{d}f(\varphi(x)), \ldots, \mathrm{d}(f\varphi^{m-2})(\varphi(x))$$

are linearly independent; due to the fact that $\mathrm{d}\varphi(x)$ is invertible, also the $m-1$ co-vectors

$$\mathrm{d}(f\varphi)(x), \ldots, \mathrm{d}(f\varphi^{m-1})(x)$$

are linearly independent. So for generic f, $\mathrm{d}f(x)$ is linearly independent of the above $m-1$ co-vectors in $T_x^\star(M)$, and hence $\mathrm{d}f(x)$ completes the basis as required. This implies that generically the property $\mathcal{P}$ holds at (x, x).

Finally, the last generic property is that for each pair of points $x \neq \tilde{x}$ in $P \cup P^1$, we have $f(x) \neq f(\tilde{x})$. This last property implies that also for $x \neq \tilde{x}$, both in $P \cup P^1$, the property $\mathcal{P}$ holds at $(x, \tilde{x})$.

Next we define the *second pre-low-periodic points* P^2 as $P^2 = \varphi^{-1}(P^1) = \varphi^{-1}(P \cup P^1) \setminus (P \cup P^1)$. From this definition it is clear that one can prove, by induction, the same type of generic properties for P^2 as for P^1. In the following we mainly need this genericity up to P^{4m}.

Statement of the generic properties for pre-low-periodic points. For a dynamical system (φ, f) we define the set $(\tilde{P}(\varphi))$ of pre-low-periodic points as the set of all periodic points (fixed points including) whose period is at most $4m$ (this is the set of low-periodic points) together with all those points which are mapped by φ^{4m} into this set of low-periodic points. Now from the above consideration it follows that in the space of C^1-dynamical systems there is an open and dense subset (of generic dynamical systems) such that for each dynamical system (φ, f) in this subset we have neighbourhoods $\mathcal{U}$ of (φ, f) in the C^1-topology and U of $\tilde{P}(\varphi)$ in M such that for each $(\tilde{\varphi}, \tilde{f}) \in \mathcal{U}$ and $x, \tilde{x} \in U$ we have:

- $\tilde{P}(\tilde{\varphi})$ is contained in U;
- if $x \neq \tilde{x}$ and if they are both in $\tilde{P}(\tilde{\varphi})$, we have $\tilde{f}(x) \neq \tilde{f}(\tilde{x})$;
- the co-vectors $\mathrm{d}\tilde{f}(x), \ldots, \mathrm{d}(\tilde{f}\tilde{\varphi}^{m-1})(x)$ form a basis of $T_x^\star(M)$, i.e. x is m-embedding;
- the property $\mathcal{P}$ holds for $(\tilde{\varphi}, \tilde{f})$ at $(x, \tilde{x})$;
- if x is a low-periodic point with period $i \leq 4m$, then $\mathrm{d}(\tilde{\varphi})^i(x)$ has no eigenvalue which is 0 or an l^{th} root of unity for $l \leq 4m$;
- $\mathrm{d}\tilde{\varphi}(x)$ is invertible.

Remark. In the above discussion, extending the generic properties to the pre-periodic points, we could just as well have stopped the process to include pre-images at a different order, say up to $\varphi^{-l}(P(\varphi))$. This will be needed in one of the arguments below. We will denote the set $\varphi^{-l}(P(\varphi))$ by $\tilde{P}^l(\varphi)$; so $\tilde{P}(\varphi) = \tilde{P}^{4m}(\varphi)$. If we want to refer to this version of the above generic properties, we will refer to *genericity on* $\tilde{P}^l(\varphi)$.

Consequences for reconstruction maps. We discuss here consequences of the above generic properties for reconstruction maps as introduced in the introduction. We also derive one further generic property which implies that generically for $k > m$ the property $\mathcal{P}_k$ holds at any pair $(x, \tilde{x})$ with x *or* $\tilde{x}$ close to a pre-low-periodic point; this means that in the rest of the proof of the density of the k-reconstruction condition, we may assume that there are no periodic point with low period.

First we recall the definition of the k-dimensional reconstruction map $\mathrm{Rec}_k : M \to \mathbb{R}^k$ for a dynamical system (φ, f):

$$\mathrm{Rec}_k(x) = (f(x), f(\varphi(x)), \ldots, f(\varphi^{k-1}(x))).$$

For $k \geq m$, the above generic properties imply that, for a sufficiently small neighbourhood U of $\tilde{P}(\varphi)$, $\mathrm{Rec}_k|_U$ is an embedding into $\mathbb{R}^k$, and hence injective.

Our first objective here is to show that with an arbitrarily small perturbation of (φ, f) we can obtain, for $k > m$, that, whenever $x \in \tilde{P}(\varphi)$ and $\tilde{x} \neq x$ we have $\mathrm{Rec}_k(x) \neq \mathrm{Rec}_k(\tilde{x})$. We call this the *injectivety property*. If this property holds for $k = m + 1$, it holds for all $k > m$. We note that this is the first instance where we need to restrict to $\mathcal{P}_k$ with $k > m$.

With a *first* small perturbation we obtain that the above generic property also holds on $\tilde{P}^{5m+1}(\varphi)$. Then there is a neighbourhood U of $\tilde{P}^{5m+1}(\varphi)$ such that $\mathrm{Rec}_{m+1}|_U$ is an embedding, and hence injective. We show that with a *second* small perturbation of f alone we can obtain the required injectivity property. So we have to show that we can obtain that the image under Rec_{m+1} of the complement of some neighbourhood of $\tilde{P}(\varphi)$ is disjoint from $\mathrm{Rec}_{m+1}(\tilde{P}(\varphi))$. Since $\mathrm{Rec}_{m+1}(\tilde{P}(\varphi))$ is 0-dimensional (it contains only a finite number of points) this means that Rec_{m+1} has to be transversal to $\mathrm{Rec}_{m+1}(\tilde{P}(\varphi))$ when restricted to the compact complement of some neighbourhood of $\tilde{P}(\varphi)$. For this we consider perturbations of f which vanish on $\tilde{P}(\varphi)$ and which are so small that $\mathrm{Rec}_{m+1}|_{U'}$ remains an embedding, where U' is some open neighbourhood such that $\tilde{P}^{5m+1}(\varphi) \subset U' \subset U$. Now we observe that in each point $x \notin U'$ we have that the points $x, \varphi(x), \ldots, \varphi^m(x)$ are pairwise different and non of them is contained in $\tilde{P}(\varphi)$. This means that the perturbations of f in these $m + 1$ different points are 'free' and independent. From this it easily follows that with such a small perturbation of f we can make Rec_{m+1}, restricted to $M - U'$ transversal with respect to $\mathrm{Rec}_{m+1}(\tilde{P}(\varphi))$. Then $\mathrm{Rec}_{m+1}(M \setminus U')$ is disjoint from $\mathrm{Rec}_{m+1}(\tilde{P}(\varphi))$.

Next we show that the injectivety property is persistent under C^1-small perturbations of (φ, f), provided (φ, f) is generic in the sense discussed before. We know already that the generic condition on $\tilde{P}(\varphi)$ implies that there are neighbourhoods $\mathcal{U}$ and U of (φ, f) and $\tilde{P}(\varphi)$ respectively such that whenever $(\tilde{\varphi}, \tilde{f}) \in \mathcal{U}$, $\tilde{P}(\tilde{\varphi}) \subset U$, and $\mathrm{Rec}_{m+1}^{(\tilde{\varphi}, \tilde{f})}$, restricted to U, is an embedding. Since $M - U$ is compact and since $\mathrm{Rec}_{m+1}(M - U)$ and $\mathrm{Rec}_{m+1}(\tilde{P}(\varphi))$ are disjoint, they have a positive distance. This means that this property of $\mathrm{Rec}_{m+1}(M - U)$ being disjoint from $\tilde{P}(\varphi)$ is indeed open, even in the C^0-topology.

This indeed implies the persistence of the injectivety property for generic (φ, f). Hence the injectivety property holds for an open and dense set of dynamical systems.

We now formulate the main conclusion of this section.

Proposition. *For C^1 dynamical systems (φ, f) on M generically the following is true:*

There are a neighbourhood $\mathcal{U}$ of (φ, f) in the C^1-topology and a neighbourhood U of $\tilde{P}(\varphi)$, the set of periodic points with period at most $4m$ together with those point which are mapped by φ^{4m} to a point with period at most $4m$, such that for any $(\tilde{\varphi}, \tilde{f}) \in \mathcal{U}$ we have

i) $\tilde{P}(\tilde{\varphi}) \subset U$;

ii) $\mathrm{Rec}_m^{(\tilde{\varphi},\tilde{f})} \mid U$ *is an embedding;*

iii) $\mathrm{Rec}_{m+1}^{(\tilde{\varphi},\tilde{f})}(U)$ *and* $\mathrm{Rec}_{m+1}^{(\tilde{\varphi},\tilde{f})}(M - U)$ *are disjoint;*

iv) whenever x and $\tilde{x}$ are in U, the property $\mathcal{P}_{m+1}$ holds for $(\tilde{\varphi}, \tilde{f})$ at $(x, \tilde{x})$.

Due to the item iii, *we conclude even that:*

iv') whenever x or *$\tilde{x}$ is in U, the property $\mathcal{P}_{m+1}$ holds for $(\tilde{\varphi}, \tilde{f})$ at $(x, \tilde{x})$.* □

We note here that whenever, in the notation of the above proposition, $x \notin U$, the points $x, \tilde{\varphi}(x), \ldots, \tilde{\varphi}^{4m}(x)$ are pairwise different. This will be important in the next section. It was also the justification of the hypothesis in the proposition in section 4.2.2 that the orbit segments $x, \varphi(x), \ldots, \varphi^{2k-2}(x)$ (and the same for $\tilde{x}$) consist of $2k - 1$ pairwise different points.

4.4 The property $\mathcal{P}$ away from the periodic points

We now come to the last part of the genericity proof. From the formulation of the property $\mathcal{P}$ it is clear that we have to investigate, what we will call *linear systems of length $l + 1$*. We will give the formal definition below, but in terms of a dynamical system (φ, f) one can give the following description. They consist of a sequence of $l + 1$ vector spaces of dimension m, like

$$T_x(M), T_{\varphi(x)}(M), \ldots, T_{\varphi^l(x)}(M)$$

connected by linear maps, like

$$\mathrm{d}\varphi(x), \mathrm{d}\varphi(\varphi(x)), \ldots, \mathrm{d}\varphi(\varphi^{l-1}(x)),$$

and with each of the vector spaces equipped with a linear function like

$$\mathrm{d}f(x), \ldots, \mathrm{d}f(\varphi^{l}(x)).$$

In particular we are interested in the question how exceptional (in the sense of co-dimension) it is that one cannot select for any $0 \leq i \leq l$ a co-basis in $T_{\varphi^i(x)}(M)$ from the co-vectors $\mathrm{d}f(\varphi^i(x)), \ldots, \mathrm{d}(f\varphi^{l-i})(\varphi^i(x))$. It will be (notationally) more convenient to formulate the results on these linear systems in a somewhat more abstract setting.

4.4.1 Linear systems

We assume that we are given an infinite collection of m-dimensional vector spaces $V_0, V_1, \ldots$. We then define a linear system of length $l+1$ as a collection C consisting of l linear maps $A_1 : V_0 \to V_1, \ldots, A_l : V_{l-1} \to V_l$ and of $l+1$ linear functions $f_0 : V_0 \to \mathbb{R}, \ldots, f_l : V_l \to \mathbb{R}$. So the set of linear systems of length $l+1$ is the vector space

$$(\oplus_{i=1}^{l}\mathrm{Lin}(V_{i-1}, V_i)) \oplus (\oplus_{i=0}^{l} V_i^{\star})$$

where $\mathrm{Lin}(V_{i-1}, V_i)$ denotes the vector space of linear maps from V_{i-1} to V_i. This means that we can speak of the co-dimension of certain properties of linear systems: for such a property there is a corresponding subset of those linear systems which have the property in question. The co-dimension of a property is then the co-dimension of the corresponding set in the vector space of all linear systems (in our considerations these subsets will always be closed algebraic subsets so that the co-dimension is well defined).

In some situations we shall also consider linear systems of length $l+1$ which are not based on the vector spaces $V_0, \ldots, V_l$ but on $V_j, \ldots, V_{j+l}$. Also we will use the notion of a *restricted linear system* of length $l+1$. This means just that the last linear map ($f_l : V_l \to \mathbb{R}$) is not included.

Definition. Let $C = (A_1, \ldots, A_l, f_0, \ldots, f_l)$ be a linear system of length $l+1$. We say that a pair of indices $0 \leq i \leq j < l$ is a *blocking pair* if the following holds:

- on the kernel K of the composition $A_{j+1} \ldots A_{i+1} : V_i \to V_{j+1}$ the restrictions of the linear functions $f_i, f_{i+1}A_{i+1}, \ldots, f_j A_j \ldots A_{i+1}$ do not contain a co-basis of K;
- there is no index j' with $i \leq j' < j$ so that the above item also holds when j is replaced by j'. □

The term *blocking pair* indicates that, even if we modify the linear system after f_j and A_{j+1}, where this modification may include extending the length of the linear system, it will never be possible that the co-vectors f_i, $f_{i+1}A_{i+1}$, $f_{i+2}A_{i+2}A_{i+1}, \dots$ contain a co-basis of V_i. Note that i, j being a blocking pair is really a property of the *restricted* linear system of length $j - i + 2$ based on the vector spaces $V_i, \dots, V_{j+1}$.

Lemma 1. *In the space of restricted linear systems of length* $s + 1$, *the set of those restricted systems, for which the indices* $0, s - 1$ *form a blocking pair, is algebraic and has co-dimension at least* $s + 1$. □

Proof. For the indices $0, s - 1$ to form a blocking pair for the restricted linear system $C = (A_1, \dots, A_s, f_0, \dots, f_{s-1})$ of length $s + 1$, the kernel K of the composition $A_s \dots A_1$ should have positive dimension.

Since the dimension of K is at most equal to the sum of the dimensions of the kernels of A_1 up to A_s, and since in the space of linear maps between m-dimensional spaces, the elements with a t-dimensional kernel form an algebraic subset of co-dimension at least t (in fact that co-dimension is t^2). This means that the condition on the maps $A_1, \dots, A_s$ that their composition has a kernel of dimension l has co-dimension at least l. If $l > s$ we even don't have to consider possible conditions on the linear functions $f_0, \dots, f_{s-1}$.

If we assume that the linear maps $A_1, \dots, A_s$ are given and that the kernel K of their composition has dimension l, then, in order that C has $0, s-1$ as a blocking pair we need that among the co-vectors f_0, $f_1A_1, \dots, f_{s-1}A_{s-1} \dots A_1$, restricted to K, there is no co-basis for K. We denote the dimension of the linear subspace of the dual $K^\star$ of K, spanned by the co-vectors $f_0|_K$, $f_1A_1|_K, \dots, f_iA_i \dots A_1|_K$ by n_i, where $i = 0, \dots, s - 1$, and define $n_{-1} = 0$. Clearly, for each $i = -1, \dots, s - 2$ we have $n_i \le n_{i+1} \le n_i + 1$. Since, in order that the indices $0, s - 1$ form a blocking pair, we need $n_{s-1} < l$. This implies that there must be a collection of at least $s - l + 1$ indices i for which $n_i = n_{i+1}$ (note that if $s - l + 1 \le 0$, then $l > s$ and, as we saw above, we know already that this corresponds to a situation which has co-dimension at least $s + 1$). For each index i with $n_i = n_{i+1}$ there are two alternatives: either f_{i+1} has to satisfy a condition, which has at least co-dimension 1, in order to make $n_{i+1} = n_i$ or f_{i+1} does not have to satisfy any condition because, due to the previous $A_1, \dots, A_{i+1}$, $f_0, \dots, f_i$, no choice of f_{i+1} could lead to n_{i+1} being $n_i + 1$. In the latter case however the indices $0, i$ would already form a blocking pair. In that case, since $i < s - 1$, the indices $0, s - 1$ cannot form a blocking pair. So the former alternative has to hold. This means that each of the $s - l + 1$ indices, for which $n_i = n_{i+1}$, represents a restriction on $f_0, \cdots, f_{s-1}$ corresponding to one co-dimension. So we conclude that the whole linear system has to satisfy a collection of conditions with total co-dimension $s + 1$: l for the restrictions on $A_1, \dots, A_s$ in order to obtain l as the dimension of K and $s - l + 1$ for the restrictions on $f_0, \dots, f_{s-1}$. (So for each dimension l of the kernel we arrive at the same total co-dimension and hence the value of l does not appear in the final result.) This comples the proof. $\square$

Lemma 2. *In the space of linear systems of length $s+1$, with $s+1 \geq m$, the set of those systems for which there is no $0 \leq i < s$ such that $0, i$ is a blocking pair and such that the co-vectors $f_0, f_1 A_1, \dots, f_s A_s \dots A_1$ do not contain a co-basis of V_0, have co-dimension at least $s - m + 2$.* $\square$

Proof. Let $C = (A_1, \dots, A_s, f_0, \dots, f_s)$ be a linear system of length $s+1$ such that there is no index $0 \leq i < s$ such that $0, i$ is a blocking pair (this is an 'open' condition) and such that the co-vectors $f_0, \dots, f_s A_s \dots A_1$ do not contain a co-basis of V_0. We denote by n_i, $i = 0, \dots, s$, the dimension of the linear subspace in $V_0^\star$ spanned by the co-vectors $f_0, f_1 A_1, \dots, f_i A_i \dots A_1$; we define $n_{-1} = 0$. As in the proof of the previous lemma we have for $i = -1, \dots, s-1$ that $n_i \leq n_{i+1} \leq n_i + 1$. By our assumption $n_s < m$. This means that the set of indices i for which $n_i = n_{i+1}$ has at least $s - m + 2$ elements. For each of these indices, assuming $A_1, \dots A_i$ and $f_0, \dots f_i$ are given, f_{i+1} has to satisfy a co-dimension 1 condition: the argument is similar to the argument we used in the proof of lemma 1: if there is no linear function $\tilde{f}_{i+1}$ on V_{i+1} such that $\tilde{f}_{i+1} A_{i+1} \dots A_1$ is linearly independent of $f_0, \dots, f_i A_i \dots A_1$, then $0, i$ is a blocking pair, contradicting our assumption. This means that f_{i+1} indeed had to satisfy a condition with at least co-dimension 1 in order to have $n_{i+1} = n_i$. So in total the co-dimension is at least $s - m + 2$. This proves the lemma. $\square$

Proposition. *In the space of linear systems of length $s+1 \geq m$ we consider the subset of (exceptional) linear systems for which there are no integers $0 \leq j < s+1$ and $0 \leq j_1 < \dots < j_m$ such that $j + j_m \leq s$ and such that*

$$f_{j+j_1} A_{j+j_1} \dots A_{j+1}, \dots, f_{j+j_m} A_{j+j_m} \dots A_{j+1}$$

is a co-basis of V_j. This subset of these exceptional linear systems has co-dimension at least $s - m + 2$. (The statement of this proposition should be compared with the definitions in section 3.) $\square$

Proof. Let $C = (A_1, \dots, A_s, f_0, \dots, f_s)$ be a linear system which belongs to the above (exceptional) set. There are two possibilities: either there is an index $0 \leq i < s$ such that $0, i$ is a blocking pair or there is no such index. If there is no such index we conclude, by assumption, that there are no integers $j = 0$ and $0 \leq j_1 < \ldots < j_m$ such that the co-vectors

$$f_{j_1} A_{j_1} \ldots A_1, \ldots, f_{j_m} A_{j_m} \ldots A_1$$

form a co-basis of V_0. This means, by lemma 2, that C belongs to the co-dimension $s - m + 2$ subset defined in that lemma. So we only have to consider the case that there is a blocking pair $0, i$.

The existence of a blocking pair $0, i$ means that $(A_1, \ldots, A_{i+1}, f_0, \ldots, f_i)$ belong to a subset of co-dimension $i + 2$. The remaining $(A_{i+2}, \ldots, A_s, f_{i+1}, \ldots, f_s)$ form a linear system of length $s - i$. If $s - i$ is smaller than m, then the co-dimension, corresponding to the blocking pair $0, i$, is already at least equal to $s - m + 2$ and we are finished. If $s - i \geq m$ we use induction on the length of the linear system. If the proposition holds for linear systems of length smaller than $s + 1$, then we conclude that we have already '$i + 2$ co-dimensions' for the blocking pair and '$s - m - i + 1$ co-dimensions' for the remaining linear system of length $s - i$. This gives in total even $s - m + 3$ co-dimensions.

So finally we only have to prove the proposition for the smallest length, which is $s + 1 = m$. In this case the proposition is a direct consequence of the lemma's 1 and 2. $\square$

This concludes the discussion of general linear systems and we return to:

4.4.2 Continuation of the genericity proof

We have to show that for generic (φ, f), for $k \geq 2m + 1$, and for each pair $(x, \tilde{x}) \in M \times M$ the property $\mathcal{P}_k$ holds. We may and do assume, without loss of generality, that $k = 2m + 1$. Since we know that the property $\mathcal{P}$ is persistent, we only have to show that we can change (φ, f), by a perturbation which is arbitrarily small in the C^1 sense, so that after the perturbation the property $\mathcal{P}_{2m+1}$ holds for all the pairs in $M \times M$. As we remarked before we may ignore periodic orbits with period at most $4m$.

We recall the definition of the meeting number $j(x, \tilde{x})$: it is the smallest integer such that $\varphi^{j(x,\tilde{x})}(x) = \varphi^{j(x,\tilde{x})}(\tilde{x})$; if no such integer exists, then $j(x, \tilde{x}) = \infty$. We denote by $\mathcal{J}^l$ the subset of $M \times M$ of the pairs $(x, \tilde{x})$ with $j(x, \tilde{x}) = l$. It is clear that $\mathcal{J}^0 = \Delta \subset M \times M$, which is an m-dimensional submanifold. For each $(x, x) \in \Delta$ we consider the linear system of length $2m + 1$ on the vector spaces $T_x(M), T_{\varphi(x)}(M), \ldots, T_{\varphi^{2m}(x)}(M)$ with linear maps $\mathrm{d}\varphi(\varphi^i(x))$, $i = 0, \ldots, 2m-1$, and linear functions $\mathrm{d}f(\varphi^i(x))$, $i = 0, \ldots, 2m$. We note that all the points $x, \varphi(x), \ldots, \varphi^{2m}(x)$ are pairwise different (due to the proposition in 4.3 we may disregard periodic points of periods up to $4m$ and their pre-images up to order $4m$, so that orbit segments of length $2m + 1$ do not 'revisit' points). So we can perturb both $\mathrm{d}f$ and $\mathrm{d}\varphi$ in all these points independently. Hence, for any property of linear systems of length $2m + 1$ with co-dimension at least

$m+1$ it is generic that it does not occur for any pair $(x, x) \in \Delta$. By the above proposition, this means that generically the property $\mathcal{P}_{2m+1}$ holds for all pairs $(x, x) \in \Delta$ because for each such x there should be some $j \leq 2m+1$ such that $\varphi^j(x)$ is $(2m-j+1)$-embedding. (In fact, the proposition in 4.4.1, applied to linear systems of length $2m+1$, yields a co-dimension $m+2$ which is one more than needed.) By persistence, the property $\mathcal{P}_{2m+1}$ also holds in a neighbourhood of Δ.

Now we consider $\mathcal{J}^1$ and will show that also there generically the property $\mathcal{P}_{2m+1}$ holds. We will show this in such a way that it will be clear how to continue by induction to $\mathcal{J}^2, \ldots, \mathcal{J}^{2m}$; after this the situation becomes different. As we observed above we can restrict ourselves to a (compact) complement of a neighbourhood of Δ. We first observe that generically the map $(\varphi, \varphi) : M \times M \to M \times M$, restricted to the complement of a neighbourhood of Δ, is transversal with respect to Δ. This means that generically $\mathcal{J}^1$ is an m-dimensional submanifold, except possibly in a small neighbourhood of $\Delta = \mathcal{J}^1$ where the propery $\mathcal{P}_{2m+1}$ holds anyway. From now on we assume $\mathcal{J}^1$ is such a submanifold.

Let $(x, \tilde{x}) \in \mathcal{J}^1$. If $f(x) \neq f(\tilde{x})$, the the property $\mathcal{P}_{2m+1}$ holds for this pair. If $f(x) = f(\tilde{x})$, we consider the linear system of length $2m$ on the tangent spaces at $\varphi(x) = \varphi(\tilde{x}), \ldots, \varphi^{2m}(x) = \varphi^{2m}(\tilde{x})$. In order that property $\mathcal{P}_{2m+1}$ does not hold in $(x, \tilde{x})$, conditions with a total co-dimension of at least $m+2$ must be satisfied: 1 co-dimension for $f(x) = f(\tilde{x})$ and $m+1$ co-dimensions for the linear system of length m to be exceptional in the sense of the proposition in section 4.4.1.

As we observed above we can continue by induction till $\mathcal{J}^{2m}$. Now we have the property $\mathcal{P}_{2m+1}$ holding on all of $\mathcal{J} = \bigcup_{i=0}^{2m} \mathcal{J}^i$, and hence also on a neighbourhood of this set. For each pair $(x, \tilde{x})$ outside $\mathcal{J}$ we have that $\varphi^i(x) \neq \varphi^i(\tilde{x})$ for $i = 0, \ldots, 2m$. So now we can apply the proposition in 4.2.2 with $k = 2m+1$ and obtain that generically for all these pairs $(x, \tilde{x})$ the corresponding reconstruction vectors are unequal: $\mathrm{Rec}_{2m+1}(x) \neq \mathrm{Rec}_{2m+1}(\tilde{x})$.

This completes the proof that generically all pairs $(x, \tilde{x})$ have the property $\mathcal{P}_{2m+1}$.

5 Reconstruction of endomorphisms

In this section we assume that $\varphi : M \to M$ is an endomorphism on the closed m-dimensional manifold M and that $f : M \to \mathbb{R}$ is a read out function, both at least C^1 and such that, for some k, the k-reconstruction condition holds, i.e. that for all pairs $(x, \tilde{x}) \in M \times M$ the property $\mathcal{P}_k$ holds. We prove here that there is a differentiable map $\pi_k : \mathcal{X}_k \to M$ such that $\pi_k Rec_k = \varphi^{k-1}$. So this holds even for $k \leq 2m$.

Lemma 1. *Under the above hypothesis there is a unique map $\pi_k : \mathcal{X}_k \to M$ such that $\varphi^{k-1} = \pi_k \operatorname{Rec}_k$.* □

Proof. The uniqueness, in case of existence, is clear. For each $\xi \in \mathcal{X}_k$ there is an $x \in M$ such that $\operatorname{Rec}_k(x) = \xi$. So then $\pi_k(\xi)$ has to be equal to $\varphi^{k-1}(x)$. The only problem with existence could be that there are different $x \neq \tilde{x}$ with the same reconstruction vectors $\operatorname{Rec}_k(x) = \operatorname{Rec}_k(\tilde{x})$. Due to the property $\mathcal{P}_k$ however, in that case we have $\varphi^{k-1}(x) = \varphi^{k-1}(\tilde{x})$ so that there is no ambiguity in the definition of $\pi_k(\xi)$. □

Lemma 2. *The map π_k is differentiable.* □

Proof. For a reconstruction vector ξ we construct a differentiable extension of π_k to a neighbourhood of ξ in $\mathbb{R}^k$. We make use of the second and equivalent formulation of the k-reconstruction condition, see the lemma in section 3. It implies the existence of an integer $j(\xi)$ and a point $p(\xi)$ such that for each $x \in \operatorname{Rec}^{-1}(\xi)$, $\varphi^{j(\xi)}(x) = p(\xi)$ and such that $p(\xi)$ is $(k - j(\xi))$-embedding. This means that there are integers $0 \leq j_1 < \ldots < j_m < (k - j(\xi))$ such that

$$f\varphi^{j_1}, \ldots, f\varphi^{j_m}$$

form a local coordinate system in a neighbourhood U of $p(\xi)$. Then we define a smooth map λ from a neighbourhood V of ξ in $\mathbb{R}^k$ to U such that for $\eta = (\eta_1, \ldots, \eta_k) \in V$, the values of $f\varphi^{j_i}(\lambda(\eta))$ agree with the coordinates $\eta_{j(\xi)+j_i}$ for $i = 1, \ldots, m$. Then it is clear that for all points

$$\tilde{x} \in W = \varphi^{-j(\xi)}(U) \bigcap \operatorname{Rec}_k^{-1}(V) \quad \text{we have} \quad \varphi^{j(\xi)}(\tilde{x}) = \lambda \operatorname{Rec}_k(\tilde{x}) .$$

So $\Lambda = \varphi^{k-j(\xi)-1}\lambda$, if necessary after restricting to a subset of V, is a smooth extension of π_k. In order to see that Λ can indeed be defined on a neighbourhood of ξ, we observe that $M \setminus W$ and hence $\operatorname{Rec}_k(M \setminus W)$ is compact and that the latter set does not contain ξ. We can take $V \setminus \operatorname{Rec}_k(M \setminus W)$ as the domain of Λ.

This completes the proof of the lemma and also of the theorem 2 as stated in the introduction. □

6 Applications: deterministic structure and the estimation of dimensions and entropies

Also in this section we assume that (φ, f) is a dynamical system on a compact manifold M and that for some k the k-reconstruction condition is satisfied.

6.1 The dynamics on $\mathcal{X}_k$

The set of k-dimensional reconstruction vectors $\mathcal{X}_k$ can also be considered as the state space of a dynamical system with map Φ (in this case we have a dynamical system without read out function) in such a way that $\mathrm{Rec}_k\varphi = \Phi\mathrm{Rec}_k$. The argument is simple:

If $w \in \mathcal{X}_k$ and $x \in M$ with $\mathrm{Rec}_k(x) = w$ then we want $\Phi(w)$ to be $\mathrm{Rec}_k(\varphi(x))$. The problem is that the element x may not be uniquely determined by w. However, in any case, the first $k-1$ coordinates of $\Phi(w)$ should be equal to the last $k-1$ coordinates of w. The last coordinate of $\Phi(w)$ should be equal to $f\varphi(\varphi^{k-1}(x))$. Though x may not be uniquely determined by w, $\varphi^{k-1}(x) = \pi_k(w)$ is uniquely determined by w; π_k is the map discussed in the previous section. So the map Φ has the explicit form

$$\Phi(w) = (w_2, \dots, w_k, f\varphi\pi_k(w)), \text{ where } w = (w_1, \dots, w_k).$$

We have derived the dynamics in $\mathcal{X}_k$ from the dynamics in M defined by φ. This can be done differently: if the k-reconstruction condition holds, then the dynamics on $\mathcal{X}_k$ can be deduced from $\mathcal{X}_{k+1}$. This is done in the following way: if $w = (w_1, \dots, w_k) \in \mathcal{X}_k$, then $\Phi(w)$ should have the form $(w_2, \dots, w_k, h(w_1, \dots, w_k))$ and it should be such that $(w_1, \dots, w_k, h(w_1, \dots, w_k))$ belongs to $\mathcal{X}_{k+1}$. From the above considerations it follows that this determines $h(w_1, \dots, w_k)$ uniquely.

So, from the k-reconstruction condition it follows that $\mathcal{X}_{k+1}$ completely determines which (finite or infinite) time series the dynamical system (φ, f) can produce.

6.2 Dimensions and entropies

We first recall the definitions of the correlation dimensions and entropies. We assume that we have a dynamical system with state space K and map ψ (also here we don't have a read out function). We assume K to be a compact metric space with metric d and ψ to be continuous. We also assume that there is a ψ-invariant Borel probability measure μ.

The *dimensions* $D_q(\mu)$, with $q \neq 1$ are defined in terms of the metric d and the measure μ, and are independent of the map ψ:

$$D_q(\mu) = \lim_{\varepsilon\to 0} \frac{\log\int(\mu(B(x,\varepsilon)))^{q-1}\mathrm{d}\mu}{(q-1)\log\varepsilon},$$

where $B(x,\varepsilon)$ denotes the ε-neighbourhood of x. The limit may not exist (in the sense that lim inf and lim sup are different) in which case one can define a lower and an upper dimension, or the limit may diverge to infinity, in which case the corresponding dimension is ∞. For $q = 1$ there is a different definition which is suggested by continuity considerations. Also for that dimension the arguments below are still valid, but we omit the details.

For the definition of the *entropies* $H_q(\psi, \mu)$ we need metrics which are derived from d and ψ: $d_i(x, y) = \max_{j=0,\dots,i-1} d(\psi^j(x), \psi^j(y))$; $B_i(x, \varepsilon)$ is the ε-neighbourhood of x with respect to the metric d_i. For $q \neq 1$ we define:

$$H_q(\psi, \mu) = \lim_{\varepsilon\to 0} \lim_{n\to\infty} \frac{\log \int (\mu(B_n(x, \varepsilon)))^{q-1} \mathrm{d}\mu}{(1-q)n}.$$

For these definitions, see also [P], [TV], and [V].

Next we introduce the notion of a *morphism between two dynamical systems*. So we let (K, ψ, μ) and (K', ψ', μ') denote two dynamical systems as above, i.e. K, K' are compact metric spaces with metrics d and d', ψ and ψ' are continuous maps on K and K' respectively, and μ and μ' are invariant Borel probability measures. A map $g : K \to K'$ is called a morphism between these dynamical systems if:

– g is continuous and even Lipschitz in the sense that for some constant C and any $x, y \in K$ we have $d'(g(x), g(y)) \leq Cd(x, y)$;

– $\psi' g = g\psi$;

– $\mu' = g_\star(\mu)$, i.e. for measurable $U' \subset K'$, we have $\mu(g^{-1}(U') = \mu'(U')$.

Proposition. *If $g : K \to K'$ is a morphism between two dynamical systems as above, then we have*

$$D_q(\mu) \geq D_q(\mu') \text{ and } H_q(\psi, \mu) \geq H_q(\psi', \mu'). \qquad \square$$

Proof. We define the functions $h_n(x, \varepsilon) = \mu(B_n(x, \varepsilon))$, see the above definitions of dimensions and entropies; the corresponding functions for the second dynamical system (on K') are denoted by h'_n. If C denotes a Lipschitz constant for g, i.e. if $d'(g(x), g(y)) \leq Cd(x, y)$ for all $x, y \in K$, then clearly $h_n(x, \varepsilon) \leq h'_n(g(x), C\varepsilon)$. This means that, for $q > 1$ we also have that

$$\int (h_n(x, \varepsilon))^{q-1} \mathrm{d}\mu(x) \leq \int (h'_n(x', C\varepsilon))^{q-1} \mathrm{d}\mu'(x').$$

So, apart from the fact that ε changed to $C\varepsilon$, this implies that $D_q(\mu) \geq D_q(\mu')$ and $H_q(\psi, \mu) \geq H_q(\psi', \mu')$. This factor C in front of the epsilon disappears in the limits defining the dimensions en entropies. This completes the proof for $q > 1$. For $q < 1$, the above inequality between integrals reverses, but also the factor $(1 - q)$ in the denominators in the definitions of the dimensions and entropies changes sign, so the outcome is the same inequality for $q < 1$. We did not spell out the definitions of D_1 and H_1, but the same arguments also lead to the same result in that case. $\square$

From the above proposition on (abstract) morphisms between dynamical systems, we obtain, by applying then to the 'morphisms' $\mathrm{Rec}_k : M \to \mathcal{X}_k$ and $\pi_k : \mathcal{X}_k \to M$ the following:

Theorem. *Let (φ, f) define a smooth dynamical system (with read out function) on the closed manifold M, which satisfies the k-reconstruction condition. Let $\Phi : \mathcal{X}_k \to \mathcal{X}_k$ define the corresponding dynamics (without read out function) on $\mathcal{X}_k$. Let μ be some φ-invariant Borel probability measure on M and let $\mu' = (\mathrm{Rec}_k)_\star(\mu)$. Then Rec_k and π_k are morphisms in the above sense between the dynamical systems (M, φ, μ) and $(\mathcal{X}_k, \Phi, \mu')$ and hence the dimensions and entropies of these systems are the same.* □

Estimates of dimensions and entropies. The estimates of dimension and entropy (for convenience we restrict to the correlation dimension D_2 and entropy H_2) of a dynamical system from time series are based on the following considerations. We consider an orbit $x_0 = \bar{x}, x_1 = \varphi(\bar{x}), \dots$ of the dynamical system on M defined by φ and assume that this orbit defines a Borel probability measure μ on M so that for each continuous function $g : M \to \mathbb{R}$ we have

$$\int_M g \mathrm{d}\mu = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} g(x_i).$$

(This is called the natural measure defined by the orbit. For the existence or nonexistence of such measures see [RU].) Such a measure μ is φ-invariant. Next we assume that the k-reconstruction condition is satisfied for (φ, f). Then the maps $\mathrm{Rec}_k : M \to \mathcal{X}_k$ and $\pi_k : \mathcal{X}_k \to M$ are morphisms between dynamical systems (the dynamics on $\mathcal{X}_k$ being defined by the map Φ as introduced in section 6.1). We want to estimate, from the time series corresponding to the orbit $x_0, x_1, \dots,$ which is $y_0 = f(x_0), y_1 = f(x_1), \dots,$ the correlation dimension and entropy of φ with respect to the measure μ.

From the above theorem it follows that this is the same as estimating these quantities of Φ with respect to the measure $\mu' = (\mathrm{Rec}_k)_\star(\mu)$. This can be done in terms of the correlation integrals $C_n(\varepsilon)$, which are defined as the probability that two 'random' reconstruction vectors of dimension n are coordinate wise within distance ε. This quantity is estimated by counting the number $\mathcal{N}_{N,n}(\varepsilon)$ of pairs (i, j) with $0 \le i < j \le N$ such that $| y_i - y_j | < \varepsilon, \dots, | y_{i+n-1} - y_{j+n-1} | < \varepsilon$. Then

$$C_n(\varepsilon) = \lim_{N\to\infty} \frac{\mathcal{N}_{N,n}(\varepsilon)}{(N(N+1)/2)}.$$

So the quantities

$$\frac{\mathcal{N}_{N,n}(\varepsilon)}{(N(N+1)/2)},$$

for N sufficiently big, are estimates for $C_n(\varepsilon)$.

The relation between these quantities and the definitions of the dimension and entropy (with $q = 2$) for Φ with respect to μ' is

$$\int_{\mathcal{X}_k} \mu'(B_n(x, \varepsilon)\mathrm{d}\mu'(x) = C_{n+k}(\varepsilon)$$

where the distance function d' on $\mathcal{X}_k$ is given by $d'((z_1, \ldots, z_k), (z'_1, \ldots, z'_k)) = \max_{1 \leq i \leq k} \mid z_i - z'_i \mid$. This means that we find for the correlation dimension and entropy the usual expressions

$$D_2 = \lim_{\varepsilon \to 0} \frac{\ln(C_n(\varepsilon))}{\ln(\varepsilon)} \text{ for } n \geq k$$

and

$$H_2 = \lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{-\ln(C_n(\varepsilon))}{n}.$$

This justifies the use of the standard algorithms, see e.g. [GPa] and [T3], for estimating the dimension and entropy of the natural measure of a given orbit also for endomorphisms.

References

[A] D. Aeyels, Generic observability of differentiable systems, SIAM J. Control and Optimisation, **19** (1981), 595–603.

[B] Th. Bröcker, Differentiable germs and catastrophes, Camb. Univ. Press, (1975).

[GPa] P. Grassberger, I. Procaccia, Measuring the strangeness of strange attractors, Physica **9D** (1983a), 189–208.

[GPb] P. Grassberger, I. Procaccia, Characterisation of strange attractors, Phys. Rev. Lett. **50** (1983b), 346–349.

[GWPL] C.G. Gibson, K. Wirthmüller, A.A. du Plessis, E.J.N. Looijenga, Topological stability of smooth mappings, Lecture Notes in Mathematics **552**, Springer-Verlag, (1976).

[H] M.W. Hirsch, Differential topology, Springer-Verlag, (1976).

[KS] H. Kantz, T. Schreiber, Nonlinear time series analysis, Cam. Univ. Press, (1997).

[L] S. Lojasiewiecz, Ensembles semi-analytiques, I.H.E.S., (1965).

[O] J.C. Oxtoby, Measure and category, Springer-Verlag, (1971).

[P] Ya.B. Pessin, Dimension theory in dynamical systems, University of Chicago Press, (1997).

[RO] C. Robinson, Dynamical systems, CRC Press, (1995).

[RU] D. Ruelle, Historical behaviour in smooth dynamical systems, in Global analysis of dynamical systems, ed.: H.W. Broer, B. Krauskopf and G. Vegter, IoP, (2001), 63–66.

[S] S. Smale, Differentiable dynamical systems, Bull. A.M.S. **73** (1967), 747–817.

[SYC] T. Sauer, J. A. Yorke, M. Casdagli, Embedology, J. Statistical Physics **65** (3/4) (1991), 579–616.

[T6] R. Thom, Un lemma sur les applications différentiables, Bol. Soc. Mat. Mex. **2** (1956), 59–71.

[T9] R. Thom, Ensembles et morphismes stratifiés, Bull. A.M.S., **75** (1969), 240–284.

[T1] F. Takens, Detecting strange attractors in turbulence, in Dynamical systems and turbulence, Warwick 1980, ed. D. A. Rand, L-S. Young, LNM **898**, Springer-Verlag, (1981).

[T3] F. Takens, Invariants related to dimension and entropy, in: Atas do 13^o colóquio bras. de mat. 1981, IMPA, Rio de Janeiro, (1983).

[TV] F. Takens, E. Verbitskiy, Generalised entropies: Rényi and correlation integral approach, Nonlinearity **11** (1998), 771–782.

[V] E.A. Verbitskiy, Generalised entropies in dynamical systems, Ph.D thesis, University of Groningen (2000).

[W5] H. Whitney, Differentiable manifolds, Ann. of Math. **37** (1935), 645–680.

[W4] H. Whitney, The self-intersections of a smooth n-manifold in $2n$-space and $(2n-1)$-space, Ann. of Math. **45** (1944), 220–293.

Floris Takens
Mathematisch Institut RUG
PObox 800, 9700 AV Groningen
NETHERLANDS

E-mail: f.takens@math.rug.nl