J. Nonlinear Sci. Vol. 13: pp. 519–577 (2003)
DOI: 10.1007/s00332-003-0534-4

Journal of Nonlinear Science

Delay Embeddings for Forced Systems. II. Stochastic Forcing

J. Stark,[1] D. S. Broomhead,[2] M. E. Davies,[3] and J. Huke[2]

[1] Department of Mathematics, Imperial College London, London, SW7 2AZ, United Kingdom
e-mail: j.stark@imperial.ac.uk

[2] Department of Mathematics, University of Manchester Institute of Science and Technology, P.O. Box 88, Manchester, M60 1QD, United Kingdom

[3] Department of Electronic Engineering, Queen Mary, University of London, Mile End Road, London E1 4NS, United Kingdom

Received July 22, 2002; accepted September 16, 2003
Online publication December 3, 2003
Communicated by T. Schreiber

Summary. Takens' Embedding Theorem forms the basis of virtually all approaches to the analysis of time series generated by nonlinear deterministic dynamical systems. It typically allows us to reconstruct an unknown dynamical system which gave rise to a given observed scalar time series simply by constructing a new state space out of successive values of the time series. This provides the theoretical foundation for many popular techniques, including those for the measurement of fractal dimensions and Liapunov exponents, for the prediction of future behaviour, for noise reduction and signal separation, and most recently for control and targeting. Current versions of Takens' Theorem assume that the underlying system is autonomous (and noise-free). Unfortunately this is not the case for many real systems. In a previous paper, one of us showed how to extend Takens' Theorem to deterministically forced systems. Here, we use similar techniques to prove a number of delay embedding theorems for arbitrarily and stochastically forced systems. As a special case, we obtain embedding results for Iterated Functions Systems, and we also briefly consider noisy observations.

1. Introduction

This paper continues the work begun in [Stark, 1999] where one of us developed techniques for proving extensions of the Takens' Embedding Theorem to forced dynamical systems. In that paper, these methods were applied to the case of forcing by a finite-dimensional deterministic system. In many applications the assumption that the forcing is deterministic is not a reasonable one, and the aim of this paper is therefore to extend these results to far more general forcing processes.

Recall that Takens' Embedding Theorem ([Takens, 1980]; see also [Aeyels, 1981] and [Sauer et al., 1991]) provides the theoretical foundation for the analysis of time series generated by nonlinear deterministic dynamical systems. Informally, it says that if we take a scalar observable φ of the state x of a deterministic dynamical system, then *typically* we can reconstruct a copy of the original system by considering blocks $(\varphi(x_t), \varphi(x_{t+\tau}), \varphi(x_{t+2\tau}), \ldots, \varphi(x_{t+(d-1)\tau}))$ of d successive observations of φ, for d sufficiently large. Here x_t is the state of the system at time t, and $\tau > 0$ is some sampling interval. Since x_t will usually be unknown, whilst $\varphi(x_t)$ is a quantity we can measure in practice, this result has stimulated a vast range of applications in fields ranging from fluid dynamics through electrical engineering to biology, medicine, and economics (for a good overview see e.g. [Ott et al., 1994]). One might even say that this one theorem has given rise to a virtually new branch of nonlinear dynamics, often informally called *chaotic time series analysis*.

However, for Takens' Theorem to be valid, we need to assume both that the dynamics is deterministic (i.e. that there is some mapping f such that $x_{t+\tau} = f(x_t)$), and that both the dynamics and the observations are autonomous (so that f and φ depend on x only). Note that by rescaling time, if necessary, we may assume that $\tau = 1$, and hence restrict t to integer values. In [Stark, 1999] we extended Takens' Theorem to the nonautonomous case where f is also a function of some other variable y_t, where y_t itself is generated by a deterministic system, so that $y_{t+1} = g(y_t)$ for some g.

Here we turn to systems driven by far more general processes. It turns out that the same approach can encompass a number of different cases. In particular the theorems proved below apply to a wide class of stochastic dynamical systems (which we can think of as deterministic systems driven by some stochastic process), to input-output systems, and to irregularly sampled systems. Input-output systems have already been considered in this context by [Casdagli, 1992], and our results here in essence prove his conjecture on embedding such systems.

Broadly speaking, our approach results in an embedding framework where both the dynamical systems and the delay map are indexed by the realization of the forcing process. This leads to results similar to the Bundle Delay Embedding Theorem (Theorem 3.2) of [Stark, 1999]; and indeed the proof of the main Stochastic Takens' Embedding Theorem here (Theorem 2.3 below) closely parallels that of Theorem 3.2 of [Stark, 1999]. We also consider a number of variations of this main theorem, including the case of Iterated Function Systems (e.g. [Norman, 1968]; [Barnsley, 1988]) and of noisy observations. It is perhaps interesting to note that the latter, at least in the case where the forcing process takes on continuous values, is by far the easiest case to prove, and amounts to little more than the classical Whitney Embedding Theorem (e.g. [Hirsch, 1976]).

We assume familiarity with the standard Takens' framework, and use the notation of [Stark, 1999]. Most of the transversality techniques employed here closely follow those developed in that paper, and we shall make use of a number of technical results and calculations derived there. We begin by describing a general framework for treating arbitrarily forced systems, and by stating the principal theorems proved in this paper.

2. Delay Embedding for Arbitrary Forcing

A convenient formalism for incorporating arbitrary forcing into a dynamical system is that of *random dynamical systems* (e.g. [Kifer, 1988]; [Arnold, 1998]). This encompasses both a wide class of noisy systems, as well as input-output systems in the terminology of [Casdagli, 1992], irregularly sampled systems, and others. In the case of discrete time systems, which is the situation that we treat here, we assume that the forcing at time $i \in \mathbb{Z}$ is specified by a variable ω_i, drawn from some appropriate space N. In the case of noisy systems, ω_i is chosen at random, with respect to some probability measure μ on N. The state of the system at time $i \in \mathbb{Z}$ is denoted by $x_i \in M$ and evolves according to

$$x_{i+1} = f(x_i, \omega_i). \tag{2.1}$$

As is usual in the standard Takens framework, we assume that M is a smooth compact manifold. The case of M noncompact can be treated (e.g. [Takens, 1980]; [Huke, 1993]), but will not be considered here further. If we think of ω_i as a parameter, then we can interpret (2.1) as a standard dynamical system with noise or forcing on the parameters. It can also be helpful to write $f_{\omega_i}(x_i)$ instead of $f(x_i, \omega_i)$. This suggests the interpretation that instead of applying the same map f every time, we choose a different map f_{ω_i} at each time step. The case of a single deterministic system f subject to additive (dynamical) noise can be included in this formalism by setting $f_{\omega_i}(x_i) = f(x_i) + \omega_i$.

In the most general case (assuming M compact), N can be taken to be the space of all maps on M [Kifer, 1988]. At the other extreme, N may be a discrete space consisting of a finite number of points, so that f_{ω_i} is chosen from a finite set of maps. This leads to the well-known example of so-called iterated function systems (e.g. [Norman, 1968]; [Barnsley, 1988]). In this paper we shall choose N somewhere between these two extremes, and make the standing assumption that N is a compact manifold.

2.1. Shift Spaces

The usual approach to nonautonomous systems of the form (2.1) is to enlarge the state space sufficiently to make the system autonomous. This is well known in the case of a periodically forced differential equation which can be transformed into an autonomous system by the addition of a dummy variable to represent time. In the context of arbitrary forcing, this is most easily accomplished through the use of an appropriate shift space (e.g. [Kifer, 1988]; [Arnold, 1998]). Thus, let $\Sigma = N^{\mathbb{Z}}$ be the space of bi-infinite sequences $\omega = (\ldots, \omega_{-1}, \omega_0, \omega_1, \ldots)$ of elements in N with the product topology. Since we assume that N is compact, the Tychonoff theorem implies that Σ is also compact. Let $\sigma\colon \Sigma \to \Sigma$ be the usual shift map

$$[\sigma(\omega)]_i = \omega_{i+1},$$

where $\omega_i \in N$ is the ith component of $\omega \in \Sigma$. Then the evolution of x_i given by (2.1) can be represented by the skew product $T\colon M \times \Sigma \to M \times \Sigma$ (see Figure 1) defined by

$$T(x, \omega) = (f(x, \omega_0), \sigma(\omega)). \tag{2.2}$$

Since the space Σ contains all possible sequences of elements in N, this gives us a very general model of systems driven by arbitrary sequences. Furthermore, if one wants to restrict interest to a particular class of input sequences, one can replace Σ by any shift-invariant closed subset. If in addition we have some probability measure μ on N, then the corresponding product measure μ_∞ on Σ is shift invariant and hence (σ, μ_∞)

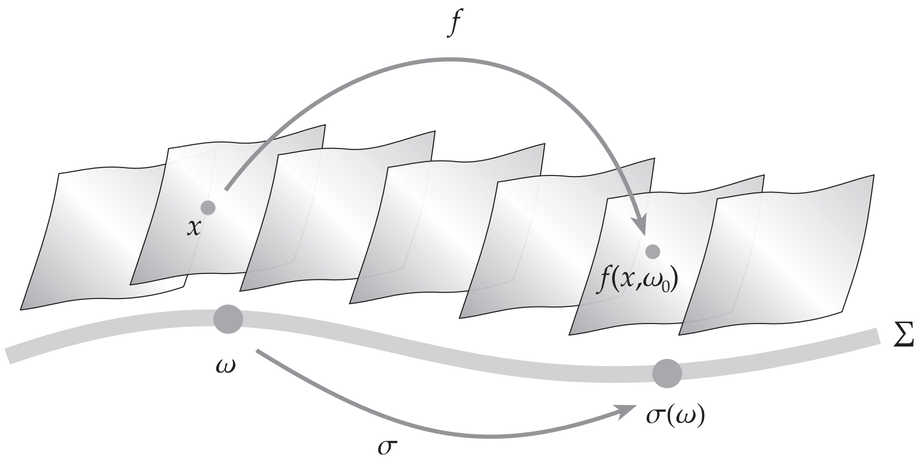

Fig. 1. Graphical representation of a random dynamical system, from [Stark, 2001].

gives rise to a (Bernoulli) stochastic process. We can also consider general σ-invariant measures μ_Σ to take account of correlations in the choice of successive ω_i (so that for instance ω is a Markov process). The restriction to σ-invariant measures corresponds to a stationarity condition on the corresponding random process. Additionally, f could be taken to be a general function $f\colon M \times \Sigma \to M$, rather than just depending on a single component ω_0 of ω. We shall not do so here, though as we shall soon see, the delay map will depend on more than a single element of ω.

If we dispense with the measure μ, the same formalism can also be used to model a deterministic system driven by an arbitrary input sequence ω. This arises frequently in communications systems where the sequence ω would represent the information being transmitted (e.g. [Broomhead et al., 1999]). Another application is to irregularly sampled time series where ω_i denotes the time between sample i and $i+1$ (see [Martin, 1998], and Example 3.5 of [Stark, 1999]).

2.2. Conjugacies for Random Dynamical Systems

The crucial question we need to address in order to develop an analogue of Takens' Theorem for random dynamical systems is what do we mean by a *delay reconstruction* of T? It is well known (e.g. [Takens, 1980]; [Stark, 1999]; [Stark, 2001]) that the fundamental property required of a reconstruction is that the reconstructed system (which we shall call F, in line with [Stark, 1999]; [Stark, 2001]) should be equivalent to the original dynamical system f under a coordinate change. In the standard Takens embedding framework therefore, we have $F = \Phi \circ f \circ \Phi^{-1}$, where Φ is the delay map.

We thus need to determine what it means for two random dynamical systems T and T' to be equivalent in this way. The most general concept is simply to require $T' = H \circ T \circ H^{-1}$ for some (invertible) coordinate change $H\colon M \times \Sigma \to M \times \Sigma$. It may not be possible to arrange for $T' = H \circ T \circ H^{-1}$ to hold for all $\omega \in \Sigma$, and hence we may either require this for only μ_Σ-almost every ω in the probabilistic setting, or for only generic ω in a topological setting. It may also be reasonable to require H to be "causal," i.e. to depend only on the past components of ω.

In the context of delay embedding, however, it turns out to be convenient to place some further restrictions on H. The space $M \times \Sigma$ will in general be infinite dimensional, and hence we have little hope of reconstructing it in a finite-dimensional delay space $\mathbb{R}^d$. Furthermore φ is a function of M only and hence attempting to reconstruct Σ using φ seems foolhardy (though note that in the case of deterministic forcing, this is in fact possible [Stark, 1999]). We therefore want to restrict ourselves to conjugacies which correspond to reconstructing only M and leaving Σ untouched. A reasonable interpretation of this is to require that the Σ component of H is the identity (Figure 2). In other words, we only consider conjugacies of the form $H = (h, Id)$ for some map h: $M \times \Sigma \to M$. If H is to be invertible for μ_Σ-almost every ω, or for generic ω, then $h_\omega = h(\bullet, \omega)$: $M \to M$ has to be invertible for μ_Σ-almost every ω, or for generic ω. We shall refer to coordinate changes of this form as *bundle conjugacies* (using the term loosely for any of h_ω, h, and H itself). Coordinate changes of this form are common in the theory of random dynamical systems (e.g. [Arnold, 1998]).

Given a skew product (2.2), it is convenient to define $\tilde{f}$: $M \times \Sigma \to M$ by $\tilde{f}(x, w) = f(x, \omega_0)$ so that $T = (\tilde{f}, \sigma)$. If similarly, $T' = (\tilde{f}', \sigma)$, then $T' = H \circ T \circ H^{-1}$ is equivalent to $(\tilde{f}', \sigma) = (h, Id) \circ (\tilde{f}, \sigma) \circ (h, Id)^{-1}$. If we denote $\tilde{f}_\omega = \tilde{f}(\bullet, \omega)$: $M \to M$, we have $(h, Id) \circ (\tilde{f}, \sigma) \circ (h, Id)^{-1}(\bullet, \omega) = (h, Id) \circ (\tilde{f}_\omega \circ h_\omega^{-1}, \sigma(\omega)) = (h_{\sigma(\omega)} \circ \tilde{f}_\omega \circ h_\omega^{-1}, \sigma(\omega))$. Hence

$$\tilde{f}'_\omega = h_{\sigma(\omega)} \circ \tilde{f}_\omega \circ h_\omega^{-1}, \tag{2.3}$$

where as usual $\tilde{f}'_\omega = \tilde{f}'(\bullet, \omega)$: $M \to M$. Note that in [Stark et al., 1997] and [Stark, 1999] we abused notation to write f_ω instead of $\tilde{f}_\omega$. Observe the similarity between (2.3) and the deterministic conjugacy $f' = h \circ f \circ h^{-1}$. Essentially all we do in the random case is index both the dynamics and the coordinate change with ω. The only slightly delicate point is the σ appearing in $h_{\sigma(\omega)}$. The reason for this is that by the time we come to apply h in equation (2.3) we have carried out one time step of the dynamics, and hence ω has moved to $\sigma(\omega)$ (Figure 2).

In applications to delay embedding we need to consider one further issue, namely that the domain and range of H will typically not be the same space. Thus, recall that in the standard Takens' framework the delay map Φ is a map Φ: $M \to \mathbb{R}^d$. The main content of Takens' Theorem is that generically this is an embedding (a diffeomorphism onto its image $\Phi(M)$) and hence we can define the coordinate change $F = \Phi \circ f \circ \Phi^{-1}$ on $\Phi(M)$. In the present context Φ will depend on ω and so for each ω can be regarded as a map Φ_ω: $M \times \Sigma \to \mathbb{R}^d$ (see section 2.3 below). Informally our concept of a Stochastic Takens' Theorem is to require Φ to be a *bundle embedding*. By this we mean that Φ_ω should be an embedding (a diffeomorphism onto its image $\Phi_\omega(M)$) for typical ω, i.e. for μ_Σ-almost every ω, or for generic ω. Note that in general the image $\Phi_\omega(M)$ will be different for each ω (but all these images will be diffeomorphic to M). The range of the map $H = (\Phi, Id)$ will thus not be $M \times \Sigma$ as above, but rather the reconstruction space $\mathbb{R}^d \times \Sigma$. When Φ is a bundle embedding, the coordinate change $T' = H \circ T \circ H^{-1}$ is defined on $H(M \times \Sigma)$. This space cannot typically be written in the form $M' \times \Sigma$ for some appropriate M', but it is bundle diffeomorphic to $M \times \Sigma$.

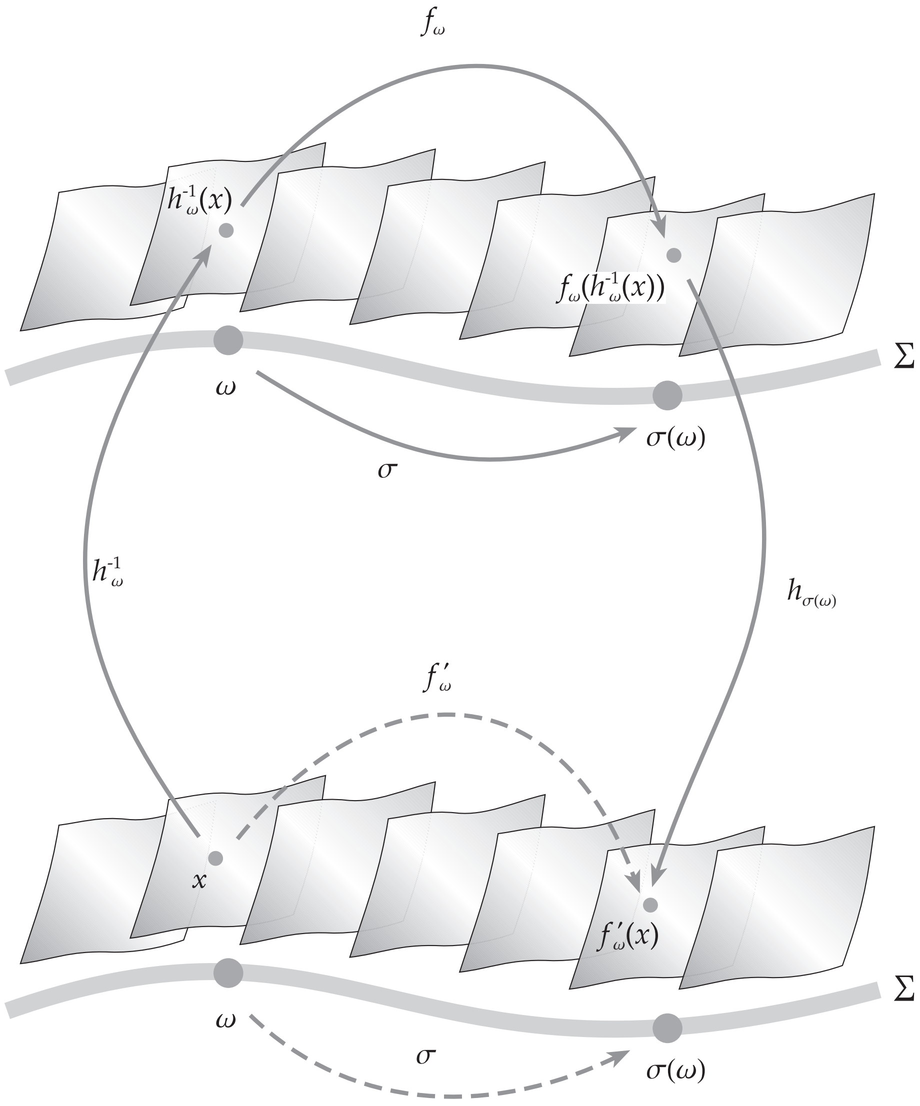

Fig. 2. Coordinate change for random dynamical systems, from [Stark, 2001].

2.3. The Stochastic Takens' Theorem

We shall now define the delay map more precisely, and give several versions of the Stochastic Takens' Theorem. Suppose that we observe the skew product (2.2) using a measurement function $\varphi\colon\ M \to \mathbb{R}$, so that the observed time series is generated by $\varphi_i = \varphi(x_i)$ (for the moment we assume that φ is independent of ω). The usual delay map can then be written as

$$\Phi_{f,\varphi}(x,\omega) = (\varphi(x), \varphi(f_{\omega_0}(x)), \varphi(f_{\omega_1\omega_0}(x)), \ldots, \varphi(f_{\omega_{d-2}\cdots\omega_0}(x)))^{\dagger},$$

where $f_{\omega_i \cdots \omega_0} = f_{\omega_i} \circ \cdots \circ f_{\omega_0}$ (as in section 3.5 of [Stark, 1999]). Observe that $\Phi_{f,\varphi}$ can be regarded as either a map $M \times N^{d-1} \to \mathbb{R}^d$ or a map $M \times \Sigma \to \mathbb{R}^d$. In the latter form, $\Phi_{f,\varphi}$ is a candidate for a bundle embedding which is equivalent to the condition that $\Phi_{f,\varphi,\omega} = \Phi_{f,\varphi}(\bullet, \omega)$ should be an embedding for μ_Σ-almost every ω, or for generic ω. In [Stark, 1999] we were able to prove that this was indeed the case for finite-dimensional deterministic forcing. Given that in the setting described in section 2.1, σ defines a deterministic dynamical system (albeit an infinite-dimensional one) and $\Phi_{f,\varphi,\omega}$ depends only on a finite number of components of ω, it should not come as a great surprise that we can modify Theorem 3.2 of [Stark, 1999] to hold in the present context. Thus, as in [Stark, 1999] denote by $\mathcal{D}^r(M \times N, M)$ the space of maps $f\colon M \times N \to M$ such that $f_y\colon M \to M$ is a C^r diffeomorphism of M for any y, where as usual $f_y(x) = f(x, y)$. Then

Theorem 2.1. *Let M and N be compact manifolds of dimension $m \geq 1$ and n respectively. Suppose that $d \geq 2m + 1$. Then for $r \geq 1$, there exists a residual set of $(f, \varphi) \in \mathcal{D}^r(M \times N, M) \times C^r(M, \mathbb{R})$ such that for any (f, φ) in this set there is an open dense set of ω in Σ such that $\Phi_{f,\varphi,\omega}$ is an embedding.*

The same approach also gives a measure theoretic version of this theorem. Since the proof ultimately relies on Sard s Theorem, we need to impose some conditions on the shift-invariant measure on Σ that describes the stochastic forcing. The most straightforward case is when we assume that this measure is a product measure μ_∞ arising from a probability measure μ on N, as discussed above in section 2.1. This gives the following statement of the theorem, already announced in [Stark et al., 1997] and [Stark, 1999]:

Theorem 2.2. *Let M and N be compact manifolds of dimension $m \geq 1$ and n respectively and let μ be a measure on N which is absolutely continuous with respect to Lebesgue measure. Suppose that $d \geq 2m + 1$. Then for $r \geq 1$, there exists a residual set of $(f, \varphi) \in \mathcal{D}^r(M \times N, M) \times C^r(M, \mathbb{R})$ such that for any (f, φ) in this set $\Phi_{f,\varphi,\omega}$ is an embedding for μ_∞ almost every ω.*

Our method of proof, however, allows for more general stochastic processes. The key property required of μ_Σ is that the marginal measure μ_{d-1} on cylinders of length $d - 1$ is absolutely continuous with respect to Lebesgue measure. Recall that μ_{d-1} is the measure on N^{d-1} defined by $\mu_{d-1}(U) = \mu_\Sigma((\pi_{d-1})^{-1}(U))$ for all measurable sets $U \subset N^{d-1}$, where $\pi_{d-1}\colon \Sigma \to N^{d-1}$ is the projection $\pi_{d-1}(\omega) = (\omega_0, \ldots, \omega_{d-2})$. We then have

Theorem 2.3 (Takens' Theorem for Stochastic Systems). *Let M and N be compact manifolds of dimension $m \geq 1$ and n respectively and let μ_Σ be an invariant measure on $\Sigma = N^{\mathbb{Z}}$ such that μ_{d-1} is absolutely continuous with respect to Lebesgue measure on N^{d-1}. Suppose that $d \geq 2m + 1$. Then for $r \geq 1$, there exists a residual set of $(f, \varphi) \in \mathcal{D}^r(M \times N, M) \times C^r(M, \mathbb{R})$ such that for any (f, φ) in this set $\Phi_{f,\varphi,\omega}$ is an embedding for μ_Σ almost every ω.*

Note that $\Phi_{f,\varphi,\omega}$ depends only on the finite number of components $(\omega_0, \dots, \omega_{d-2}) \in N^{d-1}$ and hence as we shall see in section 3.1 below it is sufficient to prove that $\Phi_{f,\varphi,\omega}$ is an embedding for μ_{d-1} almost every $(\omega_0, \dots, \omega_{d-2})$. When N is zero dimensional, and hence consists of a finite number of points, the only set which has μ_{d-1} full measure is the whole of N^{d-1}. Thus when $\dim N = 0$ the theorem implies that $\Phi_{f,\varphi,\omega}$ is an embedding for all ω. It turns out that in this case we need to use a different proof (see section 3.2 below for a more detailed explanation) to that for $\dim N > 0$, which also gives an open dense set of (f, φ), rather than merely a residual set. We thus get

Theorem 2.4 (Takens' Theorem for Iterated Function Systems). *Let M and N be compact manifolds of dimension $m \geq 1$ and $n = 0$ respectively (so that N is a finite set of points). Suppose that $d \geq 2m + 1$ and $r \geq 1$. Then there exists an open dense set of $(f, \varphi) \in \mathcal{D}^r(M \times N, M) \times \mathcal{C}^r(M, \mathbb{R})$ such that for any (f, φ) in this set $\Phi_{f,\varphi,\omega}$ is an embedding for every $\omega \in \Sigma$.*

This thus encompasses both Theorems 2.1 and 2.3 for $\dim N = 0$, and is proved separately in section 6. It is also possible to incorporate "noisy observations" in either of these results. This in fact should lead to easier proofs, since it allows us greater freedom in making perturbations. On the other hand, such a generalization does significantly complicate the notation. We therefore postpone discussion of this until section 2.5 below.

We emphasize that in common with the standard Takens' Theorem, all of these results are only valid for generic f and φ. In the case of the standard theorem the genericity conditions on f (but not on φ) can be made explicit (e.g. [Takens, 1980]; [Huke, 1993]), and in principle could be checked for any particular system. In the present context, unfortunately, it does not seem possible to characterize the set of f and φ satisfying the theorem; one can be certain only that this set is residual. This makes it impossible to verify the relevance of the theorem to any specific system or class of systems. This is, however, not an uncommon situation in nonlinear dynamics, where many important results hold only for a residual set of systems. Interestingly, if one assumes analyticity, it turns out to be possible to prove identifiability results that hold for all systems [Sontag, 2002]. It would be interesting to see if similar techniques could be applied to delay embedding problems, even if only within the context of the standard theorem.

Finally, we are grateful to a referee for pointing out the relationship of these results to Sussmann's universal input theorem. Recall that $x, x' \in M$ are said to be *distinguishable* if for some $\omega \in \Sigma$ and some $i \in \mathbb{N}$ we have $\varphi(f_{\omega_{i-1} \dots \omega_0}(x)) \neq \varphi(f_{\omega_{i-1} \dots \omega_0}(x'))$. A system is *observable* if every pair of distinct points can be distinguished, and a sequence $\omega \in \Sigma$ is called *universal for observability* if it distinguishes every pair of points. Hence, in particular, if $\Phi_{f,\varphi,\omega}$ is an embedding, then ω is universal and therefore Theorem 2.3 implies that universal sequences are residual (for generic systems). This is similar to the discrete time version of Sussmann's theorem, which states that the set of universal sequences is residual for every analytic observable system [Wang and Sontag, 1995].

2.4. Reconstructing the Dynamics

The most important consequence of the standard Takens' Theorem is that when Φ is an embedding it is possible to reconstruct a copy of the original system from successive

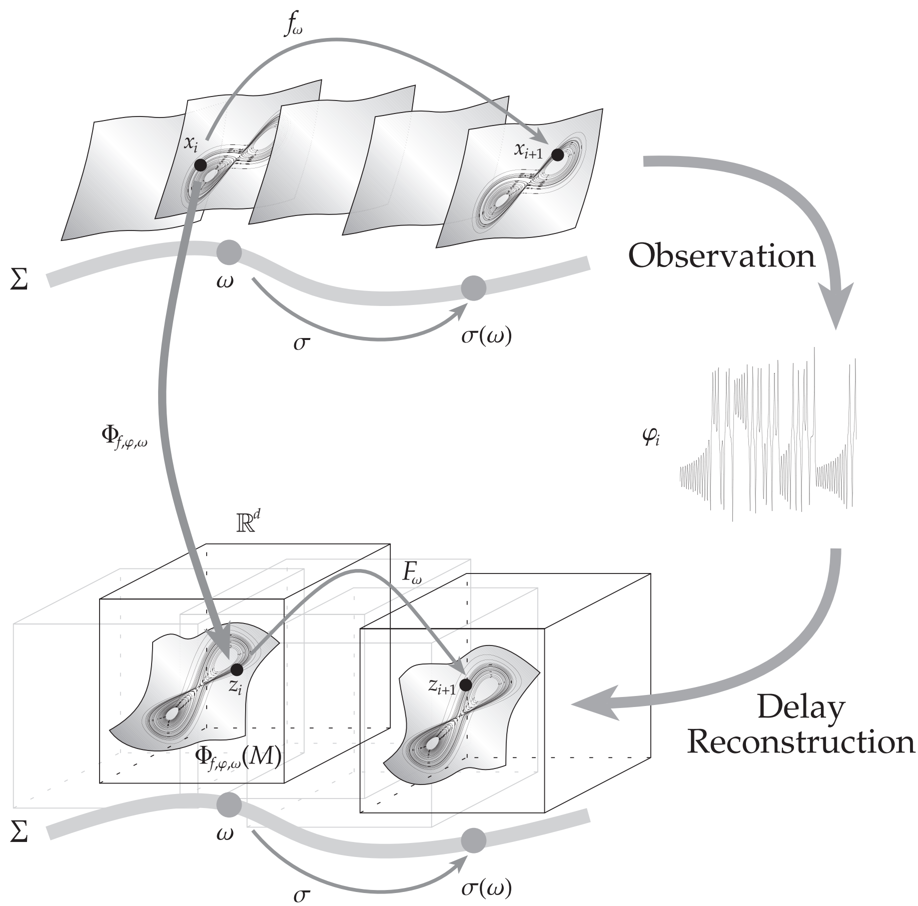

Fig. 3. Delay embedding of a random dynamical system.

observations of φ. Essentially the same situation holds in our generalized framework (Figure 3), though the term "reconstruction" is perhaps misleading in this case unless ω is known.

Thus, suppose that ω is such that $\Phi_{f,\varphi,\omega}$ and $\Phi_{f,\varphi,\sigma(w)}$ are both embeddings of M. Then the map $F_\omega = \Phi_{f,\varphi,\sigma(w)} \circ f_{\omega_0} \circ (\Phi_{f,\varphi,\omega})^{-1}$ is well defined and is a diffeomorphism between $\Phi_{f,\varphi,\omega}(M) \subset \mathbb{R}^d$ and $\Phi_{f,\varphi,\sigma(w)}(M) \subset \mathbb{R}^d$. Let $(x_i, \sigma^i(\omega))$ be an orbit of (f, σ), so that $x_{i+1} = f(x_i, \omega_i)$; recall that $\varphi_i = \varphi(x_i)$, and define $z_i = (\varphi_i, \varphi_{i+1}, \dots, \varphi_{i+d-1})$. Then $z_i = \Phi_{f,\varphi,\sigma^i(\omega)}(x_i)$ and hence

$$\begin{aligned} z_{i+1} &= \Phi_{f,\varphi,\sigma^{i+1}(\omega)}(x_{i+1}) \\ &= \Phi_{f,\varphi,\sigma^{i+1}(\omega)}(f_{\omega_i}(x_i)) \\ &= \Phi_{f,\varphi,\sigma(\sigma^i(\omega))}(f_{\omega_i}((\Phi_{f,\varphi,\sigma^i(\omega)})^{-1}(z_i))) \\ &= F_{\sigma^i(\omega)}(z_i). \end{aligned}$$

Therefore in exact analogy to the standard Takens' framework, $F_{\sigma^i(\omega)}$ is just the map which shifts a block of the time series forward by one time step, and hence $(\varphi_i, \varphi_{i+1}, \dots, \varphi_{i+d-1}) \mapsto (\varphi_{i+1}, \varphi_{i+2}, \dots, \varphi_{i+d})$ is bundle conjugate to our original dynamics f_{ω_i}. Note however, that whereas f_{ω_i} only depends on $\omega_i = \sigma^i(\omega)_0$, the map $F_{\sigma^i(\omega)}$ depends on $\omega_i, \omega_{i+1}, \dots, \omega_{i+d-1}$. Also in contrast to the standard framework, different F_ω have different domains (each a subset of $\mathbb{R}^d$ diffeomorphic to M). There is no reason in general why these should all be disjoint.

The first $d-1$ components of F_ω are trivial. If we denote the last component by G_ω: $\Phi_{f,\varphi,\omega}(M) \to \mathbb{R}$, then

$$\varphi_{i+d} = G_{\sigma^i(\omega)}(\varphi_i, \varphi_{i+1}, \dots, \varphi_{i+d-1}).$$

If we write out the dependence on $\sigma^i(\omega)$ explicitly, we get

$$\varphi_{i+d} = G(\varphi_i, \varphi_{i+1}, \dots, \varphi_{i+d-1}, \omega_i, \omega_{i+1}, \dots, \omega_{i+d-1}). \tag{2.4}$$

The existence of such a function was conjectured by [Casdagli, 1992]. From another point of view, processes of this form are well known in both signal processing and control theory (e.g. [Sontag, 1979a and b]; [Leontaritis and Billings, 1985]; [Chen and Billings, 1989]).

It is clear that estimating both G and ω in such a model is a major challenge. Note that $\omega_i, \dots, \omega_{i+d-2}$ have all been determined by time $i+d-1$, and so there is at least some hope of estimating them from previous values of the time series. By contrast ω_{i+d-1} corresponds to new uncertainty entering the system in the time step from $i+d-1$ to $i+d$.

2.5. Noisy Observations

So far we have assumed that the observations are completely noise-free. This is clearly unrealistic in many applications. In this section we therefore generalize our approach to cover the possibility of noise in the observations. We can envisage two cases: The observation noise may be independent of the noise in the dynamics, or both may depend on the same underlying stochastic process. In the first case, which is more likely to be relevant in applications, we can construct another shift space $\Sigma' = (N')^{\mathbb{Z}}$ to represent the observation dynamics, where N' is some compact manifold. The observation function is now φ: $M \times N' \to \mathbb{R}$, which gives rise to the delay map $\Phi_{f,\varphi,\omega,\eta}$ defined by

$$\Phi_{f,\varphi,\omega,\eta}(x) = (\varphi_{\eta_0}(x), \varphi_{\eta_1}(f_{\omega_0}(x)), \varphi_{\eta_2}(f_{\omega_1\omega_0}(x)), \dots, \varphi_{\eta_{d-1}}(f_{\omega_{d-2}\cdots\omega_0}(x)))^\dagger,$$

where $\varphi_{\eta_i}(x) = \varphi(x, \eta_i)$ for $\eta \in \Sigma'$. We can then require this to be an embedding for typical $(\omega, \eta) \in \Sigma \times \Sigma'$:

Theorem 2.5. *Let M, N, and N' be compact manifolds, with $m = \dim M > 0$. Suppose that $d \geq 2m+1$. Then for $r \geq 1$, there exists a residual set of $(f, \varphi) \in \mathcal{D}^r(M \times N, M) \times \mathcal{C}^r(M \times N', \mathbb{R})$ such that for any (f, φ) in this set there is an open dense set $\Sigma_{f,\omega} \subset \Sigma \times \Sigma'$ such that $\Phi_{f,\varphi,\omega,\eta}$ is an embedding for all $(\omega, \eta) \in \Sigma_{f,\varphi}$. If μ_Σ and $\mu'_{\Sigma'}$ are invariant probability measures on Σ and Σ' respectively, such that μ_{d-1} and μ'_d are absolutely continuous with respect to Lebesgue measure on N^{d-1} and $(N')^d$ respectively, then we can choose $\Sigma_{f,\varphi}$ such that $\mu_\Sigma \times \mu'_{\Sigma'}(\Sigma_{f,\varphi}) = 1$.*

As with Theorem 2.3, we need to treat the cases $\dim N = 0$ and $\dim N' = 0$ separately. Details can be found in section 7 below, which also contains the proofs of the various resulting versions of Theorem 2.5. The other possibility is that the observation noise depends on the same process as the dynamic noise. The relevant delay map is then

$$\tilde{\Phi}_{f,\varphi,\omega}(x) = (\varphi_{\omega_0}(x), \varphi_{\omega_1}(f_{\omega_0}(x)), \varphi_{\omega_2}(f_{\omega_1\omega_0}(x)), \ldots, \varphi_{\omega_{d-1}}(f_{\omega_{d-2}\cdots\omega_0}(x)))^{\dagger},$$

where $\varphi_{\omega_i}(x) = \varphi(x, \omega_i)$. We then get

Theorem 2.6. *Let M and N be compact manifolds, with $m = \dim M > 0$. Suppose that $d \geq 2m+1$. Then for $r \geq 1$, there exists a residual set of $(f, \varphi) \in \mathcal{D}^r(M \times N, M) \times \mathcal{C}^r(M \times N, \mathbb{R})$ such that for any (f, φ) in this set there is an open dense set $\Sigma_{f,\varphi} \subset \Sigma$ such that $\tilde{\Phi}_{f,\varphi,\omega}$ is an embedding for all $\omega \in \Sigma_{f,\varphi}$. If μ_Σ is an invariant probability measure on $\Sigma = N^{\mathbb{Z}}$ such that μ_d is absolutely continuous with respect to Lebesgue measure on N^d, then we can choose $\Sigma_{f,\varphi}$ such that $\mu_\Sigma(\Sigma_{f,\varphi}) = 1$.*

Again, the case $\dim N = 0$ is treated separately. Details and proofs are in section 8.

3. Structure of the Proofs of Theorems 2.1 and 2.3

The proofs of Theorems 2.1 and 2.3 are largely motivated by the proof of Theorem 3.2 of [Stark, 1999]. The main new ingredient is that the driving system is now infinite dimensional. However, since $\Phi_{f,\varphi,\omega}$ depends only on a finite number of components of ω, we can easily reduce the theorem to a finite-dimensional one. This is done in section 3.1 below. Additionally, in the measure theoretic case (Theorem 2.3), we require a measure theoretic version of the finite-dimensional Parametric Transversality Theorem ([Abraham, 1963]; [Abraham and Robbin, 1967]; or see Appendix A of [Stark, 1999]). As far as we are aware there is no published proof of this theorem, but it follows trivially by replacing the use of Smale's Density Theorem in the standard Parametric Transversality Theorem by an application of Sard's Theorem. We sketch the argument in Appendix A.

3.1. Reduction to Finite Dimensions

Since $\Phi_{f,\varphi,\omega}$ depends only on $\omega_0, \ldots, \omega_{d-2}$, it turns out to be sufficient to consider subsets of N^{d-1} rather than of Σ itself. Recall that we defined the projection $\pi_{d-1}\colon \Sigma \to N^{d-1}$ by $\pi_{d-1}(\omega) = (\omega_0, \ldots, \omega_{d-2})$. Given any $U \subset N^{d-1}$ we can define the cylinder Σ_U by $\Sigma_U = (\pi_{d-1})^{-1}(U)$ so that

$$\Sigma_U = \{\omega \in \Sigma : (\omega_0, \ldots, \omega_{d-2}) \in U\}.$$

Then

Lemma 3.1. *If U is open and dense in N^{d-1}, then Σ_U is open and dense in Σ. If U is residual in N^{d-1}, then Σ_U is residual in Σ.*

Proof. Recall that by the definition of the product topology on Σ, π_{d-1} is a continuous open mapping. Thus if U is open, then $\Sigma_U = (\pi_{d-1})^{-1}(U)$ is open. Now suppose that Σ_U is not dense. Then $\Sigma \backslash \Sigma_U$ contains some open set V. Since $\pi_{d-1}(V) \cap U = \emptyset$ and $\pi_{d-1}(V)$ is open, this means that U cannot be dense. Hence if U is dense, then so is Σ_U. If U is residual, then U contains the countable intersection of dense open sets U_i. By the above, each $(\pi_{d-1})^{-1}(U_i)$ is open and dense and hence their intersection is residual. □

If μ_Σ is a measure on Σ, recall that we defined the marginal measure μ_{d-1} on N^{d-1} by $\mu_{d-1}(U) = \mu_\Sigma((\pi_{d-1})^{-1}(U))$ for all measurable sets $U \subset N^{d-1}$. This definition immediately implies

Lemma 3.2. *If $U \subset N^{d-1}$ has full measure with respect to μ_{d-1}, then Σ_U has full measure with respect to μ_Σ.*

We also have the elementary

Lemma 3.3. *Suppose that $U \subset N^{d-1}$ is a measurable set of full Lebesgue measure (i.e. $N^{d-1} \backslash U$ has Lebesgue measure 0). Then U is dense in N^{d-1} and $\mu_{d-1}(U) = 1$ for any probability measure μ_{d-1} that is absolutely continuous with respect to Lebesgue measure on N^{d-1}.*

Proof. If U is not dense, then $N^{d-1} \backslash U$ contains an open set and hence is not of Lebesgue measure 0. Hence U cannot be of full Lebesgue measure. This contradiction implies that U must be dense if it has full Lebesgue measure. For the second part, if $N^{d-1} \backslash U$ has Lebesgue measure 0, then $\mu_{d-1}(N^{d-1} \backslash U) = 0$ by the definition of absolute continuity. Thus $\mu_{d-1}(U) = 1$, as required. □

Finally, if $\omega \in \Sigma$, define $y = \pi_{d-1}(\omega) = (\omega_0, \ldots, \omega_{d-2}) \in N^{d-1}$. Then $\Phi_{f,\varphi,\omega}(x) = \Phi_{f,\varphi,y}(x)$ where

$$\Phi_{f,\varphi,y}(x) = (\varphi(f^{(0)}(x, y)), \varphi(f^{(1)}(x, y)), \ldots, \varphi(f^{(d-1)}(x, y)))^\dagger, \tag{3.1}$$

with $f^{(i)}\colon M \times N^{d-1} \to M$ defined in an analogous fashion to section 3.2 of [Stark, 1999] by $f^{(i+1)}(x, y) = f(f^{(i)}(x, y), y_i) = f_{y_i}(f^{(i)}(x, y))$ and $f^{(0)}(x, y) = x$. Here y_i is the ith component of y, so that $y_i = \omega_i$; we adopt this notation to emphasize that y depends on only a finite number of components of ω.

Using Lemmas 3.1, 3.2, and 3.3, we can then immediately deduce Theorems 2.1 and 2.3 from

Theorem 3.4. *Let M and N be compact manifolds, with $m = \dim M > 0$ and suppose that $d \geq 2m+1$. Then for $r \geq 1$, there exists a residual set of $(f, \varphi) \in \mathcal{D}^r(M \times N, M) \times \mathcal{C}^r(M, \mathbb{R})$ such that for any (f, φ) in this set there is an open set $U_{f,\varphi} \subset N^{d-1}$ of full Lebesgue measure such that $\Phi_{f,\varphi,y}$ is an embedding for all $y \in U_{f,\varphi}$.*

3.2. The Dimension of N

Unfortunately, it turns out that to proceed further we have to treat the case of $n = \dim N = 0$ (i.e. the iterated function system case) separately from the general case of $n \geq 1$. The reason for this is that, as in Theorem 3.2 of [Stark, 1999], one of our first steps is to eliminate driving sequences ω with any repeating entries. Thus after reduction to finite dimensions, we remove those ω for which $\omega_i = \omega_j$ for any $i \neq j$ with $0 \leq i, j \leq d-2$. When $n \geq 1$, the set of such ω has zero Lebesgue measure in N^{d-1} and hence can be ignored. However, when $n = 0$ and N consists of a finite number of discrete points, this set will in general have positive measure. In effect, for $n = 0$ we have to prove the theorem for *all* ω, not just for Lebesgue almost all.

Additional insight into the $n = 0$ case comes from the observation that if N is a single point, then Theorems 2.1 and 2.3 reduce to the standard Takens' Theorem. Any proof of these theorems for $n = 0$ must therefore include a proof of the standard result. Such a proof (e.g. see section 4 of [Stark, 1999]) has to treat the short periodic orbits of f with some care. The image under the delay map of any point on such an orbit has components that are necessarily identical. This is because for such points $f^i(x) = f^j(x)$ for some $0 \leq i < j \leq 2d-1$. Each such point therefore has to be embedded individually. This can easily be done, since for generic f on a compact M there is only a finite number of short periodic orbits (this is the easy part of the Kupka-Smale Theorem; a self-contained proof is given in section 4.2.1 of [Stark, 1999]). In Theorems 2.1 and 2.3, the analogues of points on short periodic orbits are points whose images under $\Phi_{f,\varphi,y}$ are identical, that is points such that $x_i = x_j$ for some $0 \leq i < j \leq 2d-1$. Thus define

$$\tilde{P}_{\omega_0\cdots\omega_{q-1}} = \{x \in M : f_{\omega_{i-1}\ldots\omega_0}(x) = f_{\omega_{j-1}\ldots\omega_0}(x) \text{ for some } i \neq j \text{ with } 0 \leq i < j \leq q\}.$$

A technically intricate, but conceptually straightforward extension of the argument in section 5.1.1 of [Stark, 1999] can be used to show that for an open dense set of f, the set $\tilde{P}_{\omega_0\cdots\omega_{q-1}}$ consists of a finite number of points for any given $\omega_0, \omega_1, \ldots, \omega_{q-1}$. Thus, if $\dim N = 0$, so that there is only a finite number of choices of $\omega_0, \ldots, \omega_{q-1}$, the set

$$\tilde{P}^{(q)} = \bigcup_{(\omega_0\cdots\omega_{q-1})\in N^q} \tilde{P}_{\omega_0\cdots\omega_{q-1}}$$

is finite, and each point in can be dealt with individually. On the other hand if $\dim N \geq 1$, then this set is no longer finite, and a different approach is necessary, such as the one adopted in Theorem 3.2 of [Stark, 1999].

To summarize, therefore, we have to give somewhat different proofs in the cases $\dim N \geq 1$ and $\dim N = 0$ respectively. We shall outline the main ideas of the proof in the next two sections, and the detailed proofs are then given in section 5 and section 6 respectively.

3.3. Main Ideas: $\dim N \geq 1$

When $\dim N \geq 1$, Theorem 3.4 is very similar to Theorem 3.2 of [Stark, 1999], and it is not surprising that the proofs proceed along almost identical lines. The strategy is to first show that for a residual set of f and φ, $\tilde{T}\Phi_{f,\varphi}$ is transversal to the zero section in $T\mathbb{R}^d$ and $\Phi_{f,\varphi} \times \Phi_{f,\varphi}$ is transversal to the diagonal in $\mathbb{R}^d \times \mathbb{R}^d$ (with the domain of $\Phi_{f,\varphi,y} \times \Phi_{f,\varphi,y}$ being restricted to $M \times M \backslash \Delta$, where Δ is the diagonal in $M \times M$). Here $\Phi_{f,\varphi}$: $M \times N^{d-1} \to \mathbb{R}^d$ is defined by $\Phi_{f,\varphi}(x, y) = \Phi_{f,\varphi,y}(x)$, with $\Phi_{f,\varphi,y}$ given by (3.1), and $\tilde{T}\Phi_{f,\varphi}$ is the tangent map of $\Phi_{f,\varphi}$ restricted to the unit tangent bundle $\tilde{T}M = \{v \in TM : \|v\| = 1\}$ of M. We then treat y as a parameter and apply the Parametric Transversality Theorem to the maps $y \mapsto \tilde{T}\Phi_{f,\varphi,y}$ and $y \mapsto \Phi_{f,\varphi,y} \times \Phi_{f,\varphi,y}$. This shows that for a residual set of full measure of y the maps $\tilde{T}\Phi_{f,\varphi,y}$ and $\Phi_{f,\varphi,y} \times \Phi_{f,\varphi,y}$ are transversal to the zero section in $T\mathbb{R}^d$ and the diagonal in $\mathbb{R}^d \times \mathbb{R}^d$ respectively. We then apply the same dimension counting argument as was used repeatedly in [Stark, 1999] to show that if $d \geq 2m + 1$, then transversality implies nonintersection. Thus for a residual set of full measure of y the image of $\tilde{T}\Phi_{f,\varphi,y}$ cannot intersect the zero section in $T\mathbb{R}^d$ and the image of $\Phi_{f,\varphi,y} \times \Phi_{f,\varphi,y}$ (restricted to $M \times M \backslash \Delta$) cannot intersect the diagonal in $\mathbb{R}^d \times \mathbb{R}^d$. But these are precisely the statements of the immersivity and injectivity of $\Phi_{f,\varphi,y}$, respectively.

Exactly as in [Stark, 1999], the big problem is points such that $x_i = x_j$ with $i \neq j$. We proceed just as we did there, by ensuring that such points occur on a family of submanifolds $\tilde{W}_I$ of $M \times N^{d-1}$ and then dealing with each of these separately. Again, each $\tilde{W}_I$ is characterized by the set of pairs (i, j) for which $x_i = x_j$, and the more such pairs (i, j) there are, the smaller the dimension of the corresponding $\tilde{W}_I$. In particular we shall see that $\tilde{W}_{I,y} = \tilde{W}_I \cap (M \times \{y\})$ has codimension $(d - \gamma)m$ where γ is the number of distinct points in the set $\{x_0, \ldots, x_{d-1}\}$. Thus it seems plausible that generically $\Phi_{f,\varphi,y}$ should embed $\tilde{W}_{I,y}$ if $\gamma \geq 2(m - (d - \gamma)m) + 1$. But this is always satisfied if $d \geq 2m + 1$, since then $2m + 1 - 2(d - \gamma)m - \gamma \leq (d - \gamma)(1 - 2m) \leq 0$. Since the union of the $\tilde{W}_{I,y}$ over all I is $M \times \{y\}$, this means that $\Phi_{f,\varphi,y}$ should embed the whole of $M \times \{y\}$. Of course, the problem with this argument as stated is that the residual set of f and φ for which $\Phi_{f,\varphi,y}$ is an embedding will depend on y, and hence a priori the intersection of these sets over an open and dense set of y will not necessarily be residual. We deal with this as in section 6 of [Stark, 1999] by combining the construction of $\tilde{W}_I$, which itself is a transversality argument with the proof of the transversality of $\tilde{T}\Phi_{f,\varphi,y}$ and $\Phi_{f,\varphi,y} \times \Phi_{f,\varphi,y}$ on $\tilde{W}_I$. This allows a single invocation of the Parametric Transversality Theorem to yield a residual set of f and φ.

3.4. Main Ideas: **dim $N = 0$**

The overall structure closely follows the transversality proof of the standard Takens' Theorem given in [Stark, 1999], which in turn is based on Takens' original proof ([Takens, 1980]; [Huke, 1993]). We first ensure that periodic orbits of period less than $2d$ are isolated and have distinct eigenvalues, and then embed these individually. We then use the Parametric Transversality Theorem exactly as outlined above for dim $N > 0$ to show that for a dense set of f and φ, the map $\tilde{T}\Phi_{f,\varphi}$ is transversal to the zero section in $T\mathbb{R}^d$ and $\Phi_{f,\varphi} \times \Phi_{f,\varphi}$ is transversal to the diagonal in $\mathbb{R}^d \times \mathbb{R}^d$. Counting dimensions shows that if $d \geq 2m$, transversality implies nonintersection, so that $\Phi_{f,\varphi}$ is respectively immersive and injective.

The main complication is dealing with the short periodic orbits. In particular, a priori it is possible for two such orbits for two different sequences $(\omega_0, \omega_1, \ldots, \omega_{q-1})$ and

$(\omega'_0, \omega'_1, \ldots, \omega'_{q'-1})$ to have points in common. Not only does this prevent us from perturbing φ independently on each orbit when we embed the orbit, but in fact it turns out to be an obstruction if we try to mimic the proof of Lemma 4.8 of [Stark, 1999] to show that generically periodic orbits are isolated. A different way of stating this problem is that in the standard case, if x is a periodic orbit of f of minimal period q, then $f^i(x) = x$ if and only if $i = 0 (\mathrm{mod}\ q)$. This need no longer be the case in the present situation (and in fact the definition of minimal period requires some care). It turns out that to avoid this problem, we need to carry out an inductive step which simultaneously shows that if periodic orbits of period less than q are isolated and those for different sequences are distinct, then the same holds for period q orbits. This then allows us to carry the remainder of the proof in a more or less straightforward fashion.

4. Preliminary Calculations

4.1. Notation

As in section 6 of [Stark, 1999] we first prove the theorem for a sufficiently large r, and then show in section 5.5 that this implies the theorem for all $r \geq 1$. It turns out that initially we shall want f and φ to be $\mathcal{C}^{2r}$ where $r = n(d-1)$ (recall that we assume $d \geq 2m+1$ and $m \geq 1$). Also note that since $\Phi_{f,\varphi,y}$ depends continuously on y (Lemma 4.3 below) and embeddings are open in $\mathcal{C}^k(M, \mathbb{R}^d)$ (e.g. [Hirsch, 1976]), the set of y such that $\Phi_{f,\varphi,y}$ is an embedding for a fixed (f, φ) is open. Hence all that we need to prove is that this set has full Lebesgue measure.

Recall that in section 3.1, given $(f, \varphi) \in \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R})$, we defined $f^{(i)} \in \mathcal{D}^{2r}(M \times N^{d-1}, M)$ by $f^{(i+1)}(x, y) = f(f^{(i)}(x, y), y_i) = f_{y_i}(f^{(i)}(x, y))$ and $f^{(0)}(x, y) = x$. The delay map $\Phi_{f,\varphi}\colon M \times N^{d-1} \to \mathbb{R}^d$ is then given by

$$\Phi_{f,\varphi}(x, y) = (\varphi(f^{(0)}(x, y)), \varphi(f^{(1)}(x, y)), \ldots, \varphi(f^{(d-1)}(x, y)))^\dagger,$$

and for a given $y \in N^{d-1}$, we also define $\Phi_{f,\varphi,y}\colon M \to \mathbb{R}^d$ as in (3.1) by

$$\Phi_{f,\varphi,y}(x) = \Phi_{f,\varphi}(x, y). \tag{4.1}$$

Finally, let $\rho\colon \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R}) \to \mathcal{D}^r(M \times N^{d-1}, \mathbb{R}^d)$ be the map that takes (f, φ) to $\Phi_{f,\varphi}$

$$\rho(f, \varphi) = \Phi_{f,\varphi}. \tag{4.2}$$

Recall (e.g. Appendix A of [Stark, 1999]) that the *evaluation map* $ev_\rho\colon \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R}) \times M \times N^{d-1} \to \mathbb{R}^d$ of ρ is defined by $ev_\rho(f, \varphi, x, y) = \rho(f, \varphi)(x, y) = \Phi_{f,\varphi}(x, y)$. In proving immersivity, we shall also want the corresponding map for the tangent map of $\Phi_{f,\varphi}$. Since we are only interested in the immersivity of each $\Phi_{f,\varphi,y}$ individually, we want to differentiate only in the x direction. Thus $\tau\colon \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R}) \to \mathcal{VB}^{r-1}(TM \times N^{d-1}, T\mathbb{R}^d)$ is given by

$$\tau(f, \varphi) = T_1 \Phi_{f,\varphi}, \tag{4.3}$$

where T_1 is the partial tangent operator in the x direction, so that $T_1\Phi(v) = T\Phi(v, 0)$. Here, extending the notation of Appendix B.3 of [Stark, 1999], $\mathcal{VB}^{r-1}(TM \times N^{d-1}, T\mathbb{R}^d)$

is the space of maps from $TM \times N^{d-1}$ to $T\mathbb{R}^d$ that are linear on the fibres of TM, and whose dependence on $(x, y) \in M \times N^{d-1}$ is $\mathcal{C}^r$, but on $v \in T_x M$ is only $\mathcal{C}^{r-1}$. Thus $\mathcal{VB}^{r-1}(TM \times N^{d-1}, T\mathbb{R}^d)$ is a subspace of $\mathcal{C}^{r-1}(TM \times N^{d-1}, T\mathbb{R}^d)$; however since TM is not compact, the latter lacks a natural manifold structure. More usefully, note that for a given $y \in N^{d-1}$, a map in $\mathcal{VB}^{r-1}(TM \times N^{d-1}, T\mathbb{R}^d)$ is defined by its action on unit vectors in TM. We can thus also treat τ as mapping into $\mathcal{C}^{r-1}(\tilde{T}M \times N^{d-1}, T\mathbb{R})$. The evaluation map $ev_\tau\colon \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R})TM \times N^{d-1} \to \mathbb{R}^d$ is given by

$$\begin{aligned} ev_\tau(f, \varphi, v, y) &= \tau(f, \varphi)(v, y) \\ &= T_x \Phi_{f,\varphi,y}(v). \end{aligned}$$

Immersivity is defined by the condition that $T_x \Phi_{f,\varphi,y}(v) \neq 0$ for all $v \in TM$ such that $v \neq 0$. By linearity it is sufficient to consider just those v such that $\|v\| = 1$, and hence to restrict ourselves to the unit tangent bundle $\tilde{T}M = \{v \in TM : \|v\| = 1\}$. To emphasize this, as in section 3.3 above we shall denote the restriction of $T\Phi_{f,\varphi,y}$ to $\tilde{T}M$ by $\tilde{T}\Phi_{f,\varphi,y}$.

4.2. Smoothness of the Evaluation Maps

In order to apply the Parametric Transversality Theorem to $ev_{\rho\times\rho}$ and ev_τ, we need to show that these evaluation maps are smooth. This is done in a very similar fashion to the analogous results for deterministic forcing given in Appendices C.1 and C.2 of [Stark, 1999]. We also take this opportunity to compute various derivatives of the evaluation maps that we shall use later. We begin with $\hat{\rho}_i\colon \mathcal{D}^{2r}(M \times N, M) \to \mathcal{D}^r(M \times N^{d-1}, M)$ given by

$$\hat{\rho}_i(f) = f^{(i)}. \tag{4.4}$$

Lemma 4.1. *The map $\hat{\rho}_i$ is $\mathcal{C}^r$. Given a $\eta \in T_f\mathcal{D}^{2r}(M \times N, M)$, denote $\bar{\eta}_i = T_f\hat{\rho}_i(\eta)$. Then $\bar{\eta}_0 = 0$ and for $i = 1, \ldots, d-1$ we have*

$$\bar{\eta}_i(x, y) = \eta(x_{i-1}, y_{i-1}) + T_{(x_{i-1}, y_{i-1})} f(\bar{\eta}_{i-1}(x, y), 0). \tag{4.5}$$

Proof. By induction on i. Since $\hat{\rho}_0(f) = Id$ for all f, $\hat{\rho}_0$ is trivially $\mathcal{C}^r$. Suppose that $\hat{\rho}_{i-1}$ for some $i > 1$ is $\mathcal{C}^r$. Now, $f^{(i)} = f \circ (f^{(i-1)}, \pi_{i-1})$, where $\pi_i\colon M \times N^{d-1} \to N$ is the projection $\pi_i(x, y) = y_i$. Therefore $\hat{\rho}_i(f) = \chi((\hat{\rho}_{i-1}(f), \pi_{i-1}), f)$ where $\chi\colon \mathcal{D}^r(M \times N^{d-1}, M \times N) \times \mathcal{D}^{2r}(M \times N, M) \to \mathcal{D}^r(M \times N^{d-1}, M)$ is the composition $\chi(F, f) = f \circ F$. Hence $\hat{\rho}_i = \chi \circ ((\hat{\rho}_{i-1}, \hat{\pi}_{i-1}), Id)$ where $Id\colon \mathcal{D}^{2r}(M \times N, M) \to \mathcal{D}^{2r}(M \times N, M)$ is the identity $Id(f) = f$ and $\hat{\pi}_i\colon \mathcal{D}^{2r}(M \times N, M) \to \mathcal{C}^r(M \times N^{d-1}, N)$ is the constant map $\hat{\pi}_i(f) = \pi_i$. Clearly Id and $\hat{\pi}_i$ are $\mathcal{C}^\infty$ maps, and it is well known that the composition operator χ is $\mathcal{C}^r$ ([Eells, 1966]; [Foster, 1975]; [Franks, 1979], or see Theorem B.2 of [Stark, 1999]). Hence by the chain rule $\hat{\rho}_i$ is $\mathcal{C}^r$.

Since $\hat{\rho}_0(f) = Id$ for all f, we have $\bar{\eta}_0 = 0$. We differentiate $\hat{\rho}_i(f) = \chi((\hat{\rho}_{i-1}(f), \pi_{i-1}), f)$ using the chain rule and the fact that $T_{(F,f)}\chi(\zeta, \eta) = \eta \circ F + Tf \circ \zeta$ (see [Eells, 1966]; [Foster, 1975]; [Franks, 1979] or Theorem B.2 of [Stark, 1999]). This yields $T_f\hat{\rho}_i(\eta) = \eta \circ (\hat{\rho}_{i-1}(f), \pi_{i-1}) + Tf \circ (T_f\hat{\rho}_{i-1}(\eta), 0)$. Substituting $T_f\hat{\rho}_i(\eta) = \bar{\eta}_i$ and $\hat{\rho}_{i-1}(f) = f^{(i-1)}$, we obtain $\bar{\eta}_i = \eta \circ (f^{(i-1)}, \pi_{i-1}) + Tf \circ (\bar{\eta}_{i-1}, 0)$. Evaluating this at (x, y) and using the fact that $f^{(i-1)}(x, y) = x_{i-1}$ and $\pi_{i-1}(x, y) = x_{i-1}$ gives (4.5). $\square$

The corresponding evaluation map $ev_{\hat{\rho}_i}\colon \mathcal{D}^{2r}(M \times N, M) \times M \times N^{d-1} \to M$ is given by $ev_{\hat{\rho}_i}(f, x, y) = f^{(i)}(x, y) = x_i$. Note that, with one or two exceptions, we will only ever need to evaluate $Tev_{\hat{\rho}_i}$ on vectors of the form $(\eta, 0_x, 0_y)$. In other words, we need to consider only perturbations to f but not to x or y.

Corollary 4.2. *The evaluation map $ev_{\hat{\rho}_i}$ is $\mathcal{C}^r$ and $T_{(f,x,y)}ev_{\hat{\rho}_i}(\eta, 0_x, 0_y) = \bar{\eta}_i(x, y)$.*

Proof. By Corollary B.3 of [Stark, 1999] (which is a simple corollary of the smoothness of composition), the evaluation operator is smooth. More precisely $ev\colon \mathcal{D}^r(M \times N^{d-1}, M) \times M \times N^{d-1} \to M$ given by $ev(F, x, y) = F(x, y)$ is $\mathcal{C}^r$ and $T_{(F,x,y)}ev(\zeta, v, w) = \zeta(x, y) + T_{(x,y)}F(v, w)$. But $ev_{\hat{\rho}_i} = ev \circ (\hat{\rho}_i \times Id_x \times Id_y)$ and hence $ev_{\hat{\rho}_i}$ is $\mathcal{C}^r$ by the chain rule. Differentiating $ev_{\hat{\rho}_i} = ev \circ (\hat{\rho}_i \times Id_x \times Id_y)$ using the chain rule, we obtain

$$T_{(f,x,y)}ev_{\hat{\rho}_i}(\eta, v, w) = \bar{\eta}_i(x, y) + T_{(x,y)}f^{(i)}(v, w). \tag{4.6}$$

Evaluating this for $v = 0_x$, $w = 0_y$ yields $T_{(f,x,y)}ev_{\hat{\rho}_i}(\eta, 0_x, 0_y) = \bar{\eta}_i(x, y)$, as claimed. $\square$

Now define $\rho_i\colon \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R}) \to \mathcal{C}^r(M \times N^{d-1}, \mathbb{R})$ by

$$\rho_i(f, \varphi) = \varphi \circ f^{(i)}. \tag{4.7}$$

Lemma 4.3. *The map ρ_i is $\mathcal{C}^r$, with*

$$T_{(f,\varphi)}\rho_i(\eta, \xi) = \xi \circ f^{(i)} + T\varphi \circ \bar{\eta}_i. \tag{4.8}$$

The evaluation map $ev_{\rho_i}\colon \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R}) \times M \times N^{d-1} \to \mathbb{R}$ is also $\mathcal{C}^r$, and if we denote $\Xi = (f, \varphi, x, y)$, then

$$T_{\Xi}ev_{\rho_i}(\eta, \xi, v, w) = \xi(x_i) + T_{x_i}\varphi[\bar{\eta}_i(x, y) + T_{(x,y)}f^{(i)}(v, w)]. \tag{4.9}$$

Proof. We have $\rho_i = \chi' \circ (\hat{\rho}_i \times Id_\varphi)$ where χ' is the composition $\chi'\colon \mathcal{C}^r(M \times N^{d-1}, M) \times \mathcal{C}^{2r}(M, \mathbb{R}) \to \mathcal{C}^r(M \times N^{d-1}, M)$ and Id_φ is the identity on $\mathcal{C}^{2r}(M, \mathbb{R})$. As above, χ' and Id_φ are $\mathcal{C}^r$ and so is $\hat{\rho}$ by Lemma 4.1. Hence ρ_i is $\mathcal{C}^r$ by the chain rule. Differentiating $\rho_i = \chi' \circ (\hat{\rho}_i \times Id_\varphi)$, e.g. using Theorem B.2 of [Stark, 1999], we obtain $T_{(f,\varphi)}\rho_i(\eta, \xi) = \xi \circ \hat{\rho}_i(f) + T\varphi \circ T_f\hat{\rho}_i(\eta) = \xi \circ f^{(i)} + T\varphi \circ \bar{\eta}_i$, as required.

Turning to the evaluation function ev_{ρ_i}, we proceed in a similar fashion to Corollary 4.2. We have $ev_{\rho_i} = ev' \circ (\rho_i \times Id_x \times Id_y)$ where $ev'\colon \mathcal{C}^r(M \times N^{d-1}, \mathbb{R}) \times M \times N^{d-1} \to \mathbb{R}$ is the evaluation $ev'(\Phi, x, y) = \Phi(x, y)$. By Corollary B.3 of [Stark, 1999] this is $\mathcal{C}^r$ and $T_{(\Phi,x,y)}ev'(\zeta, v, w) = \zeta(x, y) + T_{(x,y)}\Phi(v, w)$. Hence ev_{ρ_i} is $\mathcal{C}^r$ by the chain rule and $T_{\Xi}ev_{\rho_i}(\eta, \xi, v, w) = T_{(f,\varphi)}\rho_i(\eta, \xi)(x, y) + T_{(x,y)}(\rho_i(f, \varphi))(v, w) = \xi \circ f^{(i)}(x, y) + T_{x_i}\varphi(\bar{\eta}_i(x, y)) + T_{x_i}\varphi(T_{(x,y)}f^{(i)}(v, w)) = \xi(x_i) + T_{x_i}\varphi[\bar{\eta}_i(x, y) + T_{(x,y)}f^{(i)}(v, w)]$. $\square$

We shall not need to evaluate $T_{\Xi}ev_{\rho_i}$ except for vectors of the form $(0_f, \xi, 0_x, 0_y)$; in other words, we consider only perturbations to φ, but not to f, x, or y. We now turn to analogous results for the derivative of $\varphi \circ f^{(i)}$ and so define $\tau_i\colon \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R}) \to \mathcal{VB}^{r-1}(TM \times N^{d-1}, T\mathbb{R})$ by

$$\tau_i(f, \varphi) = T_1(\varphi \circ f^{(i)}), \tag{4.10}$$

where as usual T_1 is the partial tangent in the x direction.

Lemma 4.4. *The operator τ_i is $\mathcal{C}^r$. Its evaluation map $ev_{\tau_i}\colon \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R}) \times TM \times N^{d-1} \to T\mathbb{R}$ is $\mathcal{C}^{r-1}$, and if we denote $\Xi' = (f, \varphi, v, y)$, then*

$$T_{\Xi'} ev_{\tau_i}(0_f, \xi, 0_v, 0_y) = \bar{\omega}(T_{x_i}\xi(v_i)),$$

where $x_i = f^{(i)}(x, y)$ and $v_i = T_{(x,y)} f^{(i)}(v, 0_y) = T_{1,(x,y)} f^{(i)}(v)$ and $\bar{\omega}$ is the canonical involution on $T(T\mathbb{R})$. Note that in [Stark, 1999] the canonical involution is denoted by ω, but here we want to avoid confusion with the usage of ω as the forcing sequence in sections 2 and 3.1.

Proof. We have $\tau_i = \sigma_1' \circ \rho_i$ where $\sigma_1'\colon \mathcal{C}^r(M \times N^{d-1}, \mathbb{R}) \to \mathcal{VB}^{r-1}(TM \times N^{d-1}, T\mathbb{R})$ is the operator that takes the partial tangent in the x direction. Then $\sigma_1'(\Phi)(w, y) = T_1\Phi(w, y) = T\Phi(w, 0_y) = \sigma'(\Phi)(w, 0_y)$ where $\sigma'\colon \mathcal{C}^r(M \times N^{d-1}, \mathbb{R}) \to \mathcal{VB}^{r-1}(TM \times TN^{d-1}, T\mathbb{R})$ is the full tangent operator. By Lemma B.11 of [Stark, 1999], σ' is $\mathcal{C}^\infty$ and $T_\Phi\sigma'(\zeta) = \bar{\omega} \circ T\zeta$. Let $\iota_y\colon \mathcal{VB}^{r-1}(TM \times TN^{d-1}, T\mathbb{R}) \to \mathcal{VB}^{r-1}(TM \times N^{d-1}, T\mathbb{R})$ be the inclusion defined by $\iota_y(\Psi)(v, y) = \Psi(v, 0_y)$. Thus $\iota_y(\Psi) = \Psi \circ (Id \times L_{N^{d-1}})$, where $L_{N^{d-1}}\colon N^{d-1} \to TN^{d-1}$ is the zero section in TN^{d-1}, i.e. $L_{N^{d-1}}(y) = 0_y$. Hence ι_y is just composition on the right with a fixed function. This is just a linear map in the local chart described in Appendix B.1 of [Stark, 1999], and hence ι_y is $\mathcal{C}^\infty$, with $T_\Psi\iota_y(\zeta)(v, y) = \zeta(v, 0_y)$ (this result appears as Proposition 4.2 in [Franks, 1979]). But $\sigma_1'(\Phi)(v, y) = \sigma'(\Phi)(v, 0_y) = \iota_y(\sigma'(\Phi))(v, y)$ so that $\sigma_1' = \iota_y \circ \sigma'$ and thus σ_1' is $\mathcal{C}^\infty$. Hence by the chain rule $\tau_i = \sigma_1' \circ \rho_i = \iota_y \circ \sigma' \circ \rho_i$ is $\mathcal{C}^r$.

Proceeding as in Corollary 4.2 and Lemma 4.3, we see that the evaluation function ev_{τ_i} is given by $ev_{\tau_i} = ev'' \circ (\tau_i \times Id_v \times Id_y)$ where $ev''\colon \mathcal{VB}^{r-1}(TM \times N^{d-1}, T\mathbb{R}) \times TM \times N^{d-1} \to T\mathbb{R}$ is the evaluation $ev''(\Psi, v, y) = \Psi(v, y)$. By Corollary B.3 of [Stark, 1999], this is $\mathcal{C}^{r-1}$ and $T_{(\Psi,v,y)}ev''(\zeta, w, w') = \zeta(v, y) + T_{(v,y)}\Psi(w, w')$. Thus ev_{τ_i} is $\mathcal{C}^{r-1}$ by the chain rule and $T_{\Xi'}ev_{\tau_i}(0_f, \xi, 0_v, 0_y) = T_{(f,\varphi)}\tau_i(0_f, \xi)(v, y)$. Differentiating $\tau_i = \iota_y \circ \sigma' \circ \rho_i$ using the chain rule gives $T_{(f,\varphi)}\tau_i(0_f, \xi) = T\iota_y \circ \bar{\omega} \circ T(T_{(f,\varphi)}\rho_i(0_f, \xi))$. By (4.8) we have $T_{(f,\varphi)}\rho_i(0_f, \xi) = \xi \circ f^{(i)}$ and hence $T_{\Xi'}ev_{\tau_i}(0_f, \xi, 0_v, 0_y) = \bar{\omega}(T\xi \circ Tf^{(i)}(v, 0_y)) = \bar{\omega}(T_{x_i}\xi \circ T_{(x,y)} f^{(i)}(v, 0_y)) = \bar{\omega}(T_{x_i}\xi(v_i))$ as required. □

5. Proof of Theorem 3.4 for dim $N \geq 1$

It turns out that initially we shall want f and φ to be $\mathcal{C}^{2r}$ where $r = n(d-1)$ (recall that we assume $d \geq 2m + 1$ and $m \geq 1$). Also note that since $\Phi_{f,\varphi,y}$ depends continuously on y (Lemma 4.3 below) and embeddings are open in $\mathcal{C}^k(M, \mathbb{R}^d)$ (e.g. [Hirsch, 1976]), the set of y such that $\Phi_{f,\varphi,y}$ is an embedding for a fixed (f, φ) is open. Hence all that we need to prove is that this set has full Lebesgue measure.

As long as $n \geq 1$, the set of $y \in N^{d-1}$ such that $y_i = y_j$ for some $i \neq j$ is closed, is nowhere dense, and has zero Lebesgue measure. We can thus ignore it and restrict ourselves to $\tilde{N}_{d-1}$ where

$$\tilde{N}_{d-1} = \{y \in N^{d-1} : y_i \neq y_j \text{ for all } i \neq j\}. \tag{5.1}$$

5.1. Partitions of the Components of Φ

As sketched in section 3.3 above, our overall strategy will be to make $ev_{\rho\times\rho}$ and ev_τ transversal to specific submanifolds of $\mathbb{R}^d \times \mathbb{R}^d$ and $T\mathbb{R}^d$. The difficulty in doing this occurs at points where $x_i = x_j$ with $i \neq j$. At such points, the ith and jth components of $\Phi_{f,\varphi}$ cannot be perturbed independently. We therefore want to define independent subsets of such components. We employ the same notation as in [Stark, 1999]. Thus, let $I = \{I_1, I_2, \ldots, I_\alpha\}$ be a partition of $\{0, \ldots, d-1\}$, and define the associated equivalence relation $\sim_I$ on $\{0, \ldots, d-1\}$ by $i \sim_I j$ if and only if i and j are in the same element of the partition. Given any such partition I, let $\mathcal{J}_I$ be a set containing precisely one element from each I_k for $k = 1, \ldots, \alpha$. There will typically be many ways to choose such a $\mathcal{J}_I$, but we arbitrarily select just one. Clearly $\mathcal{J}_I$ has α elements. Write these as $\mathcal{J}_I = \{j_1, j_2, \ldots, j_\alpha\}$ with $j_1 < j_2 < \cdots < j_\alpha$. For a given partition I, define the map $\Phi_{f,\varphi,I}\colon M \times N^{d-1} \to \mathbb{R}^\alpha$ by

$$\Phi_{f,\varphi,I}(x, y) = (\varphi(f^{(j_1)}(x, y)), \varphi(f^{(j_2)}(x, y)), \ldots, \varphi(f^{(j_\alpha)}(x, y)))^\dagger. \tag{5.2}$$

We can also write this as

$$\Phi_{f,\varphi,I}(x, y) = (\varphi(x_{j_1}), \varphi(x_{j_2}), \ldots, \varphi(x_{j_\alpha}))^\dagger,$$

where $x_i = f^{(i)}(x, y)$. This has the advantage that it emphasizes that we observe the x coordinate only, but the disadvantage that it obscures the dependence of x_i on y, which can be a potential source of errors when performing calculations. Let $\rho_I\colon \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R}) \to \mathcal{C}^r(M \times N^{d-1}, \mathbb{R}^\alpha)$ be defined in the obvious way by

$$\rho_I(f, \varphi) = \Phi_{f,\varphi,I}, \tag{5.3}$$

and for any $y \in N^{d-1}$, define $\Phi_{f,\varphi,I,y}\colon M \to \mathbb{R}^\alpha$ by

$$\Phi_{f,\varphi,I,y}(x) = \Phi_{f,\varphi,I}(x, y).$$

Note that $\rho_I = (\rho_{j_1}, \rho_{j_2}, \ldots, \rho_{j_\alpha})^\dagger$ where $\rho_i(f, \varphi) = \varphi \circ f^{(i)}$ as defined in (4.7). As an immediate consequence of this we get

Corollary 5.1. *The map ρ_I and its evaluation map ev_{ρ_I} are $\mathcal{C}^r$ for any partition I of $\{0, \ldots, d-1\}$. If $\Xi = (f, \varphi, x, y)$, then*

$$T_\Xi ev_{\rho_I}(0_f, \xi, 0_x, 0_y) = (\xi(x_{j_1}), \xi(x_{j_2}), \ldots, \xi(x_{j_\alpha}))^\dagger.$$

This corollary is the crucial result underlying our whole approach: It shows that if the points $x_{j_1}, x_{j_2}, \ldots, x_{j_\alpha}$ are distinct, then $T_\Xi ev_{\rho_I}$ is surjective and hence transversal to any submanifold of $\mathbb{R}^\alpha$.

Finally let $\tau_I\colon \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R}) \to \mathcal{C}^{r-1}(\tilde{T}M \times N^{d-1}, T\mathbb{R}^{\alpha})$ be defined by

$$\tau_I(f, \varphi) = T_1 \Phi_{f,\varphi,I}, \tag{5.4}$$

where, as above, T_1 is the partial tangent operator in the x direction. Thus

$$ev_{\tau_I}(f, \varphi, v, y) = \tau_I(f, \varphi)(v, y) = T_x \Phi_{f,\varphi,I,y}(v).$$

Similarly to ρ_I we have $\tau_I = (\tau_{j_1}, \tau_{j_2}, \ldots, \tau_{j_\alpha})^{\dagger}$, where $\tau_i(f, \varphi) = T_1(\varphi \circ f^{(i)})$ as in (4.10), giving

Corollary 5.2. *The map τ_I and its evaluation map ev_{τ_I} are $\mathcal{C}^{r-1}$ for any partition I of $\{0, \ldots, d-1\}$. If $\Xi' = (f, \varphi, v, y)$, then*

$$T_{\Xi'} ev_{\tau_I}(0_f, \xi, 0_x, 0_y) = (\bar{\omega}(T_{x_1}\xi(v_{j_1})), \bar{\omega}(T_{x_2}\xi(v_{j_2})), \ldots, \bar{\omega}(T_{x_a}\xi(v_{j_\alpha})))^{\dagger}.$$

5.2. Surjectivity of $ev_{\hat{\rho}}$

Recall from Lemma 4.1 and Corollary 4.2 that $\hat{\rho}_i\colon \mathcal{D}^{2r}(M \times N, M) \to \mathcal{D}^r(M \times N^{d-1}, M)$ given by $\hat{\rho}_i(f) = f^{(i)}$ is $\mathcal{C}^r$ and $T_{(f,x,y)} ev_{\hat{\rho}_i}(\eta, 0_x, 0_y) = \bar{\eta}_i(x, y)$ where $\bar{\eta}_i$ satisfies (4.5). Let $\hat{\rho} = (\hat{\rho}_0, \hat{\rho}_1, \ldots, \hat{\rho}_{d-1})\colon \mathcal{D}^{2r}(M \times N, M) \to \mathcal{D}^r(M \times N^{d-1}, M^d)$ and denote the corresponding evaluation map $ev_{\hat{\rho}}\colon \mathcal{D}^{2r}(M \times N, M) \times M \times N^{d-1} \to M^d$. Then $ev_{\hat{\rho}}$ is $\mathcal{C}^r$ and

Lemma 5.3. *$T_{(f,x,y)} ev_{\hat{\rho}}$ is surjective at all $(f, x, y) \in \mathcal{D}^{2r}(M \times N, M) \times M \times \tilde{N}_{d-1}$, where $\tilde{N}_{d-1}$ is defined in (5.1). More precisely, given any $(u_0, u_1, \ldots, u_{d-1}) \in T_{x_0}M \times \cdots \times T_{x_{d-1}}M$, we can find a $\eta \in T_f \mathcal{D}^{2r}(M \times N, M)$ such that $T_{(f,x,y)} ev_{\hat{\rho}}(\eta, u_0, 0_y) = (u_0, u_1, \ldots, u_{d-1})$.*

Proof. If $y \in \tilde{N}_{d-1}$, then the points $\{y_0, y_1, \ldots, y_{d-2}\}$ are all distinct, and hence $\{(x_0, y_0), (x_1, y_1), \ldots, (x_{d-2}, y_{d-2})\}$ are distinct. By Corollary C.12 of [Stark, 1999], given any set of $u_i \in T_{x_i}M$ for $i = 1, \ldots, d-1$ we can find a $\eta \in T_f \mathcal{D}^{2r}(M \times N, M)$ such that $\eta(x_{i-1}, y_{i-1}) = u_i$ for $i = 1, \ldots, d-1$. So fix some $i \in \{1, \ldots, d-1\}$ and some $u_i \in T_{x_i}M$. Choose $\eta \in T_f \mathcal{D}^{2r}(M \times N, M)$ such that $\eta(x_{i-1}, y_{i-1}) = u_i, \eta(x_i, y_i) = -T_{(x_i,y_i)} f(u_i, 0)$ if $i \neq d-1$ and $\eta(x_j, y_j) = 0$ for $j \neq i-1, i$. Then by (4.6) $T_{(f,x,y)} ev_{\hat{\rho}_i}(\eta, 0_x, 0_y) = \bar{\eta}_i(x, y) = \eta(x_{i-1}, y_{i-1}) = u_i$ and if $i \neq d-1$ then $T_{(f,x,y)} ev_{\hat{\rho}_{i+1}}(\eta, 0_x, 0_y) = \eta(x_i, y_i) + T_{(x_i,y_i)} f(\bar{\eta}_i(x, y), 0) = -T_{(x_i,y_i)} f(u_i, 0) + T_{(x_i,y_i)} f(u_i, 0) = 0$. Furthermore $\bar{\eta}_j(x, y) = 0$ for $j \neq i-1, i$ and hence $T_{(f,x,y)} ev_{\hat{\rho}}$ $(\eta, 0_x, 0_y) = (0, \ldots, 0, u_i, 0, \ldots, 0) \in T_{x_0}M \times \cdots \times T_{x_i}M \times \cdots \times T_{x_{d-1}}M$. Hence by linearity, given any $(u_1, \ldots, u_{d-1}) \in T_{x_1}M \times \cdots \times T_{x_{d-1}}M$, we can find a $\eta \in T_f \mathcal{C}^{2r}(M \times N, M)$ such that $T_{(f,x,y)} ev_{\hat{\rho}}(\eta, 0_x, 0_y) = (0, u_1, \ldots, u_{d-1})$.

It remains to treat the first component. Note that the corresponding proof in Lemma 5.12 of [Stark, 1999] contains an erroneous calculation, but the argument used here is equally valid in that case. Recall that $f^{(0)}(x, y) = x$ and hence given any $u_0 \in T_{x_0}M$, we have by (4.6) $T_{(f,x,y)} ev_{\hat{\rho}}(0_f, u_0, 0_y) = (u_0, u_1, \ldots, u_{d-1})$, for some $u_1, \ldots, u_{d-1}$ whose precise values do not concern us (in fact $u_i = T_{(x,y)} f^{(i)}(u_0, 0)$ since if $\eta = 0$

we have $\bar{\eta}_i = 0$ by (4.5)). By the first part of the proof, choose $\eta \in T_f\mathcal{D}^{2r}(M \times N, M)$ such that $T_{(f,x,y)}ev_{\hat{\rho}}(\eta, 0_x, 0_y) = (0, u_1, \ldots, u_{d-1})$. Then $T_{(f,x,y)}ev_{\hat{\rho}_0}(-\eta, u_0, 0_y) = (u_0, 0, \ldots, 0)$, and hence by linearity $T_{(f,x,y)}ev_{\hat{\rho}}$ is surjective. $\square$

5.3. Immersivity of Φ

As in [Stark, 1999], the basic idea is to make ev_τ transversal to the zero section in $T\mathbb{R}^d$ and then count dimensions. Observe that if for some $v \in \tilde{T}_x M$ we have $T_x\Phi_{f,\varphi,I,y}(v) \neq 0$ for some I, then $T_x\Phi_{f,\varphi,y}(v) \neq 0$. Therefore to prove that $\Phi_{f,\varphi}$ is immersive at x, it is sufficient to show that for all $v \in \tilde{T}_x M$ we have $T_x\Phi_{f,\varphi,I,y}(v) \neq 0$ for some I. If we define the zero section L_I in $T\mathbb{R}^\alpha$ by

$$L_I = \{v \in TM : v = 0\},$$

then our aim is to show that the image of $T_x\Phi_{f,\varphi,I,y}$ does not intersect L_I. As in [Stark, 1999], we proceed by defining a codimension $(d - \alpha)m$ submanifold of M^d by

$$\Delta_I = \{(z_0, z_1, \ldots, z_{d-1}) \in M^d : z_i = z_j \text{ if and only if } i \sim_I j\}.$$

Recall that $\hat{\rho} = (\hat{\rho}_0, \hat{\rho}_1, \ldots, \hat{\rho}_{d-1}) : \mathcal{C}^{2r}(M \times N, M) \to \mathcal{C}^r(M \times N^{d-1}, M^d)$ with $\hat{\rho}_i$ as in (4.4), and that $ev_{\hat{\rho}}$ is $\mathcal{C}^r$ by Corollary 4.2. Let ι_τ be the inclusion ι_τ: $\mathcal{C}^r(M \times N^{d-1} M^d) \to \mathcal{C}^{r-1}(\tilde{T}M \times N^{d-1}, M^d)$ given by $\iota_\tau(F) = F \circ (\tau_M \times Id)$ where τ_M: $\tilde{T}M \to M$ is the tangent bundle projection, thus $\iota_\tau(F)(v, y) = F(x, y)$ where $x \in M$ is the point such that $v \in T_x M$, i.e., $x = \tau_M(v)$. Since ι_τ is just composition with fixed maps, it is $\mathcal{C}^r$. Now define τ'_I: $\mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R}) \to \mathcal{C}^{r-1}(\tilde{T}_M \times \tilde{N}_{d-1}, T\mathbb{R}^\alpha \times M^d)$ by

$$\tau'_I(f, \varphi) = (\tau_I(f, \varphi), \iota_\tau \circ \hat{\rho}(f)).$$

Note that we have restricted the domain of $\tau'_I(f, \varphi)$ to $\tilde{T}M \times \tilde{N}_{d-1}$, where $\tilde{N}_{d-1}$ is defined in (5.1). In other words we consider only those $y \in N^{d-1}$ in which no two components are the same. Since $\tilde{N}_{d-1}$ is not compact there is no natural manifold structure on $\mathcal{C}^{r-1}(\tilde{T}M \times \tilde{N}_{d-1}, T\mathbb{R}^\alpha \times M^d)$ and we cannot speak of τ'_I being smooth. However, it is only really the evaluation map $ev_{\tau'_I}$: $\mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R}) \times \tilde{T}M \times \tilde{N}_{d-1} \to T\mathbb{R}^\alpha \times M$ that we need. This is given by

$$\begin{aligned} ev_{\tau'_I}(f, \varphi, v, y) &= (ev_{\tau_I}(f, \varphi, v, y), ev_{\hat{\rho}}(f, \tau_M(v), y)) \\ &= (\tau_I(f, \varphi)(v, y), \hat{\rho}(f)(x, y)) \\ &= (T_x\Phi_{f,\varphi,I,y}(v), (f^{(0)}(x, y), f^{(1)}(x, y), \ldots, f^{(d-1)}(x, y))), \end{aligned}$$

where $x = \tau_M(v)$. Since ev_{τ_I} (by Lemma 4.4), $ev_{\hat{\rho}}$ (by Corollary 4.2), and τ_M are all $\mathcal{C}^{r-1}$, so is $ev_{\tau'_I}$. We now claim that

Proposition 5.4. *Given any partition I of $\{0, \ldots, d-1\}$, $T_{(f,\varphi,v,y)}ev_{\tau'_I}$ is surjective at all $(f, \varphi, v, y) \in \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R}) \times \tilde{T}M \times \tilde{N}_{d-1}$ such that $\hat{\rho}(f)(x, y) \in \Delta_I$. Hence, in particular $ev_{\tau'_I}$ is transversal to $L_I \times \Delta_I$.*

Proof. Suppose that $ev_{\tau'_I}(f, \varphi, v, y) \in L_I \times \Delta_I$. We aim to evaluate $T_\Xi ev_{\tau'_I}$, where $\Xi = (f, \varphi, v, y)$. We first consider the second component. Let $ev_{\hat{\rho}_\tau}$ be the evaluation map of $\iota_\tau \circ \hat{\rho}$, so that $ev_{\hat{\rho}_\tau}(f, v, y) = ev_{\hat{\rho}}(f, \tau_M(v), y)$. Thus by the chain rule $T_{(f,v,y)} ev_{\hat{\rho}_\tau}(\eta, w, w') = T_{(f,x,y)} ev_{\hat{\rho}}(\eta, T_v \tau_M(w), w')$ where $x = \tau_M(v)$. Let $\hat{\rho}(f)(x, y) = z \in \Delta_I$. Then by Lemma 5.3, given any $u = (u_0, u_1, \ldots, u_{d-1}) \in T_z \Delta_I$, there exists a $\eta \in T_f \mathcal{D}^{2r}(M \times N, M)$ such that $T_{(f,x,y)} ev_{\hat{\rho}}(\eta, u_0, 0_y) = u$. Furthermore τ_M is a submersion, i.e. $T_v \tau_M$ is surjective for all $v \in TM$. This can be shown using local coordinates; see for instance the paragraph following Lemma B.4 of [Stark, 1999]. Hence there exists a $w \in T_v(TM)$ such that $T_v \tau_M(w) = u_0$ and hence $T_{(f,v,y)} ev_{\hat{\rho}_\tau}(\eta, w, 0_y) = u$ (so that in particular $T_{(f,v,y)} ev_{\hat{\rho}_\tau}$ is surjective).

Turning to the first component, for any $\xi \in T_\varphi \mathcal{C}^{2r}(M, \mathbb{R})$ we have by linearity

$$\begin{aligned} T_\Xi ev_{\tau_I}(\eta, \xi, w, 0_y) &= T_\Xi ev_{\tau_I}(\eta, 0_\varphi, w, 0_y) + T_\Xi ev_{\tau_I}(0_\eta, \xi, 0_v, 0_y) \\ &= \tilde{u} + T_\Xi ev_{\tau_I}(0_\eta, \xi, 0_v, 0_y) \end{aligned}$$

for some $\tilde{u} \in T(T\mathbb{R}^\alpha)$, independent of ξ. By Corollary 5.2, $T_\Xi ev_{\tau_I}(0_f, \xi, 0_x, 0_y) = (\bar{\omega}(T_{x_1}\xi(v_{j_1})), \bar{\omega}(T_{x_2}\xi(v_{j_2})), \ldots, \bar{\omega}(T_{x_\alpha}\xi(v_{j_\alpha})))^\dagger$, where $x_i = f^{(i)}(x, y)$ and $v_i = T_{(x,y)} f^{(i)}(v, 0_y)$.

Since $ev_{\hat{\rho}_\tau}(f, v, y) = ev_{\hat{\rho}}(f, \tau_M(v), y) = \hat{\rho}(f)(x, y) \in \Delta_I$, the points $x_{j_1}, x_{j_2}, \ldots, x_{j_\alpha}$ are all distinct. For a fixed y, $f^{(j_i)}$ is a diffeomorphism and $\|v\| = 1$ and hence $v_{j_i} \neq 0$ for $i = 1, \ldots, \alpha$. Thus, by Corollary C.16 of [Stark, 1999], for any $\bar{u} \in T_0(T\mathbb{R}^\alpha)$ there exists a $\xi \in T_\varphi \mathcal{C}^{2r}(M, \mathbb{R})$ such that $T_\Xi ev_{\tau_I}(0_f, \xi, 0_x, 0_y) = \bar{u} - \tilde{u}$. Hence $T_\Xi ev_{\tau_I}(\eta, \xi, w, 0_y) = \tilde{u} + (\bar{u} - \tilde{u}) = \bar{u}$ and so $T_\Xi ev_{\tau'_I}(\eta, \xi, w, 0_y) = (\bar{u}, u)$. Thus $T_\Xi ev_{\tau'_I}$ is surjective, and in particular transversal to $L_I \times \Delta_I$, as required. $\square$

Observe that $\dim \tilde{T}M \times \tilde{N}_{d-1} - \operatorname{codim} L_I \times \Delta_I = 2m - 1 + n(d-1) - (d-\alpha)m - \alpha$, and

$$\begin{aligned} 2m - 1 + n(d-1) - (d-\alpha)m - \alpha &\leq d - 2 - (d-\alpha)m - \alpha + n(d-1) \\ &\leq (d-\alpha)(1-m) - 2 + n(d-1) \\ &< n(d-1) - 1. \end{aligned} \tag{5.5}$$

Recall that $r = n(d-1)$ and hence $\dim \tilde{T}M \times \tilde{N}_{d-1} - \operatorname{codim} L_I \times \Delta_I < r - 1$. Since $ev_{\tau'_I}$ is $\mathcal{C}^{r-1}$, it satisfies the smoothness condition of the Parametric Transversality Theorem ([Abraham, 1963]; [Abraham and Robbin, 1967] or see Appendix A of [Stark, 1999]). Thus there is a residual set of $(f, \varphi) \in \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R})$ for which $\tau'_I(f, \varphi)$ is transversal to $L_I \times \Delta_I$. Fix any (f, φ) in this set, and define $\tau'_{f,\varphi,I}\colon \tilde{N}_{d-1} \to \mathcal{C}^{r-1}(\tilde{T}M, T\mathbb{R}^\alpha \times M^d)$ by

$$\tau'_{f,\varphi,I}(y)(v) = \tau'_I(f, \varphi)(v, y),$$

so that $ev_{\tau'_{f,\varphi,I}}(v, y) = \tau'_I(f, \varphi)(v, y)$. Since by our choice of (f, φ), $\tau'_I(f, \varphi)$ is transversal to $L_I \times \Delta_I$, we see that $ev_{\tau'_{f,\varphi,I}}$ is transversal to $L_I \times \Delta_I$. Also, $ev_{\tau'_{f,\varphi,I}}(v, y) = ev_{\tau'_I}(f, \varphi, v, y)$ and so $ev_{\tau'_{f,\varphi,I}}$ is $\mathcal{C}^{r-1}$. Similar to above, we have $\dim \tilde{T}M - \operatorname{codim} L_I \times \Delta_I = 2m - 1 - (d-\alpha)m - \alpha$, and hence using (5.5), $\dim \tilde{T}M - \operatorname{codim} L_I \times \Delta_I < -1 <$

$r-1$. Therefore, by the Measure Theoretic Finite Dimensional Parametric Transversality Theorem (Appendix A), $\tau'_{f,\varphi,I}(y)$ is transversal to $L_I \times \Delta_I$ for a residual set of full Lebesgue measure of y in $\tilde{N}_{d-1}$.

The dimension of $\tilde{T}M$ is $2m-1$ and that of $L_I \times \Delta_I$ is $\alpha + dm - (d-\alpha)m = (m+1)\alpha$. Using (5.5) again, we have $2m-1 < (d-\alpha)m + \alpha$ and hence $\dim \tilde{T}M + \dim L_I \times \Delta_I = 2m-1+(m+1)\alpha < (d-\alpha)m + \alpha + (m+1)\alpha = 2\alpha + dm = \dim T\mathbb{R}^\alpha \times M^d$. Now, $\dim T_v(\tilde{T}M) = \dim \tilde{T}M$, and the dimension of $T_v[\tau'_{f,\varphi,I}(y)](T_v(\tilde{T}M))$ cannot be greater than that of $T_v(\tilde{T}M)$. Hence $\dim T_v[\tau'_{f,\varphi,I}(y)](T_v(\tilde{T}M)) \leq \dim \tilde{T}M$. Suppose that the image of $\tilde{T}M$ intersected $L_I \times \Delta_I$. Denote $\tau'_{f,\varphi,I}(y)(v) = z \in L_I \times \Delta_I$. Then $\dim T_z(L_I \times \Delta_I) = \dim L_I \times \Delta_I$ and hence $\dim T_v[\tau'_{f,\varphi,I}(y)](T_v(\tilde{T}M)) + \dim T_z(L_I \times \Delta_I) < 2\alpha + dm = \dim T_z(T\mathbb{R}^\alpha \times M^d)$. Hence $T_v[\tau'_{f,\varphi,I}(y)](T_v(\tilde{T}M))$ and $T_z(L_I \times \Delta_I)$ cannot together span $T_z(T\mathbb{R}^\alpha \times M^d)$ and thus the intersection cannot be transversal.

We thus see that if $\tau'_{f,\varphi,I}(y)$ is transversal to $L_I \times \Delta_I$, then its image cannot intersect $L_I \times \Delta_I$. Hence, either $T_x\Phi_{f,\varphi,I,y}(v) \notin L_I$ or $\hat{\rho}(f)(x,y) \notin \Delta_I$. But for every $(x,y) \in M \times \tilde{N}_{d-1}$ we have $\hat{\rho}(f)(x,y) \in \Delta_I$ for some I. For this choice of I we must thus have $T_x\Phi_{f,\varphi,I,y}(v) \notin L_I$. But if $T_x\Phi_{f,\varphi,I,y}(v) \notin L_I$, then $T_x\Phi_{f,\varphi,y}(v) \neq 0$. Hence for a residual set of full Lebesgue measure of y, we have $T_x\Phi_{f,\varphi,y}(v) \neq 0$ for all $v \in \tilde{T}M$, so that $\Phi_{f,\varphi,y}$ is immersive, as required.

5.4. Injectivity of Φ

This closely parallels section 6.3 of [Stark, 1999]. As indicated in section 3.3, the strategy is to make $\Phi_{f,\varphi} \times \Phi_{f,\varphi}$ transversal to the diagonal in $\mathbb{R}^d \times \mathbb{R}^d$. As in the previous section, we need to take particular care at points x such that $x_i = x_j$ with $i \neq j$, but we now additionally also have to consider pairs $(x,x') \in (M \times M)\backslash\Delta$ such that $x_i = x'_j$ (where Δ is the diagonal in $M \times M$). As in section 6.3 of [Stark, 1999], we proceed by letting R be a subset of $\mathcal{J}_I \times \mathcal{J}_I$ (possibly empty) and setting

$$\mathcal{J}_{I,R} = \{i \in \mathcal{J}_I : (i,i') \in R \text{ for any } i' \in \mathcal{J}_I\}.$$

For each choice of R, we aim to restrict to points $(x,x') \in (M \times M)\backslash\Delta$ such that $x_i = x'_{i'}$ if and only if $(i,i') \in R$. Let β be the number of elements in R and γ the number of elements in $\mathcal{J}_{I,R}$. Since $\mathcal{J}_{I,R}$ is obtained by removing at most β elements from $\mathcal{J}_I$, we have $\gamma \geq \alpha - \beta$. If $\gamma > 0$, write $\mathcal{J}_{I,R} = \{j'_1, j'_2, \ldots, j'_\gamma\}$, with $j'_1 < j'_2 < \ldots < j'_\gamma$. Our aim will be to ensure that the sets $\{x_{j'_1}, x_{j'_2}, \ldots, x_{j'_\gamma}\}$ and $\{x'_{j'_1}, x'_{j'_2}, \ldots, x'_{j'_\gamma}\}$ are distinct. As in sections 5.6 and 6.3 of [Stark, 1999], we thus define the map $\hat{\rho}'$: $\mathcal{D}^{2r}(M \times N, M) \to \mathcal{C}^r((M \times M)\backslash\Delta \times \tilde{N}_{d-1}, M^d \times M^d)$ by

$$\begin{aligned}\hat{\rho}'(f)(x,x',y) &= (\hat{\rho}(f)(x,y), \hat{\rho}(f)(x',y)) \\ &= ((f^{(0)}(x,y), \ldots, f^{(d-1)}(x,y)), (f^{(0)}(x',y), \ldots, f^{(d-1)}(x',y))),\end{aligned}$$

and the set

$$\begin{aligned}\Delta_{I,R} = \{&(z_0, \ldots, z_{d-1}, z'_0, \ldots, z'_{d-1}) \in M^d \times M^d \colon (z_0, \ldots, z_{d-1}) \in \Delta_I, \text{ and} \\ &\text{if } (i,i') \in \mathcal{J}_I \times \mathcal{J}_I \text{ then } z_i = z'_{i'} \text{ if and only if } (i,i') \in R\}.\end{aligned}$$

Then if $ev_{\hat{\rho}'}(f, x, x', y) = \hat{\rho}'(f)(x, x', y) \in \Delta_{I,R}$, we see that the points $\{x_{j_1'}, x_{j_2'}, \ldots, x_{j_\gamma'}\}$ are all distinct and the sets $\{x_{j_1'}, x_{j_2'}, \ldots, x_{j_\gamma'}\}$ and $\{x'_{j_1'}, x'_{j_2'}, \ldots, x'_{j_\gamma'}\}$ do not intersect. Observe that for points in $\Delta_{I,R}$ we may independently set α of the points $(z_0, \ldots, z_{d-1})$. The relation R then fixes β of the points $(z'_0, \ldots, z'_{d-1})$, leaving us free to fix the remaining $d - \beta$ points. Thus the dimension of $\Delta_{I,R}$ is $(d - \beta + \alpha)m$ or in other words its codimension in $M^d \times M^d$ is $(d - \alpha_R)m$, where $\alpha_R = \alpha - \beta$. Next, define the map $\Phi_{f,\varphi,I,R}\colon M \times N^{d-1} \to \mathbb{R}^\gamma$ by

$$\begin{aligned}\Phi_{f,\varphi,I,R}(x, y) &= (\varphi(f^{(j_1')}(x, y)), \varphi(f^{(j_2')}(x, y)), \ldots, \varphi(f^{(j_\gamma')}(x, y)))^\dagger \\ &= (\varphi(x_{j_1'}), \varphi(x_{j_2'}), \ldots, \varphi(x_{j_\gamma'}))^\dagger,\end{aligned}$$

with $\Phi_{f,\varphi,I,R}(x, y) = 0$ if $\gamma = 0$ (using the convention that $\mathbb{R}^0 = \{0\}$). Then, as usual, define $\Phi_{f,\varphi,y,I,R}\colon M \to \mathbb{R}^\gamma$ by

$$\Phi_{f,\varphi,y,I,R}(x) = \Phi_{f,\varphi,I,R}(x, y).$$

Observe that if $\Phi_{f,\varphi,y,I,R}(x) \neq \Phi_{f,\varphi,y,I,R}(x')$, then $\Phi_{f,\varphi,y}(x) \neq \Phi_{f,\varphi,y}(x')$. Thus if for all $(x, x') \in (M \times M)\backslash\Delta$ we have $\Phi_{f,\varphi,y,I,R}(x) \neq \Phi_{f,\varphi,y,I,R}(x')$ for some I, R, then $\Phi_{f,\varphi,y}$ is injective on M. Define $\rho_{I,R}\colon \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R}) \to \mathcal{C}^r((M \times M)\backslash\Delta \times \tilde{N}_{d-1}, \mathbb{R}^\gamma \times \mathbb{R}^\gamma)$ by

$$\rho_{I,R}(f, \varphi)(x, x', y) = (\Phi_{f,\varphi,I,R}(x, y), \Phi_{f,\varphi,I,R}(x', y)),$$

and $\rho'_{I,R}\colon \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R}) \to \mathcal{C}^r((M \times M)\backslash\Delta \times \tilde{N}_{d-1}, \mathbb{R}^\gamma \times \mathbb{R}^\gamma \times M^d \times M^d)$ by

$$\rho'_{I,R}(f, \varphi) = (\rho_{I,R}(f, \varphi), \hat{\rho}'(f)).$$

We thus have

$$\begin{aligned}ev_{\rho'_{I,R}}(f, \varphi, x, x', y) &= (ev_{\rho_{I,R}}(f, \varphi, x, x', y), ev_{\hat{\rho}'}(f, x, x', y)) \\ &= (\rho_{I,R}(f, \varphi)(x, x', y), \hat{\rho}'(f)(x, x', y)) \\ &= (\Phi_{f,\varphi,I,R}(x, y), \Phi_{f,\varphi,I,R}(x', y), \hat{\rho}'(f)(x, x', y)) \\ &= ((\varphi(f^{(j_1')}(x, y)), \varphi(f^{(j_2')}(x, y)), \ldots, \varphi(f^{(j_\gamma')}(x, y))), \\ &\quad (\varphi(f^{(j_1')}(x', y)), \varphi(f^{(j_2')}(x', y)), \ldots, \varphi(f^{(j_\gamma')}(x', y))), \\ &\quad (f^{(0)}(x, y), \ldots, f^{(d-1)}(x, y)), (f^{(0)}(x', y), \ldots, f^{(d-1)}(x', y)) \\ &= ((\varphi(x_{j_1'}), \varphi(x_{j_2'}), \ldots, \varphi(x_{j_\gamma'})), (\varphi(x'_{j_1'}), \varphi(x'_{j_2'}), \ldots, \varphi(x'_{j_\gamma'})), \\ &\quad (x_0, \ldots, x_{d-1}), (x'_0, \ldots, x'_{d-1})).\end{aligned}$$

By Lemma 4.3 $ev_{\rho_{I,R}}$ is $\mathcal{C}^r$ and by Corollary 4.2 $ev_{\hat{\rho}'}$ is $\mathcal{C}^r$, and hence $ev_{\rho'_{I,R}}$ is $\mathcal{C}^r$. Finally, let Δ_γ be the diagonal in $\mathbb{R}^\gamma \times \mathbb{R}^\gamma$.

Proposition 5.5. *Given any partition I of $\{0, \ldots, d - 1\}$ and any subset R of $\mathcal{J}_I \times \mathcal{J}_I$ (possibly empty), $ev_{\rho'_{I,R}}$ is transversal to $\Delta_\gamma \times \Delta_{I,R}$ for all I, R.*

Proof. Denote $\Xi = (f, \varphi, x, x', y)$ and suppose that $ev_{\rho'_{I,R}}(\Xi) \in \Delta_\gamma \times \Delta_{I,R}$, for some $(x, x', y) \in (M \times M \backslash \Delta) \times \tilde{N}_{d-1}$. As in Proposition 5.4, we first consider the second component. For $y \in \tilde{N}_{d-1}$ the points $\{y_0, y_1, \ldots, y_{d-2}\}$ are disjoint, and hence $(x_i, y_i) \neq (x_{i'}, y_{i'})$, $(x'_i, y_i) \neq (x_{i'}, y_{i'})$, and $(x_i, y_i) \neq (x'_{i'}, y_{i'})$ for all $i \neq i'$. Furthermore $x \neq x'$ (recall we only consider $(x, x') \in (M \times M \backslash \Delta)$) and f_{y_i} is a diffeomorphism for all $y_i \in N$ and hence $(x_i, y_i) \neq (x'_i, y_i)$. Hence $\{(x_0, y_0), (x_1, y_1), \ldots, (x_{d-2}, y_{d-2}), (x'_0, y'_0), (x'_1, y'_1), \ldots, (x'_{d-2}, y'_{d-2})\}$ are disjoint. By Corollary C.12 of [Stark, 1999], given any set of $u_i \in T_{x_i}M$ and $u'_i \in T_{x'_i}M$ for $i = 1, \ldots, d-1$, we can find a $\eta \in T_f\mathcal{D}^{2r}(M \times N, M)$ such that $\eta(x_{i-1}, y_{i-1}) = u_i$ and $\eta(x'_{i-1}, y_{i-1}) = u'_i$ for $i = 1, \ldots, d-1$. Hence by a straightforward extension of the proof of Lemma 5.3, given any $(u_1, \ldots, u_{d-1}) \in T_{x_1}M \times \ldots \times T_{x_{d-1}}M$ and $(u'_1, \ldots, u'_{d-1}) \in T_{x'_1}M \times \ldots \times T_{x'_{d-1}}M$ we can find a $\eta \in T_f\mathcal{C}^{2r}(M \times N, M)$ such that $T_{(f,x,x',y)}ev_{\hat{\rho}'}(\eta, 0_x, 0_{x'}, 0_y) = ((0, u_1, \ldots, u_{d-1}), (0, u'_1, \ldots, u'_{d-1}))$. On the other hand, as in the second part of the proof of Lemma 5.3, $T_{(f,x,x',y)}ev_{\hat{\rho}'}(0_\eta, u_0, 0_{x'}, 0_y) = ((u_0, u_1, \ldots, u_{d-1}), (0, \ldots, 0))$ and hence given any $u_0 \in T_{x_0}M$, there exists $\eta \in T_f\mathcal{C}^{2r}(M \times N, M)$ such that $T_{(f,x,x',y)}ev_{\hat{\rho}'}(\eta, u_0, 0_{x'}, 0_y) = ((u_0, 0, \ldots, 0), (0, \ldots, 0))$. Similarly given any $u'_0 \in T_{x'_0}M$, there exists $\eta \in T_f\mathcal{C}^{2r}(M \times N, M)$ such that $T_{(f,x,x',y)}ev_{\hat{\rho}'}(\eta, 0_x, u'_0, 0_y) = ((0, \ldots, 0), (u'_0, 0, \ldots, 0))$. Hence by linearity given any $(u_0, \ldots, u_{d-1}) \in T_{x_0}M \times \ldots \times T_{x_{d-1}}M$ and $(u'_0, \ldots, u'_{d-1}) \in T_{x'_0}M \times \ldots \times T_{x'_{d-1}}M$, we can find a $\eta \in T_f\mathcal{C}^{2r}(M \times N, M)$ such that $T_{(f,x,x',y)}ev_{\hat{\rho}'}(\eta, u_0, u'_0, 0_y) = ((u_0, \ldots, u_{d-1}), (u'_0, \ldots, u'_{d-1}))$ (so that in particular $T_{(f,x,x',y)}ev_{\hat{\rho}'}$ is surjective).

Now let us turn to the other component $T_\Xi ev_{\rho_{I,R}}$ of $T_\Xi ev_{\rho'_{I,R}}$. If $\gamma = 0$, then $T(\mathbb{R}^\gamma \times \mathbb{R}^\gamma)$ is a single point, and hence the surjectivity of $T_{(f,x,x',y)}ev_{\hat{\rho}'}$ is sufficient to ensure the surjectivity of $T_\Xi ev_{\rho'_{I,R}}$. If $\gamma > 0$, then for any $\xi \in T_\varphi\mathcal{C}^{2r}(M, \mathbb{R})$ we have by linearity

$$\begin{aligned}
T_\Xi ev_{\rho_{I,R}}(\eta, \xi, u_0, u'_0, 0_y) &= T_\Xi ev_{\rho_{I,R}}(\eta, 0_\varphi, u_0, u'_0, 0_y) + T_\Xi ev_{\rho_{I,R}}(0_\eta, \xi, 0_x, 0_{x'}, 0_y) \\
&= (\tilde{u}, \tilde{u}') + T_\Xi ev_{\rho_{I,R}}(0_\eta, \xi, 0_x, 0_{x'}, 0_y),
\end{aligned}$$

for some $(\tilde{u}, \tilde{u}') \in T_{(z,z)}(\mathbb{R}^\gamma \times \mathbb{R}^\gamma) = T_z\mathbb{R}^\gamma \times T_z\mathbb{R}^\gamma$, where $\rho_{I,R}(f, \varphi)(x, x', y) = (z, z) \in \Delta_\gamma$. Note that $(\tilde{u}, \tilde{u}')$ is independent of ξ. By (4.9), $T_\Xi ev_{\rho_{I,R}}(0_\eta, \xi, 0_x, 0_{x'}, 0_y) = ((\xi(x_{j'_1}), \xi(x_{j'_2}), \ldots, \xi(x_{j'_\gamma})), (\xi(x'_{j'_1}), \xi(x'_{j'_2}), \ldots, \xi(x'_{j'_\gamma})))$. Since $\hat{\rho}'(f)(x, x', y) \in \Delta_{I,R}$, the points $x_{j_1}, x_{j_2}, \ldots, x_{j_\alpha}$ are all distinct and $\{x_{j_1}, x_{j_2}, \ldots, x_{j_\alpha}\} \cap \{x'_{j'_1}, x'_{j'_2}, \ldots, x'_{j'_\gamma}\} = \emptyset$. Thus, as in the proof of Proposition 6.2 of [Stark, 1999], given any $(\bar{u}, \bar{u}') \in T_z\mathbb{R}^\gamma \times T_z\mathbb{R}^\gamma$, by Corollary C.12 of [Stark, 1999] there exists a $\xi \in T_\varphi\mathcal{C}^{2r}(M, \mathbb{R})$ such that $T_\Xi ev_{\rho_{I,R}}(0_\eta, \xi, 0_x, 0_{x'}, 0_y) = (\bar{u} - \bar{u}', 0_z) - (\tilde{u} - \tilde{u}', 0_z)$. Then $T_\Xi ev_{\rho_{I,R}}(\eta, \xi, u_0, u'_0, 0_y) = (\tilde{u}, \tilde{u}') + (\bar{u} - \bar{u}', 0_z) - (\tilde{u} - \tilde{u}', 0_z) = (\bar{u} - \bar{u}' + \tilde{u}', \tilde{u}')$. But $(\bar{u}', \bar{u}') \in T_{(z,z)}\Delta_\gamma$ and $(\tilde{u}', \tilde{u}') \in T_{(z,z)}\Delta_\gamma$, and hence $(\bar{u}, \bar{u}') \in \text{Image}T_\Xi ev_{\rho_{I,R}} + T_{(z,z)}\Delta_\gamma$. Thus Image $T_\Xi ev_{\rho_{I,R}} + T_{(z,z)}\Delta_\gamma = T_{(z,z)}(\mathbb{R}^\gamma \times \mathbb{R}^\gamma)$, and since as we have shown above, the second component of $T_\Xi ev_{\rho'_{I,R}}$ is surjective, this implies that $ev_{\rho'_{I,R}}$ is transversal to $\Delta_\gamma \times \Delta_{I,R}$ as required. □

Now, just as in the proof of immersivity in section 5.3, all that remains is to count dimensions and apply the Parametric Transversality Theorem. We have $\dim(M \times M)\backslash\Delta\times$

$\tilde{N}_{d-1} - \operatorname{codim} \Delta_\gamma \times \Delta_{I,R} = 2m + n(d-1) - \gamma - (d-\alpha_R)m$. Since $\gamma \geq \alpha_R$, we have

$$\begin{aligned} 2m + n(d-1) - \gamma - (d-\alpha_R)m &\leq d - 1 - (d-\alpha_R)m - \alpha_R + n(d-1) \\ &\leq (d-\alpha_R)(1-m) - 1 + n(d-1) \\ &< n(d-1). \end{aligned}$$

Recall that $r = n(d-1)$ and hence $\dim(M \times M)\backslash\Delta \times \tilde{N}_{d-1} - \operatorname{codim} \Delta_\gamma \times \Delta_{I,R} < r$. Since $ev_{\rho'_{I,R}}$ is $\mathcal{C}^r$, it satisfies the smoothness condition of the Parametric Transversality Theorem ([Abraham, 1963]; [Abraham and Robbin, 1967] or see Appendix A of [Stark, 1999]). Thus there is a residual set of $(f, \varphi) \in \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R})$ for which $\rho'_{I,R}(f, \varphi)$ is transversal to $\Delta_\gamma \times \Delta_{I,R}$. Fix any (f, φ) in this set, and define $\rho'_{f,\varphi,I,R}\colon \tilde{N}_{d-1} \to \mathcal{C}^r((M \times M)\backslash\Delta, \mathbb{R}^\gamma \times \mathbb{R}^\gamma \times M^d \times M^d)$ by

$$\rho'_{f,\varphi,I,R}(y)(x, x') = \rho'_{I,R}(f, \varphi)(x, x', y).$$

Then

$$ev_{\rho'_{f,\varphi,I,R}}(x, x', y) = \rho'_{f,\varphi,I,R}(y)(x, x').$$

Since by our choice of (f, φ), $\rho'_{I,R}(f, \varphi)$ is transversal to $\Delta_\gamma \times \Delta_{I,R}$, we see that $ev_{\rho'_{f,\varphi,I,R}}$ is transversal to $\Delta_\gamma \times \Delta_{I,R}$. Also, $ev_{\rho'_{f,\varphi,I,R}}$ is $\mathcal{C}^r$ and $\dim(M \times M)\backslash\Delta - \operatorname{codim} \Delta_\gamma \times \Delta_{I,R} = 2m - \gamma - (d-\alpha_R)m$. Hence, using (5.6) we have $\dim(M \times M)\backslash\Delta - \operatorname{codim} \Delta_\gamma \times \Delta_{I,R} < 0$. Thus by the Measure Theoretic Finite Dimensional Parametric Transversality Theorem (Appendix A), $\rho'_{f,\varphi,I,R}(y)$ is transversal to $\Delta_\gamma \times \Delta_{I,R}$ for a residual set of full Lebesgue measure of y in $\tilde{N}_{d-1}$.

The dimension of $(M \times M)\backslash\Delta$ is $2m$ and that of $\Delta_\gamma \times \Delta_{I,R}$ is $\gamma + 2dm - (d-\alpha_R)m$. Using (5.6) again, we have $2m - (d-\alpha_R)m < \gamma$ and hence $\dim(M \times M)\backslash\Delta + \dim \Delta_\gamma \times \Delta_{I,R} = 2m + \gamma + 2dm - (d-\alpha_R)m < 2\gamma + 2dm = \dim \mathbb{R}^\gamma \times \mathbb{R}^\gamma \times M^d \times M^d$. Now, $\dim T_{(x,x')}(M \times M)\backslash\Delta = \dim(M \times M)\backslash\Delta$, and the dimension of $T_{(x,x')}[\rho'_{f,\varphi,I,R}(y)](T_{(x,x')}(M \times M)\backslash\Delta)$ cannot be greater than that of $T_{(x,x')}(M \times M)\backslash\Delta$. Suppose that the image of $(M \times M\backslash\Delta)$ intersected $\Delta_\gamma \times \Delta_{I,R}$. Denote $\rho'_{f,\varphi,I,R}(y)(x, x') = z \in \Delta_\gamma \times \Delta_{I,R}$. Then $\dim T_z(\Delta_\gamma \times \Delta_{I,R}) = \dim \Delta_\gamma \times \Delta_{I,R}$ and hence $\dim T_{(x,x')} [\rho'_{f,\varphi,I,R}(y)](T_{(x,x')}(M \times M)\backslash\Delta) + \dim T_z(\Delta_\gamma \times \Delta_{I,R}) < 2\gamma + 2dm = \dim \mathbb{R}^\gamma \times \mathbb{R}^\gamma \times M^d \times M^d$. Hence $T_{(x,x')}[\rho'_{f,\varphi,I,R}(y)](T_{(x,x')}(M \times M)\backslash\Delta)$ and $T_z(\Delta_\gamma \times \Delta_{I,R})$ cannot together span $T_z(\mathbb{R}^\gamma \times \mathbb{R}^\gamma \times M^d \times M^d)$ and thus the intersection cannot be transversal.

We thus see that if $\rho'_{f,\varphi,I,R}(y)$ is transversal to $\Delta_\gamma \times \Delta_{I,R}$, then its image cannot intersect $\Delta_\gamma \times \Delta_{I,R}$. Hence, either $(\Phi_{f,\varphi,I,R}(x, y), \Phi_{f,\varphi,I,R}(x', y)) \notin \Delta_\gamma$ or $\hat{\rho}'(f)(x, x', y) \notin \Delta_{I,R}$. But for every $(x, x', y) \in (M \times M)\backslash\Delta \times \tilde{N}_{d-1}$ we have $\hat{\rho}'(f)(x, x', y) \notin \Delta_{I,R}$ for some I, R. For this choice of I and R we must thus have $(\Phi_{f,\varphi,I,R}(x, y), \Phi_{f,\varphi,I,R}(x', y)) \notin \Delta_\gamma$. But by definition this means that $\Phi_{f,\varphi,I,R}(x, y) \neq \Phi_{f,\varphi,I,R}(x', y)$ and hence $\Phi_{f,\varphi} (x, y) \neq \Phi_{f,\varphi}(x', y)$. Hence for a residual set of full Lebesgue measure of y, we have $\Phi_{f,\varphi}(x, y) \neq \Phi_{f,\varphi}(x', y)$ for all $(x, x') \in (M \times M)\backslash\Delta$, so that $\Phi_{f,\varphi,y}$ is injective, as required. $\square$

5.5. Lower Degrees of Smoothness

We can deduce Theorem 3.4 in the case where f and φ are $\mathcal{C}^k$ for $1 \leq k < 2n(d-1)$ using an argument identical to that in section 6.4 of [Stark, 1999]. Fix k such that

$1 \leq k < 2n(d-1)$ and denote $\mathcal{B}^k = \mathcal{D}^k(M \times N, M) \times \mathcal{C}^k(M, \mathbb{R})$. For $(f, \varphi) \in \mathcal{B}^k$ let $\tilde{N}(f, \varphi) \subset N^{d-1}$ be the set of y such that $\Phi_{f,\varphi,y}$ is an embedding, and let $\mathcal{B}(f, \varphi) = N^{d-1} \backslash \tilde{N}(f, \varphi)$. Note that since $\Phi_{f,\varphi,y}$ depends continuously on y and embeddings are open in $\mathcal{C}^k(M, \mathbb{R}^d)$, the set $\tilde{N}(f, \varphi)$ is necessarily open. Given an $\epsilon > 0$, define the ϵ-neighbourhood of $B(f, \varphi)$ by $B(f, \varphi, \epsilon) = \{y \in N^{d-1} : \tilde{d}(y, B(f, \varphi)) < \epsilon\}$, where $\tilde{d}$ is the metric on $\tilde{N}(f, \varphi)$.

Let $\mathcal{E}^k$ be the set of (f, φ) in $\mathcal{B}^k$ for which $v(\tilde{N}(f, \varphi)) = 1$, where v is (normalized) Lebesgue measure on N^{d-1}. By sections 5.3 and 5.4, $\mathcal{E}^{n(d-1)}$ is dense in $\mathcal{B}^{n(d-1)}$. Since the latter is dense in $\mathcal{B}^k$, we see that $\mathcal{E}^{n(d-1)}$ is dense in $\mathcal{B}^k$. But $\mathcal{E}^{n(d-1)} \subset \mathcal{E}^k$, and hence $\mathcal{E}^k$ is dense in $\mathcal{B}^k$. This space is separable and hence we may choose a countable set $\{(f_i, \varphi_i) \in \mathcal{E}^k : i \in \mathbb{N}\}$ which is dense in $\mathcal{B}^k$. For any i, we have

$$B(f_i, \varphi_i) = \bigcap_{\epsilon > 0} B(f_i, \varphi_i, \epsilon),$$

and since $\mu(B(f_i, \varphi_i)) = 0$, we have $v(B(f_i, \varphi_i, \epsilon)) \to 0$ as $\epsilon \to 0$. Hence given $\delta > 0$, we may choose an $\epsilon(i, \delta) > 0$ such that $v(N^{d-1} \backslash B(f_i, \varphi_i, \epsilon(i, \delta))) > 1 - \delta$. Now, $N^{d-1} \backslash B(f_i, \varphi_i, \epsilon(i, \delta))$ is closed and hence compact. Using the continuity of $\Phi_{f,\varphi,y}$ with respect to f, φ, and y and the density of embeddings in $\mathcal{C}^k(M, \mathbb{R}^d)$, we can find an open neighbourhood $\mathcal{N}(f_i, \varphi_i, \delta)$ of (f_i, φ_i) in $\mathcal{B}^k$ such that $\Phi_{f,\varphi,y}$ is an embedding for all $(f, \varphi) \in \mathcal{N}(f_i, \varphi_i, \delta)$ and $y \in N^{d-1} \backslash B(f_i, \varphi_i, \epsilon(i, \delta))$. Since $\{(f_i, \varphi_i) : i \in \mathbb{N}\}$ is dense in $\mathcal{B}^k$, the union of $\mathcal{N}(f_i, \varphi_i, \delta)$ is open and dense in $\mathcal{B}^k$. Thus if we let δ_n be a sequence such that $\delta_n \to 0$ and define

$$\mathcal{N} = \bigcap_{n \in \mathbb{N}} \bigcup_{i \in \mathbb{N}} \mathcal{N}(f_i, \varphi_i, \delta_n),$$

we see that $\mathcal{N}$ is residual. Now if $(f, \varphi) \in \mathcal{N}$, then there exists a sequence i_n such that $(f, \varphi) \in \mathcal{N}(f_{i_n}, \varphi_{i_n}, \delta_n)$ for each n. Thus $\Phi_{f,\varphi,y}$ is an embedding for all $y \in N^{d-1} \backslash B(f_{i_n}, \varphi_{i_n}, \epsilon(i, \delta_n))$ for each n. Hence if we define

$$\tilde{N}'(f, \varphi) = \bigcup_{n \in \mathbb{N}} N^{d-1} \backslash B(f_{i_n}, \varphi_{i_n}, \epsilon(i, \delta_n)),$$

then $\tilde{N}'(f, \varphi) \subset \tilde{N}(f, \varphi)$. Since $v(N^{d-1} \backslash B(f_{i_n}, \varphi_{i_n}, \epsilon(i, \delta_n))) \to 1$ as $n \to \infty$, we have $v(\tilde{N}'(f, \varphi)) = 1$ and hence $v(\tilde{N}(f, \varphi)) = 1$. Thus $(f, \varphi) \in \mathcal{E}^k$, and so $\mathcal{N} \subset \mathcal{E}^k$. Therefore $\mathcal{E}^k$ contains a residual set, as required.

6. Proof of Theorem 2.4

As in the proof for $\dim N > 0$, we have $\Phi_{f,\varphi,\omega}(x) = \Phi_{f,\varphi,y}(x)$ where $y = (\omega_0, \ldots, \omega_{d-2}) \in N^{d-1}$ and it is thus sufficient to show that for an open dense set of $(f, \varphi) \in \mathcal{D}^r(M \times N, M) \times \mathcal{C}^r(M, \mathbb{R})$ the map $\Phi_{f,\varphi,y}$ is an embedding for every $y \in N^{d-1}$. As usual, we first prove the theorem for a sufficiently large r, and then show that this implies the theorem for all $r \geq 1$. Unlike in the previous section for $\dim N > 0$, the latter argument turns out to be elementary in the present context (as it is in the standard Takens' Theorem). In particular, embeddings are open in $\mathcal{C}^r(M, \mathbb{R}^d)$ (e.g. [Hirsch, 1976]), and for a fixed y, $\Phi_{f,\varphi,y}$ depends continuously on f and φ. Thus, for a fixed y the set of (f, φ) for which $\Phi_{f,\varphi,y}$ is an embedding is open in $\mathcal{D}^r(M \times N, M) \times \mathcal{C}^r(M, \mathbb{R})$ for any $r \geq 1$. If $\dim N = 0$, there is only a finite number of y in N^{d-1}, and hence the set of (f, φ) for which $\Phi_{f,\varphi,y}$ is an embedding for all $y \in N^{d-1}$ is also open in $\mathcal{D}^r(M \times N, M) \times \mathcal{C}^r(M, \mathbb{R})$. On the other hand $\mathcal{D}^r(M \times N, M) \times \mathcal{C}^r(M, \mathbb{R})$ is dense in $\mathcal{D}^1(M \times N, M) \times \mathcal{C}^1(M, \mathbb{R})$ for any $r \geq 1$. Hence, if we show that for some $r \geq 1$, the set of (f, φ) for which $\Phi_{f,\varphi,y}$ is an embedding for all $y \in N^{d-1}$ is dense in $\mathcal{D}^r(M \times N, M) \times \mathcal{C}^r(M, \mathbb{R})$, then this set is also dense in $\mathcal{D}^1(M \times N, M) \times \mathcal{C}^1(M, \mathbb{R})$, as required. It turns out that as in the proof of the standard Takens' Theorem in [Stark, 1999], it is sufficient to take $r = 3$. However, in order to use the calculations in section 4, and to be consistent with the proof for $\dim N > 0$, we shall work in $\mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R})$ with $r = 2$.

As already indicated in section 3.4, we begin by showing that for an open dense set of f the short periodic orbits are isolated, are distinct for different sequences, and have distinct eigenvalues. As in the proof of the standard Takens' Theorem, we shall concern ourselves with sequences of length less than $2d$.

6.1. Generic Properties of Periodic Orbits for IFS'

We begin by giving a precise definition of a periodic orbit, and its minimal period.

Definition 6.1. Given a finite sequence $(y_0, y_1, \ldots, y_{q-1}) \in N^q$, let p be the least integer with $0 < p \leq q$ for which $y_{i+p} = y_i$ for all $i = 0, \ldots, q-p-1$, or in words so that we can write $(y_0, y_1, \ldots, y_{q-1})$ as k repeats of $(y_0, y_1, \ldots, y_{p-1})$, where $q = kp$. We shall denote this by $(y_0, \ldots, y_{p-1}, y_0, \ldots, y_{p-1}, \ldots, y_0, \ldots, y_{p-1}) = (y_0, y_1, \ldots, y_{p-1})^k$. We call p the *prime length* of $(y_0, y_1, \ldots, y_{q-1})$ and refer to the initial sequence $(y_0, y_1, \ldots, y_{p-1})$ of length p as the *prime factor* of $(y_0, y_1, \ldots, y_{q-1})$. The sequence $(y_0, y_1, \ldots, y_{q-1}) \in N^q$ is called *prime* if $p = q$.

By a periodic orbit of $f \in \mathcal{D}^{2r}(M \times N^{d-1}, M)$ we mean a periodic orbit of the skew product (2.2) in the usual sense. It thus consists of an orbit $\{(x_i, y_i)\}$ such that $(x_{i+q}, y_{i+q}) = (x_i, y_i)$ for all $i \in \mathbb{Z}$, where q is the period of the orbit. Any periodic orbit is thus associated with a finite sequence $(y_0, y_1, \ldots, y_{q-1}) \in N^q$ and can be defined by

Definition 6.2. A *periodic point* of a sequence $(y_0, y_1, \ldots, y_{q-1}) \in N^q$ is a fixed point of $f_{y_{q-1}\ldots y_0}$, that is a point $x \in M$ such that $x_q = f_{y_{q-1}\ldots y_0}(x) = x$. The *minimal period* of such an orbit is kp, where k is the least positive integer such that $(f_{y_{p-1}\ldots y_0})^k(x) = x$ and $(y_0, y_1, \ldots, y_{p-1})$ is the prime factor of $(y_0, y_1, \ldots, y_{q-1})$. A periodic orbit is *hyperbolic* if $T_x f_{y_{q-1}\ldots y_0}$ has no eigenvalues on the unit circle.

Denote the set of periodic orbits for a particular sequence $(y_0, y_1, \ldots, y_{q-1})$ by

$$\mathcal{P}_{y_0\ldots y_{q-1}} = \{x \in M : f_{y_{q-1}\ldots y_0}(x) = x\}.$$

Note that unlike the standard case, it is not necessarily true that if $x \in \mathcal{P}_{y_0\ldots y_{q-1}}$ then $x_i \in \mathcal{P}_{y_0\ldots y_{q-1}}$ where $x_i = f_{y_{i-1}\ldots y_0}(x)$. Instead, we trivially have

Lemma 6.3. *Suppose that* x *is a periodic orbit of* $(y_0, y_1, \ldots, y_{q-1}) \in N^q$. *Then* x_i *is a periodic orbit for* $(y_i, y_{i+1}, \ldots, y_{q-1}, y_0, \ldots, y_{i-1}) \in N^q$.

Proof. By definition $x_0 = x_q = f_{y_{q-1}\dots y_i}(x_i)$ and $x_i = f_{y_{i-1}\dots y_0}(x_0)$. Thus $f_{y_{i-1}\dots y_0 y_{q-1}\dots y_i}(x_i) = x_i$. □

The next lemma shows that the minimal period divides q and that the orbit segment $\{x_0, x_1, \dots, x_{q-1}\}$ is just the initial segment $\{x_0, x_1, \dots, x_{kp-1}\}$ repeated r times, where $q = kpr$. This is well known for standard dynamical systems, but perhaps not entirely obvious in the present setting.

Lemma 6.4. *Suppose that x is a periodic orbit of* $(y_0, y_1, \dots, y_{q-1}) \in N^q$ *with minimal period kp. Then* $x_i = x_{(i \bmod kp)}$ *and* $y_i = y_{(i \bmod kp)}$ *for all i with* $0 \le i \le q$.

Proof. Write $q = kpr + j$, for $0 \le j < kp$, so that $j = q \bmod kp$. By definition, $(f_{y_{p-1}\dots y_0})^k(x_0) = x_0$, and since $(y_0, y_1, \dots, y_{kpr-1}) = (y_0, y_1, \dots, y_{p-1})^{kr}$, we have $x_{kpr} = f_{y_{kpr-1}\dots y_0}(x_0) = (f_{y_{p-1}\dots y_0})^{kr}(x_0) = x_0$. Hence $f_{y_{q-1}\dots y_{kpr}}(x_0) = f_{y_{q-1}\dots y_{kpr}}(x_{kpr}) = x_q = x_0$. Since q is divisible by p, so is j, say $j = k'p$, and $(y_{kpr}, y_{kpr+1}, \dots, y_{q-1}) = (y_0, y_1, \dots, y_{p-1})^{k'}$. Thus $(f_{y_{p-1}\dots y_0})^{k'}(x_0) = x_0$, with $k' < k$. But, by definition k is the least integer $k > 0$ such that $(f_{y_{p-1}\dots y_0})^k(x_0) = x_0$ and hence we must have $k' = 0$. Thus $(y_0, y_1, \dots, y_{q-1})$ consists of r' repeats of $(y_0, y_1, \dots, y_{kp-1})$ and $x_i = f_{y_{i-1}\dots y_0}(x_0) = f_{y_{i-1}\dots y_{kpr'}}(f_{y_{kp-1}\dots y_0})^{r'}(x_0) = f_{y_{i-1}\dots y_{kpr'}}(x_0)$ where $i = kpr' + j'$, with $j' = i \bmod kp$. But $(y_{kpr'}, y_{kpr'+1}, \dots, y_{i-1})$ consists of the initial j' symbols of $(y_0, y_1, \dots, y_{p-1})$; in other words $(y_{kpr'}, y_{kpr'+1}, \dots, y_{i-1}) = (y_0, y_1, \dots, y_{j'-1})$. Thus $x_i = f_{y_{j'-1}\dots y_0}(x_0) = x_{j'}$, as required. □

Note however, that unlike in the usual definition for a single f, it is possible to have $f_{y_{i-1}\dots y_0}(x) = x$ for some i such that $0 < i < kp$. This poses a serious obstacle in trying to generalize the argument of section 4.2 of [Stark, 1999] to the present case, and much of the following construction is designed to overcome this difficulty. Thus define

$$
\begin{aligned}
\mathcal{B}_1^{(q)} &= \{f \in \mathcal{D}^{2r}(M \times N^{d-1}, M) : \mathcal{P}_{y_0\dots y_{i-1}} \text{ consists of a finite number of points} \\
&\qquad \text{for all } (y_0, y_1, \dots, y_{i-1}) \in N^i \text{ for all } 0 < i < q\},\\
\mathcal{B}_2^{(q)} &= \{f \in \mathcal{D}^{2r}(M \times N^{d-1}, M) : \text{ if } \mathcal{P}_{y_0\dots y_{i-1}} \cap \mathcal{P}_{y'_0\dots y'_{i'-1}} \neq \emptyset \text{ for some } (y_0, y_1, \dots,\\
&\qquad y_{i-1}) \neq (y'_0, y'_1, \dots, y'_{i'-1}) \text{ with } 0 < i, i' < q, \text{ then } (y_0, y_1, \dots, y_{i-1}) \text{ and}\\
&\qquad (y'_0, y'_1, \dots, y'_{i'-1}) \text{ have the same prime factor}\},\\
\mathcal{B}_3^{(q)} &= \{f \in \mathcal{D}^{2r}(M \times N^{d-1}, M) : \text{ all } x \in \mathcal{P}_{y_0\dots y_{i-1}} \text{ are hyperbolic}\\
&\qquad \text{for all } (y_0, y_1, \dots, y_{i-1}) \in N^i \text{ for all } 0 < i < q\},
\end{aligned}
$$

and finally

$$\mathcal{B}^{(q)} = \mathcal{B}_1^{(q)} \cap \mathcal{B}_2^{(q)} \cap \mathcal{B}_3^{(q)},$$

with the convention that $\mathcal{B}^{(1)} = \mathcal{D}^{2r}(M \times N^{d-1}, M)$. Note that if $(y_0, y_1, \dots, y_{i-1})$ and $(y'_0, y'_1, \dots, y'_{i'-1})$ have the same prime factor and $x \in (\mathcal{P}_{y_0\dots y_{i-1}} \cap \mathcal{P}_{y'_0\dots y'_{i'-1}})$, then x is the same periodic orbit under both $(y_0, y_1, \dots, y_{i-1})$ and $(y'_0, y'_1, \dots, y'_{i'-1})$. The following lemma is trivial, but is the key step in our proof of Theorem 3.4 for $\dim N = 0$:

Lemma 6.5. *Suppose that $f \in \mathcal{B}^{(q)}$ and for some $(y_0, y_1, \ldots, y_{q-1}) \in N^q$ and some $x \in M$ we have $x_q = x_i = x_0$ for some $0 < i < q$. Then $(y_0, y_1, \ldots, y_{i-1})$ and $(y_i, y_{i-1}, \ldots, y_{q-1})$ are both powers (i.e. repeats) of the prime factor of $(y_0, y_1, \ldots, y_{q-1})$, and i is a multiple of the prime length of $(y_0, y_1, \ldots, y_{q-1})$.*

Proof. We have $f_{y_{i-1}\ldots y_0}(x_0) = x_0$ and $f_{y_{q-1}\ldots y_i}(x_0) = f_{y_{q-1}\ldots y_i}(x_i) = x_i = x_0$. Thus $x_0 \in (\mathcal{P}_{y_0\ldots y_{i-1}} \cap \mathcal{P}_{y_i\ldots y_{q-1}})$, and since both $0 < i < q$ and $0 < q-i < q$, the definition of $\mathcal{B}^{(q)}$ implies that $(y_0, y_1, \ldots, y_{i-1})$ and $(y_i, y_{i-1}, \ldots, y_{q-1})$ have the same prime factor, which we denote $(y_0, y_1, \ldots, y_{p-1})$. Hence $(y_0, y_1, \ldots, y_{i-1}) = (y_0, y_1, \ldots, y_{p-1})^k$ and $(y_i, y_{i-1}, \ldots, y_{q-1}) = (y_0, y_1, \ldots, y_{p-1})^{k'}$ for some k, k' with $i = kp$ and $q - i = k'p$. Thus $q = (k + k')p$ and $(y_0, y_1, \ldots, y_{q-1}) = (y_0, y_1, \ldots, y_{p-1})^{k+k'}$ so that $(y_0, y_1, \ldots, y_{p-1})$ is also the prime factor of $(y_0, y_1, \ldots, y_{q-1})$. □

Corollary 6.6. *Suppose that $f \in \mathcal{B}^{(q)}$ and x is a periodic orbit of $(y_0, y_1, \ldots, y_{q-1}) \in N^q$ for f. Then the minimal period of x is the least i with $0 < i \le q$ such that $x_i = x_0$.*

Proof. Let i be the least i with $0 < i \le q$ such that $x_i = x_0$. By Lemma 6.5, $(y_0, y_1, \ldots, y_{i-1}) = (y_0, y_1, \ldots, y_{p-1})^k$ for some k. Thus $(f_{y_{p-1}\ldots y_0})^k(x) = x$, and since i is the least integer such that $x_i = x_0$, k is the least integer such that $(f_{y_{p-1}\ldots y_0})^k(x) = x$. Hence $kp = i$ is the minimal period of x. □

Corollary 6.7. *Suppose that $f \in \mathcal{B}^{(q)}$ and x is a periodic orbit of $(y_0, y_1, \ldots, y_{q-1}) \in N^q$ for f with minimal period i. Then $x_j = x_{j'}$ if and only if $j = j' \bmod i$ for any $0 \le j, j' < q$. In particular the points $\{x_0, x_1, \ldots, x_{i-1}\}$ are all distinct, i.e. $x_j \ne x_{j'}$ for all $0 \le j < j' < q$.*

Proof. Suppose that $x_j = x_{j'}$ for some $0 \le j < j' < q$. By Lemma 6.4, $x_{j'} = x_{(j' \bmod i)}$ and hence we may assume without loss of generality that $j < j' \le j + i$. Recall from Lemma 6.3 that x_j is a periodic orbit of the sequence $(y_j, y_{j+1}, \ldots, y_{q-1}, y_0, \ldots, y_{j-1}) \in N^q$ for f. Since $x_{j'} = x_j$, Corollary 6.6 implies that the minimal period i' of x_j is less than or equal to $j' - j$, and hence $i' \le j' - j \le i$. But, by Lemma 6.4, $x_{i'} = x_0$ and hence Corollary 6.6 implies that the minimal period i of x_0 for the sequence $(y_0, y_1, \ldots, y_{q-1})$ is less than or equal to i', in other words $i \le i'$. Thus $i' = j' - j = i$, and $j = j' \bmod i$ as claimed. □

This corollary shows that for f in $\mathcal{B}^{(q)}$, periodic orbits behave as for standard dynamical systems. Analogously to Lemma 4.8 of [Stark, 1999], we now aim to show that $\mathcal{B}^{(q)}$ is open and dense in $\mathcal{D}^{2r}(M \times N, M)$. Since the proof here is somewhat longer, we break it up into a number of steps, dealing with each $\mathcal{B}_k^{(q)}$ in turn. The overall proof is by induction, so that we will show that if for some $q \ge 0$, $\mathcal{B}^{(q)}$ is open and dense in $\mathcal{D}^{2r}(M \times N, M)$ then each of the $\mathcal{B}_k^{(q+1)}$ is open and dense in $\mathcal{B}^{(q)}$.

For any given $y = (y_0, y_1, \ldots, y_{q-1}) \in N^q$, define $\bar{\rho}_y \colon \mathcal{D}^{2r}(M \times N^{d-1}, M) \to \mathcal{D}^r(M, M)$ by

$$\bar{\rho}_y(f) = f_{y_{q-1}\ldots y_0}.$$

Thus $ev_{\bar{\rho}_y}(f,x) = ev_{\hat{\rho}_q}(f,x,y)$ where $\hat{\rho}_q$ is as in section 4.2. Thus by Corollary 4.2, $ev_{\bar{\rho}_y}$ is $\mathcal{C}^r$ and $T_{(f,x)}ev_{\bar{\rho}_y}(\eta, 0_x) = \bar{\eta}_q(x,y)$, with $\bar{\eta}_q$ satisfying $\bar{\eta}_j(x,y) = \eta(x_{j-1}, y_{j-1}) + T_{(x_{j-1},y_{j-1})}f(\bar{\eta}_{j-1}(x,y),0)$. Since $T_{(x_{i-1},y_{i-1})}f(v,0) = T_{x_{i-1}}f_{y_{i-1}}(v)$, we have

$$\bar{\eta}_q(x,y) = \sum_{j=1}^{q} T_{x_j} f_{y_{q-1}\dots y_j}(\eta(x_{j-1}, y_{j-1})), \tag{6.1}$$

with the convention that $T_{x_j} f_{y_{q-1}\dots y_j} = Id$ if $q = j$. Now, suppose that x is periodic for $y = (y_0, y_1, \dots, y_{q-1})$. Let the minimal period of x be i, so that by Lemma 6.4 $x_j = x_{(j \bmod i)}$. Then we can rearrange the sum in (6.1) to give

$$\bar{\eta}_q(x,y) = \sum_{s=0}^{r-1} (T_{x_0} f_{y_{i-1}\dots y_0})^s \sum_{j=1}^{i} T_{x_j} f_{y_{i-1}\dots y_j}(\eta(x_{j-1}, y_{j-1})), \tag{6.2}$$

where $q = ri$. Let $T^{(r)}$ be the linear operator

$$T^{(r)} = \sum_{s=0}^{r-1} (T_{x_0} f_{y_{i-1}\dots y_0})^s \tag{6.3}$$

with the convention that $T^{(0)} = Id$. A simple calculation yields

Lemma 6.8. *Suppose that $f \in \mathcal{B}^{(q')}$ for some $0 < q \le q'$, x is periodic for $(y_0, y_1, \dots, y_{q-1}) \in N^q$, with minimal period i. Then $T^{(r)}$ is invertible.*

Proof. By considering the Jordan Normal Form of $T_{x_0} f_{y_{i-1}\dots y_0}$, we see that the eigenvalues of $T^{(r)}$ are of the form $1 + \lambda + \cdots + \lambda^{r-1}$, where λ is an eigenvalue of $T_{x_0} f_{y_{i-1}\dots y_0}$. Since $x_i = x_0$, the segment $\{x_0, x_1, \dots, x_{i-1}\}$ is a periodic orbit for the sequence $(y_0, y_1, \dots, y_{i-1})$. By definition, $\mathcal{B}^{(q')} \subset \mathcal{B}^{(q)}$, and thus $T_{x_0} f_{y_{i-1}\dots y_0}$ has no eigenvalues on the unit circle and hence in particular $\lambda^r \ne 1$. But $(1 + \lambda + \cdots + \lambda^{r-1})(1 - \lambda) = 1 - \lambda^r$, and so the eigenvalues of $T^{(r)}$ are all nonzero. □

Essentially the same proof as that of Lemma 4.8 of [Stark, 1999] gives

Lemma 6.9. *Suppose that $f \in \mathcal{B}^{(q')}$ for some $0 < q \le q'$, x is periodic for $(y_0, y_1, \dots, y_{q-1}) \in N^q$, with minimal period i. Then given any $u \in T_x M$, there exists a $\eta \in T_f \mathcal{D}^{2r}(M \times N, M)$ such that $T_{(f,x)}ev_{\bar{\rho}_y}(\eta, 0_x) = u$. Thus in particular $T_{(f,x)}ev_{\bar{\rho}_y}$ is surjective.*

Proof. By Corollary 6.7 the points $\{x_0, x_1, \dots, x_{i-1}\}$ are all distinct. As in the proof of Lemma 5.3, we can use Corollary C.12 of [Stark, 1999] to construct a $\eta \in T_f \mathcal{D}^{2r}(M \times N, M)$ such that $\eta(x_j, y_j) = 0_{x_{i+1}}$ for $j = 0, \dots, i-2$, and $\eta(x_{i-1}, y_{i-1})$ takes on whatever value we want. By Lemma 6.8, $T^{(r)}$ is invertible, and hence given $u \in T_x M$, we can choose $\eta(x_{i-1}, y_{i-1}) = (T^{(r)})^{-1}(u)$. Substituting this into (6.2) immediately gives $T_{(f,x)}ev_{\bar{\rho}_y}(\eta, 0_x) = u$, as required. □

Corollary 6.10. *If for some $q > 0$, $\mathcal{B}^{(q)}$ is open and dense in $\mathcal{D}^{2r}(M \times N, M)$, then $\mathcal{B}_1^{(q+1)}$ is open and dense in $\mathcal{B}^{(q)}$, and hence open and dense in $\mathcal{D}^{2r}(M \times N, M)$.*

Proof. If $\mathcal{B}^{(q)}$ is open and dense in $\mathcal{D}^{2r}(M \times N, M)$, then it is a Banach manifold, and $T_f\mathcal{B}^{(q)} = T_f\mathcal{D}^{2r}(M \times N, M)$ for all $f \in \mathcal{B}^{(q)}$. Let $\Delta = \{(x, x') \in M \times M : x = x'\}$ be the diagonal in $M \times M$. For any given $y = (y_0, y_1, \ldots, y_{q-1}) \in N^q$, define $\bar{\rho}$: $\mathcal{B}^{(q)} \to \mathcal{D}^r(M, M \times M)$ by

$$\bar{\rho} = (C_{Id}, \bar{\rho}_y),$$

where C_{Id}: $\mathcal{B}^{(q)} \to \mathcal{D}^r(M, M)$ is the constant map $C_{Id}(f) = Id$. Thus, $ev_{\bar{\rho}}$ is $\mathcal{C}^r$ and $T_{(f,x)}ev_{\bar{\rho}_i}(\eta, 0_x) = (0_x, ev_{\bar{\rho}_y}(\eta, 0_x))$. If $ev_{\bar{\rho}}(f, x) \in \Delta$, then $x_q = x_0$ so that x_0 is a periodic orbit for the sequence $(y_0, y_1, \ldots, y_{q-1})$. Thus, by Lemma 6.9, Image $T_{(f,x)}ev_{\bar{\rho}} = \{0_x\} \times T_xM$. On the other hand $T_{(x,x)}\Delta = \{(u, u') \in T_{(x,x)}(M \times M) : u' = u\}$ and hence Image $T_{(f,x)}ev_{\bar{\rho}} + T_{(x,x)}\Delta = T_{(x,x)}(M \times M)$.

Thus, $ev_{\bar{\rho}}$ is transversal to Δ. The codimension of Δ is m and the dimension of M is also m, so that $\dim M - \operatorname{codim} \Delta = 0 < r$. Hence, by the Parametric Transversality Theorem, $\bar{\rho}(f)$ is transversal to Δ for an open dense set of $f \in \mathcal{B}^{(q)}$. For any such f, the set $(\bar{\rho}(f))^{-1}(\Delta)$ is a codimension m submanifold of M. Thus $(\bar{\rho}(f))^{-1}(\Delta)$ has dimension 0 for an open dense set of $f \in \mathcal{B}^{(q)}$, and since M is compact, it consists of a finite number of isolated points. But $\bar{\rho}(f) = (Id, f_{y_{q-1}\ldots y_0})$ and hence $(\bar{\rho}(f))^{-1}(\Delta) = \{x \in M : f_{y_{q-1}\ldots y_0}(x) = x\} = \mathcal{P}_{y_0\ldots y_{q-1}}$. Taking the intersection over all $(y_0, y_1, \ldots, y_{q-1}) \in N^q$, we obtain an open dense set of $f \in \mathcal{B}^{(q)}$ for which each $\mathcal{P}_{y_0\ldots y_{q-1}}$ consists of a finite number of points, as required. $\square$

A somewhat more complicated version of the same line of reasoning gives

Lemma 6.11. *If for some $q > 0$, $\mathcal{B}^{(q)}$ is open and dense in $\mathcal{D}^{2r}(M \times N, M)$, then $\mathcal{B}_2^{(q+1)}$ is open and dense in $\mathcal{B}^{(q)}$, and hence open and dense in $\mathcal{D}^{2r}(M \times N, M)$.*

Proof. Fix some $y = (y_0, y_1, \ldots, y_{i-1}) \in N^i$ and $y' = (y'_0, y'_1, \ldots, y'_{i'-1}) \in N^{i'}$ such that $y \neq y'$ with $0 < i, i' \leq q$. Define $\bar{\rho}_{y,y'}$: $\mathcal{B}^{(q)} \to \mathcal{D}^r(M, M \times M \times M)$ by

$$\bar{\rho}_{y,y'}(f) = (Id, f_{y_{i-1}\ldots y_0}, f_{y'_{i'-1}\ldots y'_0}).$$

We aim to show that $ev_{\bar{\rho}_{y,y'}}$ is transversal to the bidiagonal $D' = \{(x, x', x'') \in M \times M \times M : x = x' = x''\}$. Then using Corollary 4.2 as in Corollary 6.10, we have $T_{(f,x)}ev_{\bar{\rho}_{y,y'}}(\eta, 0_x) = (0_x, \bar{\eta}_i(x, y), \bar{\eta}'_i(x, y'))$ where $\bar{\eta}_i(x, y)$ satisfies (6.1) and $\bar{\eta}'_i(x, y')$ the analogue of (6.1) for y'. Suppose that $ev_{\bar{\rho}_{y,y'}}(f, x) \in \Delta'$ so that $x = x_i = x'_{i'}$ and hence x is a periodic orbit for both y and y'. Let the minimal period for y be p and that for y' be p', with $i = rp$ and $i' = r'p'$. Rearranging the sum in (6.1), as in (6.2), gives

$$\bar{\eta}_i(x, y) = \sum_{s=0}^{r-1}(T_{x_0}f_{y_{p-1}\ldots y_0})^s \sum_{j=1}^{p} T_{x_j}f_{y_{p-1}\ldots y_j}(\eta(x_{j-1}, y_{j-1})), \tag{6.4}$$

$$\bar{\eta}'_i(x, y') = \sum_{s=0}^{r'-1}(T_{x_0}f_{y'_{p'-1}\ldots y'_0})^s \sum_{j=1}^{p'} T_{x'_{j'}}f_{y'_{p'-1}\ldots y'_{j'}}(\eta(x'_{j-1}, y'_{j-1})), \tag{6.5}$$

where $x'_j = f_{y'_{j-1}\cdots y'_0}(x)$. Now suppose that $(y_0, y_1, \dots, y_{i-1})$ and $(y'_0, y'_1, \dots, y'_{i'-1})$ do not have the same prime factor, and suppose without loss of generality that $p \le p'$. We claim that $(y_0, y_1, \dots, y_{p-1}) \neq (y'_0, y'_1, \dots, y'_{p-1})$. Suppose not. Then $x'_p = f_{y'_{p-1}\cdots y'_0}(x) = f_{y_{p-1}\dots y_0}(x) = x_p$. But p is the minimal period of x with respect to $(y_0, y_1, \dots, y_{i-1})$ and hence $x'_p = x_p = x$. Hence by Corollary 6.6, the minimal period p' of x with respect to $(y'_0, y'_1, \dots, y'_{i'-1})$ is less than or equal to p. Hence $p = p'$ and $(y_0, y_1, \dots, y_{p-1}) = (y'_0, y'_1, \dots, y'_{p'-1})$. Thus, by Lemma 6.4 $(y_0, y_1, \dots, y_{i-1}) = (y_0, y_1, \dots, y_{p-1})^r$ and $(y'_0, y'_1, \dots, y'_{i'-1}) = (y'_0, y'_1, \dots, y'_{p'-1})^{r'} = (y_0, y_1, \dots, y_{p-1})^{r'}$ and hence $(y_0, y_1, \dots, y_{i-1})$ and $(y'_0, y'_1, \dots, y'_{i'-1})$ have the same prime factor, which contradicts our assumption that they do not. Thus $(y_0, y_1, \dots, y_{p-1}) \neq (y'_0, y'_1, \dots, y'_{p-1})$.

Let k be the least integer with $0 \le k < p$ such that $y_k \neq y'_k$. Recall that $x'_0 = x = x_0$, whilst if $k > 0$, then $(y_0, y_1, \dots, y_{k-1}) = (y'_0, y'_1, \dots, y'_{k-1})$ and hence $x'_j = f_{y'_{j-1}\cdots y'_0}(x) = f_{y_{j-1}\dots y_0}(x) = x_j$ for all $0 \le j \le k$. Hence we have $(x'_j, y'_j) = (x_j, y_j)$ for all $0 \le j \le k$, but $(x'_k, y'_k) \neq (x_k, y_k)$. Furthermore, by Corollary 6.7 the points $\{x_0, x_1, \dots, x_{p-1}\}$ are all distinct as are the points $\{x'_0, x'_1, \dots, x'_{p-1}\}$. Since $x'_k = x_k$, this implies that $x'_k \neq x_j$ for all $j \neq k$ with $0 \le j < p$ and $x_k \neq x'_j$ for all $j \neq k$ with $0 \le j < p'$. Hence as in Lemmas 5.3 and 6.9 we can use Corollary C.12 of [Stark, 1999] to construct a $\eta \in T_f \mathcal{D}^{2r}(M \times N, M)$ such that $\eta(x_j, y_j) = 0_{x_{i+1}}$ for $j \neq k$ with $0 \le j < p$, $\eta(x'_{j'}, y'_{j'}) = 0_{x'_{i'+1}}$ for $j \neq k$ with $0 \le j < p'$ and $\eta(x_k, y_k)$, $\eta(x'_k, y'_k)$ take on whatever values we want. By Lemma 6.8, $T^{(r)}$ and

$$T'^{(r)} = \sum_{s=0}^{r'-1} (T_{x_0} f_{y'_{p'-1}\cdots y'_0})^s$$

are invertible, whilst $T_{x_{k+1}} f_{y_{p-1}\dots y_{k+1}}$ and $T_{x'_{k+1}} f_{y'_{p'-1}\cdots y'_{k+1}}$ are both invertible by the definition of $\mathcal{D}^{2r}(M \times N, M)$. Hence given $u, u' \in T_x M$, choose $\eta(x_k, y_k) = (T^{(r)} \circ T_{x_{k+1}} f_{y_{p-1}\dots y_{k+1}})^{-1}(u)$ and $\eta(x'_k, y'_k) = (T'^{(r)} \circ T_{x'_{k+1}} f_{y'_{p'-1}\cdots y'_{k+1}})^{-1}(u')$. Evaluating using (6.4) and (6.5) gives $T_{(f,x)} ev_{\bar{\rho}_{y,y'}}(\eta, 0_x) = (0_x, u, u')$, so that Image $T_{(f,x)} ev_{\bar{\rho}_{y,y'}} = \{0_x\} \times T_x M \times T_x M$. Since $T_{(x,x,x)}\Delta' = \{(u, u', u'') \in T_{(x,x,x)}(M \times M \times M) : u = u' = u''\}$, we have Image $T_{(f,x)} ev_{\bar{\rho}_{y,y'}} + T_{(x,x,x)}\Delta' = T_{(x,x,x)}(M \times M \times M)$ and hence $ev_{\bar{\rho}_{y,y'}}$ is transversal to Δ'.

We now apply the same dimension-counting argument as in Propositions 5.4 and 5.5. The codimension of Δ' is $2m$ and the dimension of M is m, so that $\dim M - \operatorname{codim} \Delta' < 0 < r$. Thus by the Parametric Transversality Theorem, $\bar{\rho}_{y,y'}(f) = (Id, f_{y_{i-1}\dots y_0}, f_{y'_{i'-1}\cdots y_0})$ is transversal to Δ' for an open dense set of $f \in \mathcal{B}^{(q)}$. For any such f, suppose that $(Id, f_{y_{i-1}\dots y_0}, f_{y'_{i'-1}\cdots y_0})(x) \in \Delta'$. The dimension of $T_x M$ is m and hence the dimension of $T_x(Id, f_{y_{i-1}\dots y_0}, f_{y'_{i'-1}\cdots y_0})(T_x M)$ is at most m. The dimension of $T_{(x,x,x)}\Delta'$ is also m and hence the dimension of $T_x(Id, f_{y_{i-1}\dots y_0}, f_{y'_{i'-1}\cdots y_0})(T_x M) + T_{(x,x,x)}\Delta'$ is at most $2m$. However, the dimension of $T_{(x,x,x)}(M \times M \times M)$ is $3m$ and hence $T_x(Id, f_{y_{i-1}\dots y_0}, f_{y'_{i'-1}\cdots y_0})(T_x M)$ and $T_{(x,x,x)}\Delta'$ cannot span $T_{(x,x,x)}(M \times M \times M)$. Thus the only way for $\bar{\rho}_{y,y'}(f)$ to be transversal to Δ' is for the image of $\bar{\rho}_{y,y'}(f)$ not to intersect Δ'; in other words $(Id, f_{y_{i-1}\dots y_0}, f_{y'_{i'-1}\cdots y_0})(x) \notin \Delta'$ for all $x \notin M$. We have therefore shown that if $(y_0, y_1, \dots, y_{i-1})$ and $(y'_0, y'_1, \dots, y'_{i'-1})$ do not have the same prime factor, then for an open set of $f \in \mathcal{B}^{(q)}$ we cannot have $x = x_i = x'_{i'}$. Taking the intersection of such open dense sets for all possible choices of y and y', we see that $\mathcal{B}_2^{(q+1)}$ is open and dense in $\mathcal{B}^{(q)}$, as required. □

It now remains to show that for a dense set of $f \in \mathcal{B}^{(q)}$, periodic orbits of period q are hyperbolic. Intuitively this is obvious since by the above there is a finite number of such orbits, and hence if necessary we can independently perturb the eigenvalues of each orbit to ensure that there are no eigenvalues on the unit circle. For the convenience of the reader, we give a formal proof, once again using the Parametric Transversality Theorem. We begin with a result analogous to Lemma C.9 of [Stark, 1999] (see also Lemma 4.4 above):

Lemma 6.12. *Let $\bar{\tau}_y$: $\mathcal{D}^{2r}(M \times N^{d-1}, M) \to \mathcal{VB}^{r-1}(TM, TM)$ be given by*

$$\bar{\tau}_y(f) = T_x f_{y_{q-1}\ldots y_0}$$

for some fixed $y = (y_0, y_1, \ldots, y_{q-1}) \in N^q$. Then $ev_{\bar{\tau}_y}$ is $\mathcal{C}^{r-1}$ and

$$T_{(f,v)} ev_{\bar{\tau}_y}(\eta, 0_v) = \bar{\omega}(T_x \bar{\eta}_{q,y}(v)),$$

where $\bar{\eta}_{q,y}$ is given by $\bar{\eta}_{q,y}(x) = \bar{\eta}_q(x, y)$, with $\bar{\eta}_q$ as in (4.6) and (6.1).

Proof. We have $\bar{\tau}_y = \sigma \circ \bar{\rho}_y$ where σ: $\mathcal{C}^r(M, M) \to \mathcal{VB}^{r-1}(TM, TM)$ is the tangent operator, as in Appendix B.3 of [Stark, 1999]. By Lemma B.11 of [Stark, 1999], σ is $\mathcal{C}^\infty$ and $T_f \sigma(\eta) = \bar{\omega} \circ T_\eta$. On the other hand $\bar{\rho}_y(f)(x) = \hat{\rho}_q(f)(x, y)$ where $\hat{\rho}_q$: $\mathcal{D}^{2r}(M \times N, M) \to \mathcal{D}^r(M \times N^q, M)$ is as defined in section 4.2, so that by Lemma 4.1, $\hat{\rho}_q$ is $\mathcal{C}^r$ and $T_f \hat{\rho}_q(\eta) = \bar{\eta}_q$. Thus $\bar{\rho}_y(f) = \hat{\rho}_q(f) \circ e_y$, where e_y: $M \to M \times N^q$ is just $e_y(x) = (x, y)$. Hence $\bar{\rho}_y(f) = i_y(\hat{\rho}_q(f))$, where i_y: $\mathcal{D}^r(M \times N^q, M) \to \mathcal{D}^r(M, M)$ is given by $i_y(g) = g \circ e_y$. Hence i_y is just composition on the right, and so by Proposition 4.2 of [Franks, 1979] it is $\mathcal{C}^\infty$, as in Lemma 4.4. Its tangent map is given by $T_g \iota_y(\eta)(x) = \eta(x, y)$, or in other words with a mild abuse of notation, $T_g \iota_y(\eta) = \eta \circ e_y$. We thus have $\bar{\tau}_y = \sigma \circ i_y \circ \hat{\rho}_q$ and hence $\bar{\tau}_y$ is $\mathcal{C}^r$, with $T_f \bar{\tau}_y(\eta) = \bar{\omega}(T_x \bar{\eta}_{q,y})$ where $\bar{\eta}_{q,y} = \bar{\eta}_q \circ e_y$, so that $\bar{\eta}_{q,y}(x) = \bar{\eta}_q(x, y)$.

Similarly to Lemma 4.4, the evaluation function $ev_{\bar{\tau}_y}$ is given by $ev_{\bar{\tau}_y} = ev \circ (\bar{\tau}_y \times Id_v)$ where ev: $\mathcal{VB}^{r-1}(TM, TM) \times TM \to TM$ is the evaluation $ev(\Psi, v) = \Psi(v)$. By Corollary B.3 of [Stark, 1999] this is $\mathcal{C}^{r-1}$ and $T_{(\Psi,v)} ev(\eta, w) = \eta(v) + T_v \Psi(w)$. Thus $ev_{\bar{\tau}_y}$ is $\mathcal{C}^{r-1}$ by the chain rule and $T_{(f,v)} ev_{\bar{\tau}_y}(\eta, 0_v) = T_f \bar{\tau}_y(\eta)(v) = \bar{\omega}(T_x \bar{\eta}_{q,y}(v))$, as required. □

Denote by $T_{\mathbb{C}}M$ the complexification of TM; that is, $T_{\mathbb{C}}M = \{u + iv : u, v \in TM$ such that $\tau_M(u) = \tau_M(v)\}$, where τ_M: $TM \to M$ is the tangent bundle projection, as in section 5.3, so that the condition $\tau_M(u) = \tau_M(v)$ simply requires u and v to be in the same fibre. As usual, let $\tilde{T}_{\mathbb{C}}M = \{u + iv : u, v \in TM$ such that $\tau_M(u) = \tau_M(v)$ and $\|u\|^2 + \|v\|^2 = 1\}$ be the unit bundle in $T_{\mathbb{C}}M$. Let $\mathbb{T}$ be the unit circle in $\mathbb{C}$ and define the operator $\bar{\tau}$: $\mathcal{D}^{2r}(M \times N, M) \to \mathcal{C}^{r-1}(\tilde{T}_{\mathbb{C}}M \times \mathbb{T}, T_{\mathbb{C}}M \times T_{\mathbb{C}}M)$ by

$$\bar{\tau}(f) = (\Lambda, \bar{\tau}_y(f)),$$

where $\Lambda\colon \tilde{T}_{\mathbb{C}}M \times \mathbb{T} \to T_{\mathbb{C}}M$ is the fixed map $\Lambda(w, \lambda) = \lambda w$. Thus $ev_{\bar{\tau}}(f, w, \lambda) = \bar{\tau}(f)(w, \lambda) = (\lambda w, T_x f_{y_{q-1}\dots y_0}(w))$. Let $\Delta_{\mathbb{C}} = \{(w, w') \in T_{\mathbb{C}}M \times T_{\mathbb{C}}M : w = w'\}$ be the diagonal in $T_{\mathbb{C}}M \times T_{\mathbb{C}}M$, and note that if $(w, w') \in \Delta_{\mathbb{C}}$, then w and w' must be in the same fibre of $T_{\mathbb{C}}M$. Thus if $\bar{\tau}(f)(w, \lambda) \in \Delta_{\mathbb{C}}$, then $x = f_{y_{q-1}\dots y_0}(x)$ and $T_x f_{y_{q-1}\dots y_0}(w) = \lambda w$, so that x is a periodic orbit with an eigenvalue of unit modulus, and vice versa. Hence if the image of $\bar{\tau}(f)(w, \lambda)$ does not intersect $\Delta_{\mathbb{C}}$, then any periodic orbits of the sequence $(y_0, y_1, \dots, y_{q-1})$ must be hyperbolic. We thus use the usual parametric transversality argument to show that the evaluation map of $ev_{\bar{\tau}}$ is transversal to $\Delta_{\mathbb{C}}$, and then count dimensions to show that transversality implies nonintersection.

Lemma 6.13. *If for some $q > 0$, $\mathcal{B}^{(q)}$ is open and dense in $\mathcal{D}^{2r}(M \times N, M)$, then $ev_{\bar{\tau}}$ is transversal to $\Delta_{\mathbb{C}}$.*

Proof. Suppose $ev_{\bar{\tau}}(f, w, \lambda) \in \Delta_{\mathbb{C}}$. Then x is a periodic orbit of $(y_0, y_1, \dots, y_{q-1})$ with λ and eigenvalue of $T_x f_{y_{q-1}\dots y_0}$. This implies that x must have minimal period q, since if not, it would be a periodic orbit for $(y_0, y_1, \dots, y_{i-1})$ with $0 \le i < q$. Since $f \in \mathcal{B}^{(q)}$, by definition $T_x f_{y_{i-1}\dots y_0}$ has no eigenvalues of unit modulus. But $T_x f_{y_{q-1}\dots y_0} = (T_x f_{y_{i-1}\dots y_0})^r$, where $q = ir$, and therefore the eigenvalues of $T_x f_{y_{q-1}\dots y_0}$ are of the form μ^r, where μ is an eigenvalue of $T_x f_{y_{i-1}\dots y_0}$. Since $\|\lambda\| = 1$, this contradicts our assumption that $ev_{\bar{\tau}}(f, w, \lambda) \in \Delta_{\mathbb{C}}$.

By Lemma 6.12 we have $T_{(f,w,\lambda)} ev_{\bar{\tau}}(\eta, 0_w, 0_\lambda) = (0_{\lambda w}, \bar{\omega}(T_x \bar{\eta}_{q,y}(w)))$. Differentiating (6.1) gives

$$\begin{aligned} T_x \bar{\eta}_{q,y} &= \sum_{j=1}^{q} T_x (T f_{y_{q-1}\dots y_j} \circ \eta_{y_{j-1}} \circ f_{y_{j-2}\dots y_0}) \\ &= \sum_{j=1}^{q} T_{\eta(y_{j-1}, x_{j-1})} (T f_{y_{q-1}\dots y_j}) \circ T_{x_{j-1}} \eta_{y_{j-1}} \circ T_x f_{y_{j-2}\dots y_0}. \end{aligned} \tag{6.6}$$

Given $w \in T_{\mathbb{C},x}M$, as usual denote $w_j = T_x f_{y_{j-2}\dots y_0}(w)$. By Corollary 6.7 the points $\{x_0, x_1, \dots, x_{q-1}\}$ are all distinct and thus by Corollary C.16 of [Stark, 1999], given any $u \in T_{w_q}(T_{\mathbb{C}}M)$ there exists a $\eta \in T_f \mathcal{D}^{2r}(M \times N, M)$ such that $\bar{\omega}(T_{x_{q-1}} \eta_{y_{q-1}}(w_{q-1})) = u$, and $T_{x_{j-1}} \eta_{y_{j-1}}(w_{j-1}) = 0$ for $0 < j < q$. Therefore, since $T(T f_{y_{q-1}\dots y_j})$ is linear, we have $T_{\eta(y_{j-1}, x_{j-1})}(T f_{y_{q-1}\dots y_j})(T_{x_{j-1}} \eta_{y_{j-1}}(w_{j-1})) = 0$ for $0 < j < q$, so that

$$\bar{\omega}(T_x \bar{\eta}_{q,y}(w)) = \bar{\omega}(T_{x_{q-1}} \eta_{y_{q-1}}(w_{q-1})) = u.$$

Thus Image $T_{(f,w,\lambda)} ev_{\bar{\tau}} = \{0_{\lambda w}\} \times T_w(T_{\mathbb{C}}M)$. On the other hand, $T_{(w,w)} \Delta_{\mathbb{C}} = \{(u, u') \in T_{(w,w)}(T_{\mathbb{C}}M \times T_{\mathbb{C}}M) : u = u'\}$ and thus Image $T_{(f,w,\lambda)} ev_{\bar{\tau}} + T_{(w,w)} \Delta_{\mathbb{C}} = T_{(w,w)}(T_{\mathbb{C}}M \times T_{\mathbb{C}}M)$. Hence $ev_{\bar{\tau}}$ is transversal to $\Delta_{\mathbb{C}}$, as required. □

Corollary 6.14. *If for some $q > 0$, $\mathcal{B}^{(q)}$ is open and dense in $\mathcal{D}^{2r}(M \times N, M)$, then $\mathcal{B}_3^{(q+1)}$ is open and dense in $\mathcal{B}^{(q)}$, and hence open and dense in $\mathcal{D}^{2r}(M \times N, M)$.*

Proof. It is sufficient to consider orbits with period q since by definition if $f \in \mathcal{B}^{(q)}$ then any periodic orbit for $(y_0, y_1, \dots, y_{i-1})$ with $0 < i < q$ is hyperbolic. As in

Lemma 6.10, if $\mathcal{B}^{(q)}$ is open and dense in $\mathcal{D}^{2r}(M \times N, M)$, then it is a Banach manifold, and $T_f\mathcal{B}^{(q)} = T_f\mathcal{D}^{2r}(M \times N, M)$ for all $f \in \mathcal{B}^{(q)}$. The (real) dimension of $T_{\mathbb{C}}M$ is $3m$ and hence the codimension of $\Delta_{\mathbb{C}}$ is $3m$, whilst the dimension of $\tilde{T}_{\mathbb{C}}M \times \mathbb{T}$ is also $3m$, so that $\dim \tilde{T}_{\mathbb{C}}M \times \mathbb{T} - \operatorname{codim} \Delta_{\mathbb{C}} = 0 < r - 1$. Since $ev_{\bar{\tau}}$ is $\mathcal{C}^{r-1}$, by the Parametric Transversality Theorem, $\bar{\tau}(f)$ is transversal to $\Delta_{\mathbb{C}}$ for an open dense set of $f \in \mathcal{B}^{(q)}$. For any such f, suppose that $\bar{\tau}(f)(w, \lambda) \in \Delta_{\mathbb{C}}$, for some $(w, \lambda) \in \tilde{T}_{\mathbb{C}}M \times \mathbb{T}$. The dimension of $T_{(w,\lambda)}(\tilde{T}_{\mathbb{C}}M \times \mathbb{T})$ is $3m$ and hence the dimension of Image $T_{(w,\lambda)}(\bar{\tau}(f)) = T_{(w,\lambda)}(\bar{\tau}(f))(T_{(w,\lambda)}(\tilde{T}_{\mathbb{C}}M \times \mathbb{T}))$ is at most $3m$. The dimension of $T_{(w,w)}\Delta_{\mathbb{C}}$ is also $3m$ and the dimension of $T_{(w,w)}(T_{\mathbb{C}}M \times T_{\mathbb{C}}M)$ is $6m$. However, note that since $T_x f_{y_{q-1}\ldots y_0}(w)$ is linear, for any $\alpha \in \mathbb{C}$ we have $\bar{\tau}(f)(\alpha w, \lambda) = (\alpha\lambda w, \alpha T_x f_{y_{q-1}\ldots y_0}(w)) = \alpha\bar{\tau}(f)(w, \lambda)$. Hence if $\bar{\tau}(f)(w, \lambda) \in \Delta_{\mathbb{C}}$, then $\bar{\tau}(f)(\alpha w, \lambda) \in \Delta_{\mathbb{C}}$ for all $\alpha \in \mathbb{T}$. Thus the intersection of Image $T_{(w,\lambda)}(\bar{\tau}(f))$ and $T_{(w,w)}\Delta_{\mathbb{C}}$ has (real) dimension at least 1, and so the dimension of Image $T_{(w,\lambda)}(\bar{\tau}(f)) + T_{(w,w)}\Delta_{\mathbb{C}}$ is at most $6m-1$. Thus if $\bar{\tau}(f)(w, \lambda) \in \Delta_{\mathbb{C}}$, then $\bar{\tau}(f)$ cannot be transversal to $\Delta_{\mathbb{C}}$ at (w, λ). Hence $\bar{\tau}(f)(w, \lambda) \notin \Delta_{\mathbb{C}}$ for all $(w, \lambda) \in \tilde{T}_{\mathbb{C}}M \times \mathbb{T}$, or in other words any periodic orbits of minimal period q must be hyperbolic. $\square$

Corollary 6.15. *$\mathcal{B}^{(q)}$ is open and dense in $\mathcal{D}^{2r}(M \times N, M)$ for all $q > 0$.*

Proof. By induction on q. The result holds for $q = 1$ trivially by definition. Suppose that it holds for some $q > 1$. Then by Corollaries 6.10, 6.14, and Lemma 6.11, each of $\mathcal{B}_1^{(q+1)}$, $\mathcal{B}_2^{(q+1)}$, and $\mathcal{B}_3^{(q+1)}$ are open and dense in $\mathcal{B}^{(q)}$, and hence in $\mathcal{D}^{2r}(M \times N, M)$. Thus $\mathcal{B}^{(q+1)} = B_1^{(q+1)} \cap B_2^{(q+1)} \cap B_3^{(q+1)}$ is open and dense in $\mathcal{D}^{2r}(M \times N, M)$, thereby completing the inductive step. $\square$

Next, we use Lemma 6.12, and a similar (but more convoluted) argument to Lemma 6.13 and Corollary 6.14 to show that generically the eigenvalues of any periodic orbit are distinct and of multiplicity one. We proceed in two stages, defining

$$\begin{aligned} B_4^{(q)} = \quad & \{f \in \mathcal{D}^{2r}(M \times N^{d-1}, M) : \text{ if } x \in \mathcal{P}_{y_0\ldots y_{i-1}} \text{ for some } (y_0, y_1, \ldots, y_{i-1}) \in N^i \\ & \text{with } 0 < i < q, \text{ then the Jordan Normal Form of } T_x f_{y_{i-1}\ldots y_0} \text{ is diagonal}\}, \\ B_5^{(q)} = \quad & \{f \in \mathcal{D}^{2r}(M \times N^{d-1}, M) : \text{ if } x \in \mathcal{P}_{y_0\ldots y_{i-1}} \text{ for some } (y_0, y_1, \ldots, y_{i-1}) \in N^i \\ & \text{with } 0 < i < q \text{ and } w, w' \text{ are eigenvectors of } T_x f_{y_{i-1}\ldots y_0} \text{ with the same} \\ & \text{eigenvalue, then } w' = \alpha w \text{ for some } \alpha \in \mathbb{C}\}. \end{aligned}$$

We begin by recalling that the Whitney (or direct) sum of $T_{\mathbb{C}}M$ with itself is $T_{\mathbb{C}}M \bigoplus T_{\mathbb{C}}M = \{(w, w') \in T_{\mathbb{C}}M \times T_{\mathbb{C}}M : \tau_M(w) = \tau_M(w')\}$, and similarly for unit vectors $\tilde{T}_{\mathbb{C}}M \bigoplus \tilde{T}_{\mathbb{C}}M = \{(w, w') \in \tilde{T}_{\mathbb{C}}M \times \tilde{T}_{\mathbb{C}}M : \tau_M(w) = \tau_M(w')\}$. Thus $T_{\mathbb{C}}M \bigoplus T_{\mathbb{C}}M$ has (real) dimension $5m$ and $\tilde{T}_{\mathbb{C}}M \bigoplus \tilde{T}_{\mathbb{C}}M$ has dimension $5m - 2$. Let $\Delta_{\mathbb{C}}^{(4)} = \{(w, w', w'', w''') \in (T_{\mathbb{C}}M \bigoplus T_{\mathbb{C}}M) \times (T_{\mathbb{C}}M \bigoplus T_{\mathbb{C}}M) : w = w'', w' = w'''\}$ and define $\bar{\tau}_{\mathrm{JNF}}$: $\mathcal{D}^{2r}(M \times N, M) \to \mathcal{C}^{r-1}(\tilde{T}_{\mathbb{C}}M \bigoplus \tilde{T}_{\mathbb{C}}M \times \mathbb{C}, (T_{\mathbb{C}}M \bigoplus T_{\mathbb{C}}M) \times (T_{\mathbb{C}}M \bigoplus T_{\mathbb{C}}M))$ by

$$\bar{\tau}_{\mathrm{JNF}}(f)(w, w', \lambda) = (\lambda w, w + \lambda w', T_x f_{y_{q-1}\ldots y_0}(w), T_x f_{y_{q-1}\ldots y_0}(w')).$$

Recall that $\tau_M \circ T f_{y_{q-1}\ldots y_0} = f_{y_{q-1}\ldots y_0} \circ \tau_M$ and hence if $\tau_M(w) = \tau_M(w') = x$, then $\tau_M(T_x f_{y_{q-1}\ldots y_0}(w)) = x_q = \tau_M(T_x f_{y_{q-1}\ldots y_0}(w'))$ and thus $\bar{\tau}_{\mathrm{JNF}}(f)(w, w', \lambda) \in$

$(T_{\mathbb{C}}M \bigoplus T_{\mathbb{C}}M) \times (T_{\mathbb{C}}M \bigoplus T_{\mathbb{C}}M)$ as claimed. If $\bar{\tau}_{\mathrm{JNF}}(f)(w, w', \lambda) \in \Delta_{\mathbb{C}}^{(4)}$, w is an eigenvector of $T_x f_{y_{q-1}\ldots y_0}$ with eigenvalue λ, whilst w' satisfies $T_x f_{y_{q-1}\ldots y_0}(w') - \lambda w' = w$, so that it is a generalized eigenvector for λ. As usual, we want to prove that this does not occur for a dense open set of $f \in \mathcal{B}^{(q)}$. We do this by showing that $ev_{\bar{\tau}_{\mathrm{JNF}}}$ is transversal to $\Delta_{C}^{(4)}$, so that by the Parametric Transversality Theorem $\bar{\tau}_{\mathrm{JNF}}(f)$ is transversal to $\Delta_{\mathbb{C}}^{(4)}$ for a dense open set of f. We then use the usual dimension counting argument to show that transversality can only occur if the image of $\bar{\tau}_{\mathrm{JNF}}(f)$ does not intersect $\Delta_{\mathbb{C}}^{(4)}$. Thus, by Lemma 6.12 $ev_{\bar{\tau}_{\mathrm{JNF}}}$ is $\mathcal{C}^{r-1}$ and

$$T_{(f,w,w',\lambda)} ev_{\bar{\tau}_{\mathrm{JNF}}}(\eta, 0_w, 0_{w'}, 0_\lambda) = (0_{\lambda w}, 0_{w+\lambda w'}, \bar{\omega}(T_x \bar{\eta}_{q,y}(w), \bar{\omega}(T_x \bar{\eta}_{q,y}(w')).$$

It turns out to be convenient to consider only periodic orbits of minimal period q, and hence to organize the proof inductively. We thus begin with

Lemma 6.16. *If for some $q \geq 0$, $\mathcal{B}^{(q)} \cap \mathcal{B}_4^{(q)}$ is open and dense in $\mathcal{B}^{(q)}$ and hence in $\mathcal{D}^{2r}(M \times N, M)$ then the evaluation map $ev_{\bar{\tau}_{\mathrm{JNF}}}\colon (\mathcal{B}^{(q)} \cap \mathcal{B}_4^{(q)}) \times \tilde{T}_{\mathbb{C}}M \bigoplus \tilde{T}_{\mathbb{C}}M \times \mathbb{C} \to (T_{\mathbb{C}}M \bigoplus T_{\mathbb{C}}M) \times (T_{\mathbb{C}}M \bigoplus T_{\mathbb{C}}M)$ is transversal to $\Delta_{\mathbb{C}}^{(4)}$.*

Proof. Similarly to Corollary 6.10, since $\mathcal{B}^{(q)} \cap \mathcal{B}_4^{(q)}$ is open and dense in $\mathcal{D}^{2r}(M \times N, M)$, it is a Banach manifold, and $T_f(\mathcal{B}^{(q)} \cap \mathcal{B}_4^{(q)}) = T_f\mathcal{D}^{2r}(M \times N, M)$ for all $f \in (\mathcal{B}^{(q)} \cap \mathcal{B}_4^{(q)})$. Suppose $ev_{\bar{\tau}_{\mathrm{JNF}}}(f, w, w', \lambda) \in \Delta_{\mathbb{C}}^{(4)}$, so that $x = \tau_M(w) = \tau_M(w')$ is periodic of period q. As in Lemma 6.13, suppose the minimal period of x is i, with $q = ri$, with $r > 1$. Since $f \in (\mathcal{B}^{(q)} \cap \mathcal{B}_4^{(q)})$, the Jordan Normal Form of $T_x f_{y_{i-1}\ldots y_0}$ is diagonal. Hence, the Jordan Normal Form of $T_x f_{y_{q-1}\ldots y_0} = (T_x f_{y_{i-1}\ldots y_0})^r$ is also diagonal, contradicting $ev_{\bar{\tau}_{\mathrm{JNF}}}(f, w, w', \lambda) \in \Delta_{\mathbb{C}}^{(4)}$. Hence the minimal period of x must be q, so that by Corollary 6.7 the points $\{x_0, x_1, \ldots, x_{q-1}\}$ are all distinct. By Corollary C.18 of [Stark, 1999], the mapping from η to $(T_{x_0}\eta_{y_0}, T_{x_1}\eta_{y_1}, \ldots, T_{x_{q-1}}\eta_{y_{i-1}})$ is therefore a submersion. Thus, we can choose a $\eta \in T_f(\mathcal{B}^{(q)} \cap \mathcal{B}_4^{(q)})$ such that $T_{x_{j-1}}\eta_{y_{j-1}} = 0$ for $0 < j < q$ and $T_{x_{q-1}}\eta_{y_{q-1}}$ is whatever linear map we require. Since $\lambda w = T_x f_{y_{q-1}\ldots y_0}(w)$ and $T_x f_{y_{q-1}\ldots y_0}(w') - \lambda w' = w$, we have $w \neq \alpha w'$ for all $\alpha \in \mathbb{T}$. This is because if $w' = \alpha w$ for some $\alpha \in \mathbb{T}$, then $w = T_x f_{y_{q-1}\ldots y_0}(\alpha w) - \lambda \alpha w = 0$ by the linearity of $T_x f_{y_{q-1}\ldots y_0}$, contradicting $\|w\| = 1$ for all $w \in \tilde{T}_{\mathbb{C}}M$. Since $f_{y_{j-1}\ldots y_0}$ is a diffeomorphism, we have $w_j \neq \alpha w_j'$ for all $\alpha \in \mathbb{T}$, for all $0 \leq j \leq q$, where $w_j = T_x f_{y_{j-1}\ldots y_0}(w)$ and $w_j' = T_x f_{y_{j-1}\ldots y_0}(w')$. Thus, given any $u \in T_{w_q}(T_{\mathbb{C}}M)$ and $u' \in T_{w_q'}(T_{\mathbb{C}}M)$ we can find a $\eta \in T_f(\mathcal{B}^{(q)} \cap \mathcal{B}_4^{(q)})$ such that $\bar{\omega}(T_{x_{q-1}}\eta_{y_{q-1}}(w_{q-1})) = u$, $\bar{\omega}(T_{x_{q-1}}\eta_{y_{q-1}}(w_{q-1}')) = u'$, and $T_{x_{j-1}}\eta_{y_{j-1}}(w_{j-1}) = T_{x_{j-1}}\eta_{y_{j-1}}(w_{j-1}') = 0$ for $0 < j < q$. Using (6.6), this gives

$$T_{(f,w,w',\lambda)} ev_{\bar{\tau}_{\mathrm{JNF}}}(\eta, 0_w, 0_{w'}, 0_\lambda) = (0_{\lambda w}, 0_{w+\lambda w'}, u, u').$$

But $T_{(\lambda w, w+\lambda w', \lambda w, w+\lambda w')}\Delta_{\mathbb{C}}^{(4)} = \{(u, u', u'', u''') \in T_{(\lambda w, w+\lambda w', \lambda w, w+\lambda w')}(T_{\mathbb{C}}M \bigoplus T_{\mathbb{C}}M) \times (T_{\mathbb{C}}M \bigoplus T_{\mathbb{C}}M) : u = u'', u' = u''\}$. Thus Image $T_{(f,w,w',\lambda)} ev_{\bar{\tau}_{\mathrm{JNF}}} + T_{(\lambda w, w+\lambda w', \lambda w, w+\lambda w')}\Delta_{\mathbb{C}}^{(4)} = T_{(\lambda w, w+\lambda w', \lambda w, w+\lambda w')}(T_{\mathbb{C}}M \bigoplus T_{\mathbb{C}}M) \times (T_{\mathbb{C}}M \bigoplus T_{\mathbb{C}}M)$. Hence $ev_{\bar{\tau}_{\mathrm{JNF}}}$ is transversal to $\Delta_{\mathbb{C}}^{(4)}$, as required. □

Corollary 6.17. *$\mathcal{B}^{(q)} \cap \mathcal{B}_4^{(q)}$ is open and dense in $\mathcal{D}^{2r}(M \times N, M)$.*

Proof. By induction on q. The lemma holds trivially for $q = 1$ by definition. Suppose that it holds for some $q > 1$. The codimension of $\Delta_{\mathbb{C}}^{(4)}$ is $5m$ and the dimension of $\tilde{T}_{\mathbb{C}}M \bigoplus \tilde{T}_{\mathbb{C}}M \times \mathbb{C}$ is $5m$, so that $\dim \tilde{T}_{\mathbb{C}}M \bigoplus \tilde{T}_{\mathbb{C}}M \times \mathbb{C} - \operatorname{codim} \Delta_{\mathbb{C}}^{(4)} = 0 < r - 1$, and $ev_{\bar{\tau}_{\mathrm{JNF}}}$ is $\mathcal{C}^{r-1}$. Thus by the Parametric Transversality Theorem, $\bar{\tau}_{\mathrm{JNF}}(f)$ is transversal to $\Delta_{\mathbb{C}}^{(4)}$ for an open dense set of $f \in (\mathcal{B}^{(q)} \cap \mathcal{B}_4^{(q)})$. For any such f, suppose that $\bar{\tau}_{\mathrm{JNF}}(f)(w, w', \lambda) \in \Delta_{\mathbb{C}}^{(4)}$ for some (w, w', λ). Then $\bar{\tau}_{\mathrm{JNF}}(f)(\alpha w, \alpha w', \lambda) = (\lambda\alpha w, \alpha w + \lambda\alpha w', \alpha T_x f_{y_{q-1}\dots y_0}(w), \alpha T_x f_{y_{q-1}\dots y_0}(w')) = (\lambda\alpha w, \alpha w + \lambda\alpha w', \lambda\alpha w, \alpha(w + \lambda w')) \in \Delta_{\mathbb{C}}^{(4)}$ for any $\alpha \in \mathbb{T}$. Hence the intersection of Image $T_{(w,w',\lambda)}(\bar{\tau}_{\mathrm{JNF}}(f))$ and $T_{(\lambda w, w+\lambda w', \lambda w, w+\lambda w')}\Delta_{\mathbb{C}}^{(4)}$ has dimension at least 1. Thus the dimension of Image $T_{(w,w',\lambda)}(\bar{\tau}_{\mathrm{JNF}}(f)) + T_{(\lambda w, w+\lambda w', \lambda w, w+\lambda w')}\Delta_{\mathbb{C}}^{(4)}$ is at most $10m - 1$ and so Image $T_{(w,w',\lambda)}$ $(\bar{\tau}_{\mathrm{JNF}}(f)) + T_{(\lambda w, w+\lambda w', \lambda w, w+\lambda w')}\Delta_{\mathbb{C}}^{(4)}$ cannot span $T_{(\lambda w, w+\lambda w', \lambda w, w+\lambda w')}((T_{\mathbb{C}}M \bigoplus T_{\mathbb{C}}M) \times (T_{\mathbb{C}}M \bigoplus T_{\mathbb{C}}M))$. Since $\bar{\tau}_{\mathrm{JNF}}(f)$ is transversal to $\Delta_{\mathbb{C}}^{(4)}$, this implies that the image of $\bar{\tau}_{\mathrm{JNF}}(f)$ cannot intersect $\Delta_{\mathbb{C}}^{(4)}$, and hence no eigenvalue of $T_x f_{y_{q-1}\dots y_0}$ has a generalized eigenvector. Thus $\mathcal{B}^{(q+1)} \cap \mathcal{B}_4^{(q+1)}$ is open and dense in $\mathcal{B}^{(q)} \cap \mathcal{B}_4^{(q)}$, and hence open and dense in $\mathcal{D}^{2r}(M \times N, M)$, thereby completing the inductive step. □

We now go on to show that $\mathcal{B}^{(q)} \cap \mathcal{B}_5^{(q)}$ is dense in $\mathcal{D}^{2r}(M \times N, M)$. This requires a small modification of $\bar{\tau}_{\mathrm{JNF}}$, namely $\bar{\tau}_2$: $\mathcal{D}^{2r}(M \times N, M) \to \mathcal{C}^{r-1}(\tilde{T}_{\mathbb{C}}M \bigoplus \tilde{T}_{\mathbb{C}}M \times \mathbb{C}, (T_{\mathbb{C}}M \bigoplus T_{\mathbb{C}}M) \times (T_{\mathbb{C}}M \bigoplus T_{\mathbb{C}}M))$, given by

$$\bar{\tau}_2(f)(w, w', \lambda) = (\lambda w, \lambda w', T_x f_{y_{q-1}\dots y_0}(w), T_x f_{y_{q-1}\dots y_0}(w')).$$

If $\bar{\tau}_2(f)(w, w', \lambda) \in \Delta_{\mathbb{C}}^{(4)}$, then $\lambda w = T_x f_{y_{q-1}\dots y_0}(w)$ so that $x_q = \tau_M(T_x f_{y_{q-1}\dots y_0}(w)) = \tau_M(\lambda w) = x$, and so x is periodic with period q. Furthermore w and w' are both (unit) eigenvectors of $T_x f_{y_{q-1}\dots y_0}$ with the same eigenvalue λ. We want to show that this cannot happen if w and w' are independent, that is $w \neq \alpha w'$ for all $\alpha \in \mathbb{T}$, and as usual we employ a transversality argument. Thus, by Lemma 6.12 $ev_{\bar{\tau}_2}$ is $\mathcal{C}^{r-1}$ and

$$T_{(f,w,w',\lambda)} ev_{\bar{\tau}_2}(\eta, 0_w, 0_{w'}, 0_\lambda) = (0_{\lambda w}, 0_{\lambda w'}, \bar{\omega}(T_x \bar{\eta}_{q,y}(w), \bar{\omega}(T_x \bar{\eta}_{q,y}(w')).$$

Unfortunately, however, for $ev_{\bar{\tau}_2}$ to be transversal to $\Delta_{\mathbb{C}}^{(4)}$ we need to restrict to $(\tilde{T}_{\mathbb{C}}M \bigoplus \tilde{T}_{\mathbb{C}}M)\backslash\tilde{\Delta}$, where $\tilde{\Delta} = \{(w, w') \in \tilde{T}_{\mathbb{C}}M \bigoplus \tilde{T}_{\mathbb{C}}M : w \neq \alpha w' \text{ for some } \alpha \in \mathbb{T}\}$ is the set of pairs of linearly dependent vectors in $\tilde{T}_{\mathbb{C}}M \bigoplus \tilde{T}_{\mathbb{C}}M$. Note that since w and w' both have unit norm, it is sufficient to consider $\alpha \in \mathbb{T}$, rather than $\alpha \in \mathbb{C}$. Then,

Lemma 6.18. *If for some $q \geq 0$, $\mathcal{B}^{(q)} \cap \mathcal{B}_5^{(q)}$ is open and dense in $\mathcal{B}^{(q)}$ and hence in $\mathcal{D}^{2r}(M \times N, M)$, then the evaluation map $ev_{\bar{\tau}_2}$: $(\mathcal{B}^{(q)} \cap \mathcal{B}_5^{(q)}) \times ((\tilde{T}_{\mathbb{C}}M \bigoplus \tilde{T}_{\mathbb{C}}M)\backslash\tilde{\Delta}) \times \mathbb{C} \to (T_{\mathbb{C}}M \bigoplus T_{\mathbb{C}}M) \times (T_{\mathbb{C}}M \bigoplus T_{\mathbb{C}}M)$ is transversal to $\Delta_{\mathbb{C}}^{(4)}$.*

Proof. This is identical to the proof of Lemma 6.16, except that $ev_{\bar{\tau}_2}(f, w, w', \lambda) \in \Delta_{\mathbb{C}}^{(4)}$ no longer implies that $w \neq \alpha w'$ for all $\alpha \in \mathbb{T}$. This condition now follows by definition by excluding points in $\tilde{\Delta}$. □

Corollary 6.19. *$\mathcal{B}^{(q)} \cap \mathcal{B}_4^{(q)} \cap \mathcal{B}_5^{(q)}$ is open and dense in $\mathcal{D}^{2r}(M \times N, M)$.*

Proof. By induction on q. The lemma holds trivially for $q = 1$ by definition. Suppose that it holds for some $q > 1$. We first show that $\mathcal{B}^{(q+1)} \cap \mathcal{B}_5^{(q+1)}$ is dense in $\mathcal{B}^{(q)}$, closely following the argument in Corollary 6.17. The dimension of $((\tilde{T}_{\mathbb{C}}M \bigoplus \tilde{T}_{\mathbb{C}}M)\backslash\tilde{\Delta}) \times \mathbb{C}$ is $5m$, whilst the codimension of $\Delta_{\mathbb{C}}^{(4)}$ is $5m$ so that $\dim((\tilde{T}_{\mathbb{C}}M \bigoplus \tilde{T}_{\mathbb{C}}M)\backslash\tilde{\Delta}) \times \mathbb{C} - \operatorname{codim}\Delta_{\mathbb{C}}^{(4)} = 0 < r - 1$, and $ev_{\bar{\tau}_2}$ is $\mathcal{C}^{r-1}$. Thus by the Parametric Transversality Theorem, $\bar{\tau}_2(f)$ is transversal to $\Delta_{\mathbb{C}}^{(4)}$ for a residual set of $f \in \mathcal{B}^{(q)}$. Note that since $((\tilde{T}_{\mathbb{C}}M \bigoplus \tilde{T}_{\mathbb{C}}M)\backslash\tilde{\Delta}) \times \mathbb{C}$ is not closed, we do not get transversality on an open set in this case. For any such f, suppose that $\bar{\tau}_2(f)(w, w', \lambda) \in \Delta_{\mathbb{C}}^{(4)}$ for some (w, w', λ). The dimension of $T_{(w,w',\lambda)}(((\tilde{T}_{\mathbb{C}}M \bigoplus \tilde{T}_{\mathbb{C}}M)\backslash\tilde{\Delta}) \times \mathbb{C})$ is $5m$ and hence the dimension of Image $T_{(w,w',\lambda)}(\bar{\tau}_2(f))$ is at most $5m$. The dimension of $T_{(\lambda w,\lambda w',\lambda w,\lambda w')}\Delta_{\mathbb{C}}^{(4)}$ is also $5m$ and the dimension of $T_{(\lambda w,\lambda w',\lambda w,\lambda w')}((T_{\mathbb{C}}M \bigoplus T_{\mathbb{C}}M) \times (T_{\mathbb{C}}M \bigoplus T_{\mathbb{C}}M))$ is $10m$. However, similarly to Corollaries 6.14 and 6.17, if $\bar{\tau}_2(f)(w, w', \lambda) \in \Delta_{\mathbb{C}}^{(4)}$, then $\bar{\tau}_2(f)(\alpha w, \alpha' w', \lambda) \in \Delta_{\mathbb{C}}^{(4)}$ for any $\alpha, \alpha' \in \mathbb{T}$. Hence the intersection of Image $T_{(w,w',\lambda)}(\bar{\tau}_2(f))$ and $T_{(\lambda w,\lambda w',\lambda w,\lambda w')}\Delta_{\mathbb{C}}^{(4)}$ has dimension at least 2, and so the dimension of Image $T_{(w,w',\lambda)}(\bar{\tau}_2(f)) + T_{(\lambda w,\lambda w',\lambda w,\lambda w')}\Delta_{\mathbb{C}}^{(4)}$ is at most $10m-2$. Thus $\bar{\tau}_2(f)$ cannot be transversal to $\Delta_{\mathbb{C}}^{(4)}$ at (w, w', λ) if $\bar{\tau}_2(f)(w, w', \lambda) \in \Delta_{\mathbb{C}}^{(4)}$, so that $\bar{\tau}_2(f)(w, w', \lambda) \notin \Delta_{\mathbb{C}}^{(4)}$ for all $(w, w', \lambda) \in ((\tilde{T}_{\mathbb{C}}M \bigoplus \tilde{T}_{\mathbb{C}}M)\backslash\tilde{\Delta}) \times \mathbb{C}$. This implies that if w and w' are both eigenvectors of $T_x f_{y_{i-1}\dots y_0}$ with the same eigenvalue λ, then $w = \alpha w'$ for some $\alpha \in \mathbb{C}$. Thus, we have shown that $\mathcal{B}^{(q+1)} \cap \mathcal{B}_5^{(q+1)}$ is dense in $\mathcal{B}^{(q)}$, and hence dense in $\mathcal{D}^{2r}(M \times N, M)$.

Now, suppose that $f \in (\mathcal{B}^{(q+1)} \cap \mathcal{B}_4^{(q+1)} \cap \mathcal{B}_5^{(q+1)})$. Since the periodic orbits of f of period q or less are all hyperbolic, the Implicit Function Theorem implies that there is an open neighbourhood of f in $\mathcal{D}^{2r}(M \times N, M)$ which has the same number of periodic orbits of each period, and furthermore the position of these orbits depends continuously on f. Thus, the Jacobian $T_x f_{y_{i-1}\dots y_0}$ of any such periodic orbits varies continuously with f, and hence so do its eigenvalues. There is thus an open neighbourhood of f in which the eigenvalues of independent eigenvectors are distinct. This neighbourhood is thus in $\mathcal{B}_5^{(q+1)}$, since $\mathcal{B}^{(q+1)} \cap \mathcal{B}_4^{(q+1)}$ is open in $\mathcal{D}^{2r}(M \times N, M)$ by Corollary 6.17, $\mathcal{B}^{(q+1)} \cap \mathcal{B}_4^{(q+1)} \cap \mathcal{B}_5^{(q+1)}$ is open in $\mathcal{D}^{2r}(M \times N, M)$. We have shown above that $\mathcal{B}^{(q+1)} \cap \mathcal{B}_5^{(q+1)}$ is dense in $\mathcal{D}^{2r}(M \times N, M)$, and again by Corollary 6.17 $\mathcal{B}^{(q+1)} \cap \mathcal{B}_4^{(q+1)}$ is dense in $\mathcal{D}^{2r}(M \times N, M)$. Thus $\mathcal{B}^{(q+1)} \cap \mathcal{B}_4^{(q+1)} \cap \mathcal{B}_5^{(q+1)}$ is open and dense in $\mathcal{D}^{2r}(M \times N, M)$, completing the inductive step. □

Finally, it turns out to be convenient to place one further restriction on the periodic orbits of f, namely that f does not map one periodic orbit to another in too small a number of iterates. This is a generalization of Lemma 6.11 and will be proved by a similar argument. Although not strictly necessary, this property does considerably simplify the proofs of Propositions 6.25, 6.26, and 6.30 below. To motivate the precise formulation we employ, note that hitherto, given a sequence $(y_0, y_1, \dots, y_{q-1}) \in N^q$, we have considered orbits periodic with respect to the whole sequence, so that $x_q = x$. However, when we come to the proof of Theorem 3.4, we are concerned specifically with orbits where $x_i = x_j$ for some $0 \le i \neq j \le q - 1$, without the preceding segment $(y_0, y_1, \dots, y_{i-1})$ and the following segment $(y_j, y_{j+1}, \dots, y_{q-1})$ necessarily bearing any relation to the periodic part $(y_i, y_{i+1}, \ldots, y_{j-1})$. We thus define

Definition 6.20. Given a sequence $(y_0, y_1, \ldots, y_{q-1}) \in N^q$ and some $i, j, k \in \mathbb{Z}$ with $0 \le i \ne j \le q-1$, define $k \bmod (i, j)$ to be the unique $k' \in \{i, i+1, \ldots, j-1\}$ such that $k' = k \bmod (j-i)$. Then let

$$\begin{aligned} b^-(i, j) &= \min\{k \in \{0, 1, \ldots, i\} : y_{k'} = y_{(k' \bmod (i,j))} \text{ for all } k \le k' \le i\}, \\ b^+(i, j) &= \max\{k \in \{j, j+1, \ldots, q\} : y_{k'} = y_{(k' \bmod (i,j))} \text{ for all } j-1 \le k' \le k-1\}. \end{aligned}$$

We shall call the segment $(y_{b^-(i,j)}, y_{b^-(i,j)+1}, \ldots, y_{b^+(i,j)-1})$ the *maximal periodic segment* of (i, j).

The asymmetry in the definition of $b^-(i, j)$ and $b^+(i, j)$ arises because y_j plays no role in determining whether or not $x_i = x_j$. Note that we have $y_k = y_{(k \bmod (i,j))}$ for all $k \in \{b^-(i, j), \ldots, b^+(i, j) - 1\}$, but $y_{b^-(i,j)-1} \ne y_{((b^-(i,j)-1) \bmod (i,j))}$ (if $b^-(i, j) > 0$) and $y_{b^+(i,j)} \ne y_{(b^+(i,j) \bmod (i,j))}$ (if $b^+(i, j) < q$). Using Lemma 6.3, it is straightforward to see that if x_i is periodic for $(y_i, y_{i+1}, \ldots, y_{j-1})$ under some $f \in \mathcal{D}^{2r}(M \times N^{d-1}, M)$, then the first of these relations also holds for x_k.

Lemma 6.21. *Given $f \in \mathcal{D}^{2r}(M \times N^{d-1}, M)$, $x \in M$, and $(y_0, y_1, \ldots, y_{q-1}) \in N^q$ such that $x_i \in \mathcal{P}_{y_i \ldots y_{j-1}}$ for some $0 \le i \ne j \le q-1$, then $x_k = x_{(k \bmod (i,j))}$ for all $k \in \{b^-(i, j), \ldots, b^+(i, j)\}$.*

Proof. Suppose $k' \in \{i, i+1, \ldots, j-1\}$. Then by Lemma 6.3, $x_{k'}$ is periodic for $(y_{k'}, y_{k'+1}, \ldots, y_{j-1}, y_i, y_{i+1}, \ldots, y_{k'-1})$, so that $f_{y_{k'-1} \ldots y_{j-1} y_i \ldots y_{k'}}(x_{k'}) = x_{k'}$. First consider a $k \in \{b^-(i, j), \ldots, b^+(i, j)\}$ such that $k' = k \bmod (i, j)$ and $k > k'$. Then $y_{k''} = y_{(k'' \bmod (i,j))}$ for all $k' \le k'' < k$ and hence $(y_{k'}, y_{k'+1}, \ldots, y_{k-1}) = (y_{k'}, y_{k'+1}, \ldots, y_{j-1}, y_i, y_{i+1}, \ldots, y_{k'-1})^r$ for some r. Thus $x_k = f_{y_{k-1} \ldots y_{k'}}(x_{k'}) = (f_{y_{k'-1} \ldots y_{j-1} y_i \ldots y_{k'}})^r(x_{k'}) = x_{k'}$. Similarly, if $k < k'$, then $(y_k, y_{k+1}, \ldots, y_{k'-1}) = (y_{k'}, y_{k'+1}, \ldots, y_{j-1}, y_i, y_{i+1}, \ldots, y_{k'-1})^r$ for some r. Then, since $f_{y_k \ldots y_{k'-1}}$ is invertible, $x_k = (f_{y_{k'-1} \ldots y_k})^{-1}(x_{k'}) = ((f_{y_{k'-1} \ldots y_{j-1} y_i \ldots y_{k'}})^r)^{-1}(x_{k'}) = ((f_{y_{k'-1} \ldots y_{j-1} y_i \ldots y_{k'}})^{-1})^r(x_{k'}) = x_{k'} = x_{(k \bmod (i,j))}$. Hence in both cases, $x_k = x_k = x_{(k \bmod (i,j))}$ as required. $\square$

If we now restrict to $f \in \mathcal{B}^{(q)}$, we also have $x_{b^-(i,j)-1} \ne x_{((b^-(i,j)-1) \bmod (i,j))}$ (if $b^-(i, j) > 0$) and $x_{b^+(i,j)+1} \ne x_{((b^+(i,j)+1) \bmod (i,j))}$ (if $b^+(i, j) < q$), in analogy to the relations for $y_{b^-(i,j)-1}$ and $y_{b^+(i,j)}$. However, we shall in fact need the stronger

Lemma 6.22. *Suppose $f \in \mathcal{B}^{(q)}$ and $x_i \in \mathcal{P}_{y_i \ldots y_{j-1}}$ for some $(y_0, y_1, \ldots, y_{q-1}) \in N^q$. Then $\{x_0, \ldots, x_{b^-(i,j)-1}\} \cap \{x_{b^-(i,j)}, \ldots, x_{b^+(i,j)}\} = \emptyset$, and similarly $\{x_{b^+(i,j)+1}, \ldots, x_q\} \cap \{x_{b^-(i,j)}, \ldots, x_{b^+(i,j)}\} = \emptyset$.*

Proof. We show that $x_k \notin \{x_{b^-(i,j)}, \ldots, x_{b^+(i,j)}\}$ for all $0 \le k < b^-(i, j)$ by contradiction; the proof for $b^+(i, j) < k \le q$ is similar. So suppose that $x_k = x_{k'}$ with $0 \le k < b^-(i, j) \le k' \le b^+(i, j)$. By Lemma 6.21, $x_{k'} = x_{k''}$ for some $i \le k'' < j$ and thus by Lemma 6.3, $x_k = x_{k'} = x_{k''}$ is periodic for $(y_{k''}, y_{k''+1}, \ldots, y_{j-1}, y_i, \ldots, y_{k''-1})$. On the other hand $x_k = x_{k''}$ implies that x_k is periodic for $(y_k, y_{k+1}, \ldots, y_{k''-1})$. Hence by the definition of $\mathcal{B}^{(q)}$, $(y_{k''}, y_{k''+1}, \ldots, y_{j-1}, y_i, \ldots, y_{k''-1})$ and $(y_k, y_{k+1}, \ldots, y_{k''-1})$ are multiples of the same prime factor, of length say p, so that $j - i = ps$ and $k'' - k = ps'$. This prime factor must agree with the last p entries of $(y_k, y_{k+1}, \ldots, y_{k''-1})$, and we may thus write it as $(y_{k''-p}, y_{k''-p+1}, \ldots, y_{k''-1})$. Thus $(y_{k''}, y_{k''+1}, \ldots, y_{j-1}, y_i, \ldots, y_{k''-1}) = (y_{k''-p}, y_{k''-p+1}, \ldots, y_{k''-1})^s$ and $(y_k, y_{k+1}, \ldots, y_{k''-1}) = (y_{k''-p}, y_{k''-p+1}, \ldots, y_{k''-1})^{s'}$. Working upwards from k'' in the first of these relations, we see that $y_{k''+r} = y_{k''-p+(r \bmod p)}$ for all $r \in \{k'', \ldots, j-1\}$, whilst working backwards yields $y_r = y_{(r \bmod (k''-p, k''-1))}$ for all $r \in \{i, \ldots, k''-1\}$. But $y_{k''-p+(r \bmod p)} = y_{((k''+r) \bmod (k''-p, k''-1))}$ and so $y_r = y_{(r \bmod (k''-p, k''-1))}$ for all $r \in \{i, \ldots, j-1\}$. Similarly, working downwards from $k''-1$ in the second relation gives $y_r = y_{(r \bmod (k''-p, k''-1))}$ for all $r \in \{k, \ldots, k''-1\}$. But given any $r \in \{k, \ldots, k''-1\}$, by definition $(r \bmod (i, j) - r)$ is divisible by $(j - i)$ and hence by p, whilst $(r \bmod (k''-p, k''-1) - r)$ is also divisible by p. Hence $(r \bmod (i, j) - r \bmod (k''-p, k''-1))$ is divisible by p, so that $r \bmod (k''-p, k''-1) = (r \bmod (i, j)) \bmod (k''-p, k''-1)$. Since $(r \bmod (i, j)) \in \{i, \ldots, j-1\}$, this implies by the above that $y_{(r \bmod (i,j))} = y_{(r \bmod (k''-p, k''-1))}$, and since $y_r = y_{(r \bmod (k''-p, k''-1))}$, we have $y_r = y_{(r \bmod (i,j))}$ for all $r \in \{k, \ldots, k''-1\}$. Thus $b^-(i, j) \leq k$, which contradicts our initial assumption that $k < b^-(i, j)$. $\square$

Finally, we can show that for a dense open set of $f \in \mathcal{B}^{(q)}$ we cannot have more than one periodic segment on an orbit of length q. Thus define

$$\mathcal{B}_6^{(q)} = \{f \in \mathcal{D}^{2r}(M \times N^{d-1}, M) : \text{if } x \in \mathcal{P}_{y_0 \ldots y_{j-1}} \text{ with } 0 < j < q \text{ for some } (y_0, y_1, \ldots, y_{q-1}) \in N^q \text{ and for some } j < i' < j' < q \text{ we have } b(0, j) < b(i', j'), \text{ then } f_{y_{i'-1} \ldots y_0}(x) \notin \mathcal{P}_{y_{i'} \ldots y_{j'-1}}\}.$$

Note that if $f \in \mathcal{B}_6^{(q)}$ and $x_i \in \mathcal{P}_{y_i \ldots y_{j-1}}$, then we cannot have $f_{y_{i'-1} \ldots y_i}(x_i) \in \mathcal{P}_{y_{i'} \ldots y_{j'-1}}$ for any i', j' such that $j < i' < j' < q$ and $b^-(i, j) < b^-(i', j')$, and hence it is sufficient to have $x \in \mathcal{P}_{y_0 \ldots y_{i-1}}$ in the definition of $\mathcal{B}_6^{(q)}$, rather than the apparently more general $x_i \in \mathcal{P}_{y_i \ldots y_{j-1}}$. We want to show that $\mathcal{B}^{(q)} \cap \mathcal{B}_6^{(q)}$ is open and dense in $\mathcal{D}^{2r}(M \times N, M)$. Define $\bar{\rho}_4$: $\mathcal{D}^{2r}(M \times N^{d-1}, M) \to \mathcal{D}^r(M, M \times M \times M \times M)$ by

$$\bar{\rho}_4(f) = (Id, f_{y_{j-1} \ldots y_0}, f_{y_{i'-1} \ldots y_0}, f_{y_{j'-1} \ldots y_0}),$$

similarly to $\bar{\rho}_{y,y'}$ in Lemma 6.11. Observe that if $\bar{\rho}_4(f)(x) \in \Delta_4$, where $\Delta_4 = \{(x, x', x', x''') \in M \times M \times M \times M : x = x', x'' = x'''\}$, then $x = x_j$ and $x_{i'} = x_{j'}$, so that $x \in \mathcal{P}_{y_0 \ldots y_{j-1}}$ and $x_{i'} \in \mathcal{P}_{y_{i'} \ldots y_{j'-1}}$. As usual, we therefore aim to prove that $ev_{\bar{\rho}_4}$ is transversal to Δ_4, and then use a dimension counting argument to show that $\bar{\rho}_4(f)$ can be transversal to Δ_4 only if the image of $\bar{\rho}_4(f)$ does not intersect Δ_4:

Lemma 6.23. *$\mathcal{B}^{(q)} \cap \mathcal{B}_6^{(q)}$ is open and dense in $\mathcal{B}^{(q)}$ and hence in $\mathcal{D}^{2r}(M \times N, M)$.*

Proof. First observe that $x_{i'} \in \mathcal{P}_{y_{i'} \ldots y_{j'-1}}$ if and only if $x_{b^-(i',j')} \in \mathcal{P}_{y_{b^-(i',j')} \cdots y_{b^-(i',j')+j'-i'-1}}$. Hence we may assume without loss of generality that $b^-(i', j') = i'$. As in Corollary 6.10, Corollary 4.2 implies that $ev_{\bar{\rho}_4}$ defined as above is $\mathcal{C}^r$ and $T_{(f,x)} ev_{\bar{\rho}_4}(\eta, 0_x) = (0_x, \bar{\eta}_j(x, y), \bar{\eta}_{i'}(x, y), \bar{\eta}_{j'}(x, y))$ where $\bar{\eta}_i(x, y)$ satisfies (6.1). We aim to show that

$ev_{\bar{\rho}_4}\colon \mathcal{B}^{(q)} \times M \to M \times M \times M \times M$ is transversal to Δ_4. Suppose that $ev_{\bar{\rho}_4}(f, x) \in \Delta_4$, so that $x \in \mathcal{P}_{y_0 \ldots y_{i-1}}$ and $x_{i'} \in \mathcal{P}_{y_{i'} \ldots y_{j'-1}}$. Let the minimal period of x for $(y_0, y_1, \ldots, y_{j-1})$ be p and of $x_{i'}$ for $(y_{i'}, y_{i'+1}, \ldots, y_{j'-1})$ be p'. Using (6.2), we have

$$\bar{\eta}_j(x, y) = \sum_{i=1}^{p} T^{(s)}(T_{x_i} f_{y_{p-1} \ldots y_i}(\eta(x_{i-1}, y_{i-1}))),$$

with $T^{(s)}$ invertible by Lemma 6.8. As usual, by Corollary 6.7 the points $\{x_0, x_1, \ldots, x_{p-1}\}$ are all distinct and so given a $u \in T_{x_j} M$, we can construct a $\eta \in T_f \mathcal{B}^{(q)}$ using Corollary C.12 of [Stark, 1999] such that $\eta(x_i, y_i) = 0_{x_{i+1}}$ for $i = 0, \ldots, p-2$, and $\eta(x_{p-1}, y_{p-1}) = (T^{(s)})^{-1}(u)$, so that $\bar{\eta}_i(x, y) = u$. The points $\{(x_p, y_p), \ldots, (x_{b^+(0,j)-1}, y_{b^+(0,j)-1})\}$ are all determined by $x_i = x_{(i \bmod p)}$ and $y_i = y_{(i \bmod p)}$. Hence this choice of $\eta \in T_f \mathcal{B}^{(q)}$ fixes a value of $\bar{\eta}_{b^+(0,j)}(x, y) \in T_{x_{b^+(0,j)}} M$ (which we do not need to evaluate). By Corollary 6.7, $x_{b^+(0,j)} \neq x_i$ for any $i \neq b^+(0, j) \bmod p$, so that $(x_{b^+(0,j)}, y_{b^+(0,j)}) \neq (x_i, y_i)$ for all $i \neq b^+(0, j) \bmod p$. On the other hand, by the definition of $b^+(0, j)$ we have $y_{b^+(0,j)} \neq y_{(b^+(0,j) \bmod p)}$, and hence $(x_{b^+(0,j)}, y_{b^+(0,j)}) \neq (x_{(b^+(0,j) \bmod p)}, y_{(b^+(0,j) \bmod p)})$. Thus $(x_{b^+(0,j)}, y_{b^+(0,j)}) \notin \{(x_0, y_0), (x_1, y_1), \ldots, (x_{b^+(0,j)-1}, y_{b^+(0,j)-1})\}$, and hence we can choose the value of $\eta(x_{b^+(0,j)}, y_{b^+(0,j)})$ independently of $\eta(x_i, y_i)$ for $i = 0, \ldots, b^+(0, j) - 1$. Taking $\eta(x_{b^+(0,j)}, y_{b^+(0,j)}) = -T_{(x_{b^+(0,j)}, y_{b^+(0,j)})} f(\bar{\eta}_{b^+(0,j)}(x, y), 0)$ gives $\bar{\eta}_{b^+(0,j)+1}(x, y) = 0_{x_{b^+(0,j)+1}}$ by (4.5). Furthermore by Lemma 6.22, $\{x_{b^+(0,j)+1}, \ldots, x_q\} \cap \{x_0, x_1, \ldots, x_{b^+(0,j)}\} = \emptyset$, and hence we can choose $\eta(x_i, y_i) = 0_{x_{i+1}}$ for $i = b^+(0, j) + 1, \ldots, q-1$, which gives $\bar{\eta}_i(x, y) = 0_{x_i}$ for $i = b^+(0, j) + 1, \ldots, q$. Hence, this choice of $\eta \in T_f \mathcal{B}^{(q)}$ gives $T_{(f,x)} ev_{\bar{\rho}_4}(\eta, 0_x) = (0_x, u, 0_{x_{i'}}, 0_{x_{j'}})$.

Turning to the last two components of $T_{(f,x)} ev_{\bar{\rho}_4}$, by Lemma 6.22 and our assumption that $b^-(i', j') = i'$, we have $\{x_0, x_1, \ldots, x_{i'-1}\} \cap \{x_{i'}, \ldots, x_{j'-1}\} = \emptyset$. Hence we may choose a $\eta' \in T_f \mathcal{B}^{(q)}$ such that $\eta'(x_i, y_i) = 0_{x_{i+1}}$ for $i = 0, \ldots, i'-1$, and $\eta'(x_i, y_i)$ takes on whatever values we require for $i = j', \ldots, j' + p' - 1$. For such a choice, using (6.1), we have $\bar{\eta}_{i'}(x, y) = 0_{x_{i'}}$ and

$$\bar{\eta}_{j+i'}(x, y) = \sum_{i=i'+1}^{i'+p'} T^{(s')}(T_{x_i} f_{y_{i'+p'-1} \ldots y_i}(\eta(x_{i-1}, y_{i-1}))),$$

with $T^{(s')}$ invertible by Lemma 6.8. Thus, given $u' \in T_{x_{j'}} M$, we take $\eta' \in T_f \mathcal{B}^{(q)}$ so that $\eta(x_i, y_i) = 0_{x_{i+1}}$ for $i = i', \ldots, i' + p' - 2$, and $\eta(x_{i'+p'-1}, y_{i'+p'-1}) = (T^{(s')})^{-1}(u')$, so that $\bar{\eta}_{j'}(x, y) = u'$. This gives $T_{f,x} ev_{\bar{\rho}_4}(\eta', 0_x) = (0_x, 0_{x_j}, 0_{x_{i'}}, u')$. Hence we have $(0_x, u, 0_{x_{i'}}, u') \in T_{(f,x)} ev_{\bar{\rho}_4}(T_{(f,x)}(\mathcal{B}^{(q)} \times M))$ for all $u \in T_{x_j} M$ and $u' \in T_{x_{j'}} M$. But $(u, u, u', u') \in T_{(x,x,x',x')} \Delta_4$ and so $T_{(f,x)} ev_{\bar{\rho}_4}(T_{(f,x)}(\mathcal{B}^{(q)} \times M)) + T_{(x,x,x',x')} \Delta_4$ spans $T_{(x,x,x',x')}(M \times M \times M \times M)$. Thus $ev_{\bar{\rho}_4}$ is transversal to Δ_4. The dimension of M is m and the codimension of Δ_4 is $2m$ and hence $\dim M - \operatorname{codim} \Delta_4 < 0 < r$. Since $ev_{\bar{\rho}_4}$ is $\mathcal{C}^r$, the Parametric Transversality Theorem implies that $\bar{\rho}_4(f)$ is transversal to Δ_4 for an open dense set of $f \in \mathcal{B}^{(q)}$. Choose any such f and suppose $\bar{\rho}_4(f)(x) \in \Delta_4$. Then the dimension of $T_x \bar{\rho}_4(f)(T_x M)$ is at most m and the dimension of $T_{(x,x,x',x')} \Delta_4$ is $2m$. Hence Image $T_x \bar{\rho}_4(f) + T_{(x,x,x',x')} \Delta_4$ cannot span $T_{(x,x,x',x')}(M \times M \times M \times M)$, which has dimension $4m$. Thus for $\bar{\rho}_4(f)$ to be transversal to Δ_4 we must have $\bar{\rho}_4(f)(x) \notin \Delta_4$ for all $x \in M$, so that $f \in \mathcal{B}_6^{(q)}$. Hence $\mathcal{B}^{(q)} \cap \mathcal{B}_6^{(q)}$ is open and dense in $\mathcal{B}^{(q)}$ as required. □

6.2. Embedding the Short Periodic Orbits

Given the definitions in the previous sections, define

$$\tilde{\mathcal{B}}^{(q)} = \mathcal{B}^{(q)} \cap \mathcal{B}_4^{(q)} \cap \mathcal{B}_5^{(q)} \cap \mathcal{B}_6^{(q)}.$$

From now on, we shall restrict ourselves to an $f \in \tilde{\mathcal{B}}^{(2d)}$ and aim to construct an open dense set of φ in $\mathcal{C}^{2r}(M, \mathbb{R})$ for which $\Phi_{f,\varphi,y}$ is an embedding for all $y \in N^{d-1}$, by analogy to the Unstated Takens' Theorem of [Huke, 1993] (Theorems 2.2 and 4.7 of [Stark, 1999]). In fact, as indicated at the beginning of section 6, it is sufficient to prove that such φ are dense, and since there is only a finite number of y in N^{d-1}, we can consider a single y at a time.

As already indicated in section 3.2, the main difficulty occurs at those points $x \in M$ such that $x_i = x_j$ for some $0 \le i < j \le d-1$. In the previous section we showed that for $f \in \tilde{\mathcal{B}}^{(2d)}$ the number of such points was finite, and our next aim is to construct an open dense set $\mathcal{A}_f$ of φ in $\mathcal{C}^{2r}(M, \mathbb{R})$ for which $\Phi_{f,\varphi,y}$ embeds this finite set for all $y \in N^{d-1}$. Thus recall from section 3.2 that we define

$$\tilde{\mathcal{P}}_{y_0 \ldots y_{q-1}} = \{x \in M : f_{y_{i-1} \ldots y_0}(x) = f_{y_{j-1} \ldots y_0}(x) \text{ for some } i \neq j \text{ with } 0 \le i < j \le q\}.$$

Thus in terms of the previous section,

$$\tilde{\mathcal{P}}_{y_0 \ldots y_{q-1}} = \bigcup_{0 \le i \le j < q} (f_{y_{i-1} \ldots y_0})^{-1}(\mathcal{P}_{y_i \ldots y_{j-1}}).$$

Since $f_{y_{i-1} \ldots y_0}$ is invertible and $\mathcal{P}_{y_i \ldots y_{j-1}}$ is a finite set if $f \in \mathcal{B}^{(q')}$ for any $0 < q \le q'$, this means that $\tilde{\mathcal{P}}_{y_0 \ldots y_{q-1}}$ is finite for each $(y_0, y_1, \ldots, y_{q-1}) \in N^q$ and hence

$$\tilde{\mathcal{P}}^{(q)} = \bigcup_{(y_0, \ldots, y_{q-1}) \in N^q} \tilde{\mathcal{P}}_{y_0 \ldots y_{q-1}}$$

is finite if $f \in \mathcal{B}^{(q')}$ for any $0 < q \le q'$. Note that if $x \in \tilde{\mathcal{P}}^{(q)}$ then it is possible to have $f_{y_{i-1} \ldots y_0}(x) \in \mathcal{P}_{y_i \ldots y_{j-1}}$ and $f_{y'_{i'-1} \ldots y'_0}(x) \in \mathcal{P}_{y'_{i'} \ldots y'_{j'-1}}$ for some $(y_0, y_1, \ldots, y_{j-1}) \neq (y'_0, y'_1, \ldots, y'_{j'-1})$. One could exclude this using a proof similar to that of Lemma 6.23, but in fact we do not need to do this. We are now in a position to give the analogues of Propositions 4.2 and 4.3 of [Stark, 1999].

Proposition 6.24. *The set of φ such that φ is $1-1$ on $\tilde{\mathcal{P}}^{(2d-1)}$ is open and dense in $\mathcal{C}^{2r}(M, \mathbb{R})$.*

Proof. This is intuitively obvious since $\tilde{\mathcal{P}}^{(2d-1)}$ consists of a finite number of points, and so we can perturb φ independently on each point to ensure that it takes a distinct value there. Furthermore, if φ is injective, then clearly so is any φ' in an open neighbourhood of φ. A formal proof is identical to the proof of Proposition 4.2 in section 4.2.2 of [Stark, 1999]. □

Note that since φ is the first component of $\Phi_{f,\varphi,y}$, this means that $\Phi_{f,\varphi,y}$ is also injective on $\tilde{\mathcal{P}}^{(2d-1)}$ for an open dense set of φ, i.e. $\Phi_{f,\varphi,y}(x) \neq \Phi_{f,\varphi,y}(x')$ for all $x, x' \in \tilde{\mathcal{P}}^{(2d-1)}$ such that $x \neq x'$. The reason that we want injectivity on $\tilde{\mathcal{P}}^{(2d-1)}$ rather than just $\tilde{\mathcal{P}}^{(d-1)}$ is that this simplifies the treatment of injectivity in the general case below. By contrast, we only need $\Phi_{f,\varphi,y}$ to be immersive on $\tilde{\mathcal{P}}^{(d-1)}$. In fact, it turns out to be simplest to phrase this in terms of each individual $\tilde{\mathcal{P}}_{y_0 \dots y_{d-2}}$:

Proposition 6.25. *Suppose that $x \in \tilde{\mathcal{P}}_{y_0 \dots y_{d-2}}$ for some $y = (y_0, y_1, \dots, y_{d-2}) \in N^{d-1}$; then the set of φ such that $T_x \Phi_{f,\varphi,y}$ has rank m is open and dense in $\mathcal{C}^{2r}(M, \mathbb{R})$.*

Proof. The kth component of $T_x \Phi_{f,\varphi,y}$ is given by $a_k = T_{x_k} \varphi \circ T_x f_{y_{k-1} \dots y_0}$ and we thus need to show that m such linear maps are independent. Since $x \in \tilde{\mathcal{P}}_{y_0 \dots y_{d-2}}$, we have $x_i = x_j$ for some $0 \leq i < j \leq d-1$. We shall separately treat the two cases where $b^+(i,j) - b^-(i,j) \geq m$, and $b^+(i,j) - b^-(i,j) < m$. Thus first consider the former, which is simply the proof of Proposition 4.3 in section 4.2.3 of [Stark, 1999] applied to the maximal periodic segment $\{x_{b^-(i,j)}, \dots, x_{b^+(i,j)}\}$. Without loss of generality, we may assume that $b^-(i,j) = i$ (this has no effect on the details of the proof, but serves to considerably simplify the notation). Let p be the minimal period of x_i for $(y_i, y_{i+1}, \dots, y_{j-1})$. Given any k such that $i \leq k \leq i+m-1$, since $i+m-1 < b^+(i,j)$, we have $x_k = x_{(k \bmod (i,i+p))}$ by Lemma 6.21. Hence $a_k = a_{(k \bmod (i,i+p))} \circ A^r$ where r is such that $k = k \bmod (i, i+p) + rp$, and $A = T_{x_i} f_{y_{i+p-1} \dots y_i}$. Thus $(a_i, a_{i+1}, \dots, a_{i+m-1}) = (a_i, a_{i+1}, \dots, a_{i+p-1}, a_i \circ A, a_{i+1} \circ A, \dots, a_{i+p-1} \circ A, \dots, a_{i+s} \circ A^{r'})$, where $s = (m-1-i) \bmod p$ and $m-1 = i+s+rp$. By [Huke, 1993] (see Lemma 4.9 of [Stark, 1999]) this implies that $(a_i, a_{i+1}, \dots, a_{i+m-1})$ are linearly independent, and hence $T_x \Phi_{f,\varphi,y}$ has rank m, for a dense open set of φ.

We now turn to the case $b^+(i,j) - b^-(i,j) < m$. Then since $f \in \tilde{\mathcal{B}}^{(2d)}$ we cannot have $x_{i'} = x_{j'}$ for any i', j' such that $0 \leq i' < j' < b^-(i,j)$, or such that $b^+(i,j) < i' < j' \leq d-1$. Nor can we have $x_{i'} = x_{j'}$ with $0 \leq i' < b^-(i,j)$ and $b^+(i,j) < j' \leq d-1$ since in that case we would have $b^-(i',j') < b^-(i,j) < b^+(i,j) < b^+(i',j')$. Since $f \in \mathcal{B}_2^{(q)}$, this means that $(y_{b^-(i,j)}, y_{b^-(i,j)+1}, \dots, y_{b^+(i,j)-1})$ and $(y_{b^-(i',j')}, y_{b^-(i',j')+1}, \dots, y_{b^+(i',j')-1})$ have the same prime factor, violating the maximality of $(y_{b^-(i,j)}, y_{b^-(i,j)+1}, \dots, y_{b^+(i,j)-1})$. Thus $\{x_0, \dots, x_{b^-(i,j)-1}\} \cup \{x_{b^+(i,j)+1}, \dots, x_{d-1}\}$ are distinct. This set consists of $b^-(i,j) + d - 1 - b^+(i,j) > m-1$ points since $d > 2m$ and $b^+(i,j) - b^-(i,j) < m$. Hence $\{x_0, \dots, x_{b^-(i,j)-1}\} \cup \{x_{b^+(i,j)+1}, \dots, x_{d-1}\}$ consists of at least m distinct points. It is now intuitively obvious that for a dense open set of φ, we can choose $T_{x_k}\varphi$ on these points so that $(a_0, \dots, a_{b^-(i,j)-1}; a_{b^+(i,j)+1}, \dots, a_{d-1})$ has rank m. Formally, this follows from the fact, proved in Corollary C.18 of [Stark, 1999], that the map $\varphi \mapsto (T_{x_0}\varphi, T_{x_1}\varphi, \dots, T_{x_{b^-(i,j)-1}}\varphi, T_{x_{b^+(i,j)+1}}\varphi, \dots, T_{x_{d-1}}\varphi)$ is a submersion. But $(a_0, \dots, a_{b^-(i,j)-1}; a_{b^+(i,j)+1}, \dots, a_{d-1})$ is a subset of the components of $T_x \Phi_{f,\varphi,y}$, and hence if it has rank m, so has $T_x \Phi_{f,\varphi,y}$. □

We now define $\mathcal{A}_f$ to be the set of φ such that φ is $1-1$ on $\tilde{\mathcal{P}}^{(2d-1)}$ and $T_x \Phi_{f,\varphi,y}$ has rank m for all $x \in \tilde{\mathcal{P}}_{y_0 \dots y_{d-2}}$. By the previous two propositions, $\mathcal{A}_f$ is open and dense in

$\mathcal{C}^{2r}(M, \mathbb{R})$. Now, as in Lemma 4.4 of [Stark, 1999], let

$$\mathcal{E}^{2r} = \{(f, \varphi) \in \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R}) : f \in \tilde{\mathcal{B}}^{(2d)}, \varphi \in \mathcal{A}_f\}.$$

The same proof as in [Stark, 1999] shows that $\mathcal{E}^{2r}$ is open and dense in $\mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R})$. It is thus a Banach manifold and $T_{(f,\varphi)}\mathcal{E}^{2r} = T_{(f,\varphi)}\mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R})$ for any $(f, \varphi) \in \mathcal{E}^{2r}$. From now on, we shall restrict ourselves to $\mathcal{E}^{2r}$.

6.3. Immersivity of Φ

This is identical to the proof for the standard Takens' Theorem in section 4 of [Stark, 1999]. For a given $y \in N^{d-1}$, define τ_y: $\mathcal{E}^{2r} \to \mathcal{C}^{r-1}(\tilde{T}M, T\mathbb{R}^d)$ by

$$\tau_y(f, \varphi) = T\Phi_{f,\varphi,y}, \tag{6.7}$$

where $\Phi_{f,\varphi,y}$ is defined in (4.1) by $\Phi_{f,\varphi,y}(x) = \Phi_{f,\varphi}(x, y)$. Thus $ev_{\tau_y}(f, \varphi, v) = ev_\tau(f, \varphi, v, y)$, where $\tau(f, \varphi) = T_1\Phi_{f,\varphi}$ as in (4.3). Thus by Lemma 4.4, ev_{τ_y} is $\mathcal{C}^{r-1}$ and $T_{(f,\varphi,v)}ev_{\tau_y}(0_f, \xi, 0_x) = (\bar{\omega}(T_{x_0}\xi(v_0)), \bar{\omega}(T_{x_1}\xi(v_1)), \ldots, \bar{\omega}(T_{x_{d-1}}\xi(v_{d-1})))^\dagger$, where as usual $v_i = T_x f_{y_{i-1}\ldots y_0}(v)$, $x = \tau_M(v)$, and $x_i = f_{y_{i-1}\ldots y_0}(x, y)$. Let L be the 0 section in $T\mathbb{R}^d$, and then, exactly as in Proposition 4.5 of [Stark, 1999], we have

Proposition 6.26. *Given any $y \in N^{d-1}$, ev_{τ_y} is transversal to L.*

Proof. Suppose $ev_{\tau_y}(f, \varphi, v) \in L$, so that $T_x\Phi_{f,\varphi,y}(v) = 0$. Since $v \in \tilde{T}M$, we have $\|v\| = 1$ and hence $v \neq 0$. Thus rank $T_x\Phi_{f,\varphi,y} < m$, and since $\varphi \in \mathcal{A}_f$, we cannot have $x \in \tilde{\mathcal{P}}_{y_0\ldots y_{d-2}}$. Thus the points $\{x_0, \ldots, x_{d-1}\}$ are distinct. Furthermore, since $f_{y_{i-1}\ldots y_0}$ is a diffeomorphism, $v_i \neq 0$ for all i. Hence, by Corollary C.16 of [Stark, 1999] $T_{(f,\varphi,v)}ev_{\tau_y}$ is surjective, and hence ev_{τ_y} is transversal to L. $\square$

Now, as usual, $\dim \tilde{T}M = 2m - 1$, $\operatorname{codim} L = d$ and so if $d \geq 2m - 1$, we have $\dim \tilde{T}M - \operatorname{codim} L < 0 < r - 1$. Thus by the Parametric Transversality Theorem, $\tau_y(f, \varphi)$ is transversal to L for an open dense set of $(f, \varphi) \in \mathcal{E}^{2r}$ and hence an open dense set in $\mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R})$. Fix any such (f, φ), and suppose that $\tau_y(f, \varphi)(v) = T_x\Phi_{f,\varphi,y}(v) \in L$ for some $v \in \tilde{T}M$. Then $\dim T_v(\tilde{T}M) = 2m - 1$, and so the dimension of $T_v[T\Phi_{f,\varphi,y}](T_v(\tilde{T}M))$ cannot be greater than $2m - 1$. On the other hand $\dim T_uL = d$, and $\dim T_u(T\mathbb{R}^d) = 2d$, for any $u \in T\mathbb{R}^d$. Thus if $d \geq 2m - 1$, then dim Image $T_v(T_x\Phi_{f,\varphi,y}) + \dim T_uL < \dim T_u(T\mathbb{R}^d)$, so that Image $T_v(T_x\Phi_{f,\varphi,y}) + T_uL$ cannot span $T_u(T\mathbb{R}^d)$. Thus the only way for $\tau_y(f, \varphi) = T\Phi_{f,\varphi,y}$ to be transversal to L is for the image of $T\Phi_{f,\varphi,y}$ not to intersect L. Thus $T_x\Phi_{f,\varphi,y}(v) \neq 0$ for all $v \in \tilde{T}M$, i.e. $\Phi_{f,\varphi,y}$ is immersive for an open dense set of $(f, \varphi) \in \mathcal{E}^{2r}$. Taking the intersection over the finite number of $y \in N^{d-1}$, we obtain an open dense set for which $\Phi_{f,\varphi,y}$ is immersive for all y.

6.4. Injectivity of Φ

Similarly to section 4 of [Stark, 1999], define ρ_y: $\mathcal{E}^{2r} \to \mathcal{C}^r((M \times M)\backslash\Delta, \mathbb{R}^d \times \mathbb{R}^d)$ for a given $y \in N^{d-1}$ by

$$\rho_y(f, \varphi)(x, x') = (\Phi_{f,\varphi,y}(x), \Phi_{f,\varphi,y}(x')),$$

where $\Delta = \{(x, x') \in M \times M : x = x'\}$ is the diagonal in $M \times M$. Then $ev_{\rho_y}(f, \varphi, x, x') = (ev_\rho(f, \varphi, x, y), ev_\rho(f, \varphi, x', y))$ where $\rho(f, \varphi) = \Phi_{f,\varphi}$ is as in (4.2). Thus ev_{ρ_y} is $\mathcal{C}^r$ by Lemma 4.3 and $T_{(f,\varphi,x,x')}ev_{\rho_y}(0_\eta, \xi, 0_x, 0_{x'}) = ((\xi(x), \xi(x_1), \ldots, \xi(x_{d-1})), (\xi(x'), \xi(x'_1), \ldots, \xi(x'_{d-1})))^\dagger$. Let Δ_d be the diagonal in $\mathbb{R}^d \times \mathbb{R}^d$, and suppose that $ev_{\rho_y}(f, \varphi, x, x') \in \Delta_d$, so that $\Phi_{f,\varphi,y}(x) = \Phi_{f,\varphi,y}(x')$. Then in particular $\varphi(x) = \varphi(x')$ but $x \neq x'$ and thus by the definition of $\mathcal{A}_f$, at least one of x or x' cannot be in $\tilde{\mathcal{P}}^{(2d-1)}$. We shall suppose without loss of generality that this is x. Then

Lemma 6.27. *Suppose $f \in \tilde{\mathcal{B}}^{(2d)}$, $x \neq x'$, $x \notin \tilde{\mathcal{P}}^{(2d-1)}$, and for some $i, i' \in \{0, 1, \ldots, d-1\}$ we have $x_i = x'_{i'}$. Then $x_i \neq x'_{i''}$ for all $i'' \in \{0, 1, \ldots, d-1\}$ such that $i' \neq i''$ and $x'_{i'} \neq x_{i''}$ for all $i'' \in \{0, 1, \ldots, d-1\}$ such that $i \neq i''$.*

Proof. If $x_i = x'_{i''}$ for some i'' such that $i' \neq i''$, then $x'_{i'} = x'_{i''}$ and hence $x'_{i'} = x_i$ is a periodic point of period less than $2d$, contradicting $x \notin \tilde{\mathcal{P}}^{(2d-1)}$. Similarly if $x'_{i'} = x_{i''}$ for some i'' such that $i \neq i''$, then $x_i = x'_{i'} = x_{i''}$, so that x_i is periodic, again contradicting $x \notin \tilde{\mathcal{P}}^{(2d-1)}$. □

Given any $i \in \{0, 1, \ldots, d-1\}$, it thus makes sense to define a *chain* starting at $i = i_0$ to be a sequence $(i_0, i_1, \ldots, i_{k-1})$ with $i_j \in \{0, 1, \ldots, d-1\}$ and $x_{i_j} = x'_{i_{j+1}}$ for all $j \in \{0, 1, \ldots, k-1\}$.

Lemma 6.28. *Suppose $f \in \tilde{\mathcal{B}}^{(2d)}$, $x \neq x'$, $x \notin \tilde{\mathcal{P}}^{(2d-1)}$, and $(i_0, i_1, \ldots, i_{k-1})$ is a chain starting at $i = i_0$. Then $\{i_0, i_1, \ldots, i_{k-1}\}$ are distinct, i.e. $i_j \neq i_{j'}$ for any $0 \leq i \neq i' \leq k-1$.*

Proof. Since $f_{y_{i-1}\ldots y_0}$ is a diffeomorphism for all i, we have $x_i = x'_i$ for all $i \in \{0, 1, \ldots, d-1\}$ and hence $i_j \neq i_{j+1}$ for all $j \in \{0, 1, \ldots, \kappa - 2\}$. Hence either $i_j < i_{j+1}$ or $i_j > i_{j+1}$; suppose the former. Then we claim that $i_{j+1} < i_{j+2}$. Suppose not, so that $i_{j+1} > i_{j+2}$, and denote $i_j = i$, $i_{j+1} = i'$ and $i_{j+2} = i''$, so that $i' > i$ and $i' > i''$. Then $x_{i'} = f_{y_{i'-1}\ldots y_i}(x_i)$, $x'_{i'} = f_{y_{i'-1}\ldots y_{i''}}(x'_{i''})$ and $x_i = x'_{i'}$, $x_{i'} = x'_{i''}$. This implies that $f_{y_{i'-1}\ldots y_{i''} y_{i'-1}\ldots y_i}(x_i) = f_{y_{i'-1}\ldots y_{i''}}(f_{y_{i'-1}\ldots y_i}(x_i)) = f_{y_{i'-1}\ldots y_{i''}}(x_{i'}) = f_{y_{i'-1}\ldots y_{i''}}(x'_{i''}) = x'_{i'} = x_i$. Hence x_i is periodic for $(y_i, y_{i+1}, \ldots, y_{i'-1}, y_{i''}, y_{i''+1}, \ldots, y_{i'-1})$, contradicting $x \notin \tilde{\mathcal{P}}^{(2d-1)}$. Similarly if $i_j > i_{j+1}$ then $i_{j+1} > i_{j+2}$. Repeating this argument for $j = 0, 1, \ldots, \kappa - 2$, we have either $i_0 < i_1 < \ldots < i_{\kappa-1}$ or $i_0 > i_1 > \ldots > i_{\kappa-1}$. □

Since $\{i_0, i_1, \ldots, i_{\kappa-1}\} \subset \{0, 1, \ldots, d-1\}$, this lemma implies that $\kappa \leq d$, and in particular, every chain has to be finite. Thus we may define a *maximal chain* to be one that cannot be extended (in either direction), i.e. a chain such that $x'_{i_0} \notin \{x_0, x_1, \ldots, x_{d-1}\}$ and $x_{i_{k-1}} \notin \{x'_0, x'_1, \ldots, x'_{d-1}\}$. Observe that if $x_i \notin \{x'_0, x'_1, \ldots, x'_{d-1}\}$ and $x'_i \notin \{x_0, x_1, \ldots,$

$x_{d-1}\}$, then the maximal chain starting at i is (i). Not every i has a maximal chain starting from it, but

Lemma 6.29. *Suppose $f \in \tilde{\mathcal{B}}^{(2d)}$, $x \neq x'$, and $x \notin \tilde{\mathcal{P}}^{(2d-1)}$. Then the maximal chains give a disjoint partition of $\{0, 1, \ldots, d-1\}$, i.e. every $i \in \{0, 1, \ldots, d-1\}$ lies in precisely one maximal chain.*

Proof. Given a chain $(i_0, i_1, \ldots, i_{\kappa-1})$, we have either $x_{i_{k-1}} \notin \{x'_0, x'_1, \ldots, x'_{d-1}\}$ or $x_{i_{k-1}} = x'_{i'}$ for some $i' \in \{0, 1, \ldots, d-1\}$, in which case $(i_0, i_1, \ldots, i_{\kappa-1}, i')$ is also a chain and by Lemma 6.27 the choice of i' is unique. Similarly, either $x'_{i_0} \notin \{x_0, x_1, \ldots, x_{d-1}\}$ or $x'_{i_0} = x_{i''}$ for some unique $i'' \in \{0, 1, \ldots, d-1\}$ and $(i'', i_0, i_1, \ldots, i_{\kappa-1})$ is a chain. Since by Lemma 6.28 every chain has length less than or equal to d, we can extend $(i_0, i_1, \ldots, i_{\kappa-1})$ only a finite number of times in either direction. When the chain can no longer be extended, it is maximal, and hence $(i_0, i_1, \ldots, i_{\kappa-1})$ is a subchain of a unique maximal chain. Every i lies in at least the one point chain (i), and hence it lies in at least one maximal chain. But by Lemma 6.27, two distinct maximal chains cannot intersect, and hence each i lies in a unique maximal chain. ◻

We now have the tools required to prove

Proposition 6.30. *Given any $y \in N^{d-1}$, ev_{ρ_y} is transversal to Δ_d.*

Proof. Suppose $ev_{\rho_y}(f, \varphi, x, x') \in \Delta_d$, so that as above without loss of generality we have $x \notin \tilde{\mathcal{P}}^{(2d-1)}$. Denote $z = \Phi_{f,\varphi,y}(x) = \Phi_{f,\varphi,y}(x')$, so that $ev_{\rho_y}(f, \varphi, x, x') = (z, z)$ and for convenience write $\Xi = (f, \varphi, x, x')$. Let e_i be a basis vector for $T_{z_i}\mathbb{R}$ and $e^{(i)}$ be the vector $(0, \ldots, e_i, \ldots, 0) \in T_z\mathbb{R}^d$, so that $e^{(0)}, e^{(1)}, \ldots, e^{(d-1)}$ forms a basis for $T_z\mathbb{R}^d$. Denote by S the span of Image $T_\Xi ev_{\rho_y} + T_{(z,z)}\Delta_d$. To show that ev_{ρ_y} is transversal to Δ_d, we thus need to prove that $S = T_{(z,z)}(\mathbb{R}^d \times \mathbb{R}^d) = T_z\mathbb{R}^d \times T_z\mathbb{R}^d$. Since $(u, u) \in T_{(z,z)}\Delta_d$ for any $u \in T_z\mathbb{R}^d$, it is sufficient to show that $(e^{(i)}, 0) \in S$ for all $i \in \{0, 1, \ldots, d-1\}$. By Lemma 6.29 the maximal chains form a partition of $\{0, 1, \ldots, d-1\}$ and hence it is sufficient to show that $(e^{(i_j)}, 0) \in S$ for all $j \in \{0, 1, \ldots, \kappa-1\}$ where $(i_0, i_1, \ldots, i_{\kappa-1})$ is a maximal chain. We do this inductively, starting from the end of the chain. Since $x \notin \tilde{\mathcal{P}}^{(2d-1)}$, the points $\{x_0, \ldots, x_{d-1}\}$ are distinct, and $x_{i_{k-1}} \notin \{x'_0, x'_1, \ldots, x'_{d-1}\}$ because the chain is maximal. Thus as in Proposition 5.5, by Corollary C.12 of [Stark, 1999] there exists a $\xi \in T_\varphi\mathcal{C}^r(M, \mathbb{R})$ such that $\xi(x_{i_{k-1}}) = e_{i_{k-1}}$, $\xi(x_i) = 0$ for all $i \neq i_{\kappa-1}$, and $\xi(x'_i) = 0$ for all i. Thus $T_\Xi ev_{\rho_y}(0_\eta, \xi, 0_x, 0_{x'}) = (e^{(i_{\kappa-1})}, 0)$, and hence $(e^{(i_{\kappa-1})}, 0) \in S$. Since $(e^{(i_{\kappa-1})}, e^{(i_{\kappa-1})}) \in T_{(z,z)}\Delta_d \subset S$, we also have $(e^{(i_{\kappa-1})}, e^{(i_{\kappa-1})}) - (e^{(i_{\kappa-1})}, 0) = (0, e^{(i_{\kappa-1})}) \in S$. Now $x_{i_{k-2}} = x'_{i_{k-1}}$ and by Lemma 6.27, $x_{i_{k-2}} \neq x'_i$ for all $i \neq i_{\kappa-1}$ and $x_{i_{k-2}} \neq x_i$ for all $i \neq i_{\kappa-2}$. Thus, as above, we can find a $\xi \in T_\varphi\mathcal{C}^r(M, \mathbb{R})$ such that $T_\Xi ev_{\rho_y}(0_\eta, \xi, 0_x, 0_{x'}) = (e^{(i_{\kappa-2})}, e^{(i_{\kappa-1})})$, and hence $(e^{(i_{\kappa-2})}, e^{(i_{\kappa-1})}) \in S$. We have already shown that $(0, e^{(i_{\kappa-1})}) \in S$, and so $(e^{(i_{\kappa-2})}, e^{(i_{\kappa-1})}) - (0, e^{(i_{\kappa-1})}) = (e^{(i_{\kappa-2})}, 0) \in S$. Repeating this argument $\kappa - 1$ times, we show that $(e^{(i_j)}, 0) \in S$ for all $j \in \{0, 1, \ldots, k-1\}$, as required. ◻

Now, as usual, we apply the Parametric Transversality Theorem and count dimensions. We have $\dim(M \times M)\backslash\Delta = 2m$, $\operatorname{codim}\Delta_d = d$, and so if $d \geq 2m - 1$, we have $\dim(M \times M)\backslash\Delta - \operatorname{codim}\Delta_d < 0 < r$. Thus by the Parametric Transversality Theorem, $\rho_y(f, \varphi)$ is transversal to Δ_d for a residual set of $(f, \varphi) \in \mathcal{E}^{2r}$ and hence a residual set in $\mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M, \mathbb{R})$ (we do not get an open set since $(M \times M)\backslash\Delta$ is not closed, but as indicated at the beginning of section 6, density is sufficient). Fix any such (f, φ), and suppose that $\rho_y(f, \varphi)(x, x', y) = (\Phi_{f,\varphi,y}(x), \Phi_{f,\varphi,y}(x')) \in \Delta_d$ for some $(x, x') \in (M \times M)\backslash\Delta$. Then $\dim T_{(x,x')}((M \times M)\backslash\Delta) = 2m$, and so the dimension of $T_{(x,x')}[\Phi_{f,\varphi,y} \times \Phi_{f,\varphi,y}](T_{(x,x')}((M \times M)\backslash\Delta))$ cannot be greater than $2m$. On the other hand $\dim T_{(z,z)}\Delta_d = d$, and $\dim T_{(z,z)}(\mathbb{R}^d \times \mathbb{R}^d) = 2d$. Thus if $d \geq 2m - 1$, then $\dim \operatorname{Image} T_{(x,x')}[\Phi_{f,\varphi,y} \times \Phi_{f,\varphi,y}] + \dim T_{(z,z)}\Delta_d < \dim T_{(z,z)}(\mathbb{R}^d \times \mathbb{R}^d)$, so that Image $T_{(x,x')}[\Phi_{f,\varphi,y} \times \Phi_{f,\varphi,y}] + \dim T_{(z,z)}\Delta_d$ cannot span $T_{(z,z)}(\mathbb{R}^d \times \mathbb{R}^d)$. Thus the only way for $\rho_y(f, \varphi) = \Phi_{f,\varphi,y} \times \Phi_{f,\varphi,y}$ to be transversal to Δ_d is for the image of $\Phi_{f,\varphi,y} \times \Phi_{f,\varphi,y}$ not to intersect Δ_d. Thus $\Phi_{f,\varphi,y}(x) \neq \Phi_{f,\varphi,y}(x')$ for all $x \neq x'$, i.e. $\Phi_{f,\varphi,y}$ is injective for a residual set of $(f, \varphi) \in \mathcal{E}^{2r}$. Taking the intersection over the finite number of $y \in N^{d-1}$, we obtain a residual set for which $\Phi_{f,\varphi,y}$ is injective for all y, as required.

7. Theorem 2.5: Variations and Proofs

Since $\Phi_{f,\varphi,\omega,\eta}$ depends only on $\omega_0, \ldots, \omega_{d-2}$ and $\eta_0, \ldots, \eta_{d-1}$, then as for Theorem 2.3, it is sufficient to consider

$$\Phi_{f,\varphi,y,z}(x) = (\varphi(f^{(0)}(x, y), z_0), \varphi(f^{(1)}(x, y), z_1), \ldots, \varphi(f^{(d-1)}(x, y), z_{d-1}))^{\dagger},$$

for $(y, z) \in N^{d-1} \times (N')^d$. The reduction to this case is carried out exactly as in section 3.1, and the details are therefore left to the reader.

7.1. The Dimension of N and N′

Depending on the dimensionality of N and N', a number of versions of Theorem 2.5 are possible. By far the easiest case is when $\dim N' > 0$. Then $\eta_0, \ldots, \eta_{d-1}$ are distinct for $\mu'_{\Sigma'}$-almost all η, and hence the maps $\varphi_{\eta_0}, \varphi_{\eta_1}, \ldots, \varphi_{\eta_{d-1}}$ are independent. Theorem 2.5 then amounts to little more than the standard Whitney Embedding Theorem (e.g. [Hirsch, 1976]), and in fact can be strengthened to give

Theorem 7.1. *Let M, N, and N' be compact manifolds, with $m = \dim M > 0$ and $\dim N' > 0$. Suppose that $d \geq 2m + 1$ and $r \geq 1$. Given any $f \in \mathcal{D}^r(M \times N, M)$ there exists a residual set of $\varphi \in \mathcal{C}^r(M \times N', \mathbb{R})$ such that for any φ in this set there is an open dense set $\Sigma_{f,\varphi} \subset \Sigma \times \Sigma'$ such that $\Phi_{f,\varphi,\omega,\eta}$ is an embedding for all $(\omega, \eta) \in \Sigma_{f,\varphi}$. If μ_Σ and $\mu'_{\Sigma'}$ are invariant probability measures on Σ and Σ' such that μ_{d-1} and μ'_d are absolutely continuous with respect to Lebesgue measure on N^{d-1} and $(N')^d$ respectively, then we can choose $\Sigma_{f,\varphi}$ such that $\mu_\Sigma \times \mu'_{\Sigma'}(\Sigma_{f,\varphi}) = 1$.*

Using the same procedure as in section 3.1 this follows from

Theorem 7.2. *Let M, N, and N' be compact manifolds, with* $m = \dim M > 0$ *and* $\dim N' > 0$. *Suppose that $d \geq 2m+1$ and $r \geq 1$, and let ν_{d-1} and ν'_d denote normalized Lebesgue measure on N^{d-1} and $(N')^d$ respectively. Given any $f \in \mathcal{D}^r(M \times N, M)$ there exists a residual set of $\varphi \in \mathcal{C}^r(M \times N', \mathbb{R})$ such that for any φ in this set there is an open dense set $U_{f,\varphi} \subset N^{d-1} \times (N')^d$ such that $\nu_{d-1} \times \nu'_d(U_{f,\varphi}) = 1$ and the delay map $\Phi_{f,\varphi,y,z}$ is an embedding for all $(y, z) \in U_{f,\varphi}$.*

We give an easy direct proof of this, independent of the other proofs in this paper. Essentially, this proof amounts to a transversality proof of the Whitney Embedding Theorem (e.g. [Hirsch, 1976]). It is sketched in section 7.2 below. Observe that a minor modification of this proof can be used to show that given any $f \in \mathcal{D}^r(M \times N, M)$ and any $y \in N^{d-1}$ there is a residual set of $\varphi \in \mathcal{C}^r(M \times N', \mathbb{R})$ for which $\Phi_{f,\varphi,y,z}$ is an embedding for an open dense set of z of full measure. However, this residual set of φ depends on y, and hence this version does not seem to be useful from the point of applications.

An alternative approach to Theorem 2.5 is to repeat the transversality arguments of sections 5.3 and 5.4 (when $\dim N > 0$) and sections 6.3 and 6.4 (when $\dim N = 0$), replacing φ by $\varphi_{j_1}, \varphi_{j_2}, \ldots, \varphi_{j_\alpha}$, where $\varphi_j(x) = \varphi(x, z_j)$. This works for any dimension of N', but we obtain slightly stronger results when $\dim N' = 0$, and hence we restrict ourselves to this case.

Theorem 7.3. *Let M, N, and N' be compact manifolds, with* $m = \dim M > 0$, $\dim N > 0$, *and* $\dim N' = 0$. *Suppose that $d \geq 2m+1$ and $r \geq 1$. Then there exists a residual set of $(f, \varphi) \in \mathcal{D}^r(M \times N, M) \times \mathcal{C}^r(M \times N', \mathbb{R})$ such that for any (f, φ) in this set there is an open dense set $\Sigma_{f,\varphi} \subset \Sigma$ such that $\Phi_{f,\varphi,\omega,\eta}$ is an embedding for all $(\omega, \eta) \in \Sigma_{f,\varphi} \times \Sigma'$. If μ_Σ is an invariant probability measure on Σ, such that μ_{d-1} is absolutely continuous with respect to Lebesgue measure on N^{d-1}, then we can choose $\Sigma_{f,\varphi}$ such that $\mu_\Sigma(\Sigma_{f,\varphi}) = 1$.*

As usual this is deduced from the final dimensional version, using the procedure given in section 3.1:

Theorem 7.4. *Let M, N, and N' be compact manifolds, with* $m = \dim M > 0$, $\dim N > 0$, *and* $\dim N' = 0$. *Suppose that $d \geq 2m+1$ and $r \geq 1$, and let ν_{d-1} be normalized Lebesgue measure on N^{d-1}. Then there exists a residual set of $(f, \varphi) \in \mathcal{D}^r(M \times N, M) \times \mathcal{C}^r(M \times N', \mathbb{R})$ such that for any (f, φ) in this set there is an open dense set $U_{f,\varphi} \subset N^{d-1}$ such that $\nu_{d-1}(U_{f,\varphi}) = 1$ and the delay map $\Phi_{f,\varphi,y,z}$ is an embedding for all $(y, z) \in U_{f,\varphi}(N')^d$.*

The proof is sketched in section 7.3 below. Finally we consider the case $\dim N = 0$ and $\dim N' = 0$:

Theorem 7.5. *Let M, N, and N' be compact manifolds, with* $m = \dim M > 0$, $\dim N = \dim N' = 0$. *Suppose that $d \geq 2m+1$ and $r \geq 1$. Then there exists an open dense set of $(f, \varphi) \in \mathcal{D}^r(M \times N, M) \times \mathcal{C}^r(M \times N', \mathbb{R})$ such that for any (f, φ) in this set $\Phi_{f,\varphi,\omega,\eta}$ is an embedding for all $(\omega, \eta) \in \Sigma \times \Sigma'$.*

This immediately follows from the final dimensional version, which is proved in section 7.4 below.

Theorem 7.6. *Let M, N, and N' be compact manifolds, with $m = \dim M > 0$, $\dim N = \dim N' = 0$. Suppose that $d \geq 2m+1$ and $r \geq 1$. Then there exists an open dense set of $(f, \varphi) \in \mathcal{D}^r(M \times N, M) \times \mathcal{C}^r(M \times N', \mathbb{R})$ such that for any (f, φ) in this set $\Phi_{f,\varphi,y,z}$ is an embedding for all $(y, z) \in N^{d-1} \times (N')^d$.*

The proofs in this section all follow a familiar pattern: construction of an appropriate evaluation function, proof of its transversality, application of the Parametric Transversality Theorem, and finally dimension counting. As usual we work with a sufficiently smooth f and φ, and reduce to $\mathcal{C}^1$ at the end, and since $F_{f,\varphi,y,z}$ is continuous in all its parameters, and embeddings are open, it is enough to prove that appropriate sets are dense and/or of full measure. We leave most such repetitious technical details to the reader, and focus on constructing the necessary evaluation functions and showing their transversality.

7.2. *Sketch Proof for* $\dim N' > 0$

Although essentially independent of previous proofs in this paper, the proof of Theorem 7.2 still uses familiar ingredients. We start by defining

$$\tilde{N}'_d = \{z \in (N')^d : z_i \neq z_j \text{ for all } i \neq j\},$$

by analogy to (5.1). Since $\mu'_d(\tilde{N}'_d) = 1$, we can restrict ourselves to $z \in \tilde{N}'_d$. As usual, denote $\varphi_{z_i}(x) = \varphi(x, z_i)$. Given $f \in \mathcal{D}^{2r}(M \times N, M)$, define $\tau_f\colon \mathcal{C}^{2r}(M \times N', \mathbb{R}) \to \mathcal{C}^{r-1}(\tilde{T}M \times N^{d-1}\tilde{N}'_d, T\mathbb{R}^d)$ by

$$\tau_f(\varphi)(v, y, z) = T_x \Phi_{f,\varphi,y,z}(v).$$

Then ev_{τ_f} is $\mathcal{C}^{r-1}$ and $T_{(\varphi,v,y,z)} ev_{\tau_f}(\xi, 0_v, 0_y, 0_z) = (\bar{\omega}(T_{x_0}\xi(v_0, z_0)), \bar{\omega}(T_{x_1}\xi(v_1, z_1)), \ldots, \bar{\omega}(T_{x_{d-1}}\xi(v_{d-1}, z_{d-1})))^{\dagger}$. Since the points $\{z_0, \ldots, z_{d-1}\}$ are distinct, we can choose each $T_{x_i}\xi(v_i, z_i)$ independently, and hence $T_{(\varphi,v,y,z)} ev_{\tau_f}$ is surjective. Thus ev_{τ_f} is transversal to the zero section L in $T\mathbb{R}^d$. By the Parametric Transversality Theorem, therefore, there is a residual set of $\varphi \in \mathcal{C}^{2r}(M \times N', \mathbb{R})$ such that $\tau_f(\varphi)$ is transversal to L. Fix any such φ and define $\tau_{f,\varphi}\colon N^{d-1} \times \tilde{N}'_d \to \mathcal{C}^{r-1}(\tilde{T}M, T\mathbb{R}^d)$ by $\tau_{f,\varphi}(y, z) = T\Phi_{f,\varphi,y,z}$. Then $ev_{\tau_{f,\varphi}}$ is $\mathcal{C}^{r-1}$ and $ev_{\tau_{f,\varphi}}(y, z, v) = \tau_{f,\varphi}(y, z)(v) = \tau_f(\varphi)(v, y, z)$, and hence $ev_{\tau_{f,\varphi}}$ is transversal to L. Applying the Measure Theoretic Finite Dimensional Parametric Transversality Theorem (Appendix A), we obtain an open dense set of (y, z) of full $\mu_{d-1} \times \mu'_d$ measure for which $T\Phi_{f,\varphi,y,z}$ is transversal to L. But the dimension of $T_v(\tilde{T}M)$ is $2m-1$, that of $T_w L$ is d, and hence the dimension of the span of the image of $T_v(T\Phi_{f,\varphi,y,z})$ and $T_w L$ is at most $2m + d - 1 < 2d$ if $d \geq 2m+1$. Hence as usual, transversality implies that $T\Phi_{f,\varphi,y,z}$ does not intersect L and hence $\Phi_{f,\varphi,y,z}$ is immersive.

To prove injectivity, we define $\rho_f\colon \mathcal{C}^{2r}(M \times N', \mathbb{R}) \to \mathcal{C}^r((M \times M)\backslash\Delta \times N^{d-1} \times \tilde{N}'_d, \mathbb{R}^d \times \mathbb{R}^d)$ by

$$\rho_f(\varphi)(x, x', y, z) = (\Phi_{f,\varphi,y,z}(x), \Phi_{f,\varphi,y,z}(x')).$$

Thus ev_{ρ_f} is $\mathcal{C}^r$ and $T_{(\varphi,x,x',y,z)}ev_{\rho_f}(\xi, 0_x, 0_{x'}, 0_y, 0_z) = ((\xi(x, z_0), \xi(x_1, z_1), \ldots, \xi(x_{d-1}, z_{d-1})), (\xi(x', z_0), \xi(x'_1, z_1), \ldots, \xi(x'_{d-1}, z_{d-1})))^\dagger$. Since for $z \in \tilde{N}'_d$, the points $\{z_0, \ldots, z_{d-1}\}$ are distinct, and $x_i \neq x'_i$ for all i, the points $\{(x_0, z_0), (x_1, z_1), \ldots, (x_{d-1}, z_{d-1}), (x'_0, z_0), (x'_1, z_1), \ldots, (x'_{d-1}, z_{d-1})\}$ are all distinct. Therefore $T_{(\varphi,x,x',y,z)}ev_{\rho_f}$ is surjective, so that ev_{ρ_f} is transversal to the diagonal Δ_d in $\mathbb{R}^d \times \mathbb{R}^d$. By the Parametric Transversality Theorem, there is a residual set of $\varphi \in \mathcal{C}^{2r}(M \times N', \mathbb{R})$ such that $\rho_f(\varphi)$ is transversal to Δ_d. Fix any such φ and define $\rho_{f,\varphi}\colon N^{d-1} \times \tilde{N}'_d \to \mathcal{C}^r((M \times M)\backslash\Delta, \mathbb{R}^d)$ by $\rho_{f,\varphi}(y, z) = \Phi_{f,\varphi,y,z} \times \Phi_{f,\varphi,y,z}$, so that $\rho_{f,\varphi}(y, z)(x, x') = \rho_f(\varphi)(x, x', y, z)$. Then $ev_{\rho_{f,\varphi}}$ is $\mathcal{C}^r$ and $ev_{\rho_{f,\varphi}}(y, z, x, x') = \rho_f(\varphi)(x, x', y, z)$ and hence $ev_{\rho_{f,\varphi}}$ is transversal to Δ_d. Applying the Measure Theoretic Finite Dimensional Parametric Transversality Theorem (Appendix A), we obtain an open dense set of (y, z) of full $\mu_{d-1} \times \mu'_d$ measure for which $\Phi_{f,\varphi,y,z} \times \Phi_{f,\varphi,y,z}$ is transversal to Δ_d. The dimension of $T_{(x,x')}(M \times M)\backslash\Delta = 2m$ and that of $T_{(w,w)}\Delta_d$ is d, and hence the dimension of the span of the image of $T_{(x,x')}(\Phi_{f,\varphi,y,z} \times \Phi_{f,\varphi,y,z})$ and $T_{(w,w)}\Delta_d$ is at most $2m + d < 2d$ if $d \geq 2m + 1$. Hence as usual, transversality implies that $\Phi_{f,\varphi,y,z} \times \Phi_{f,\varphi,y,z}$ does not intersect Δ_d, and hence $\Phi_{f,\varphi,y,z}(x) \neq \Phi_{f,\varphi,y,z}(x')$ for all $x \neq x'$, as required.

7.3. Sketch Proof for dim $N > 0$

The main difference between Theorems 2.3 and 2.5 is that in the latter, φ has an additional argument. If anything, this enhances our ability to make perturbations and hence should not affect the transversality of τ'_I and $\rho'_{I,R}$, which are key to the proof of Theorem 2.3 in sections 5.3 and 5.4. This is indeed the case and the transversality arguments in Propositions 5.4 and 5.5 can be repeated almost verbatim, replacing φ by $\varphi_{j_1}, \varphi_{j_2}, \ldots, \varphi_{j_\alpha}$ and ξ by $\xi_{j_1}, \xi_{j_2}, \ldots, \xi_{j_\alpha}$, to yield a proof of Theorem 2.5 for the case dim $N > 0$. However, since in the case dim $N' > 0$ we have already given a proof in the previous section of a slightly stronger result, we shall restrict ourselves here to the case dim $N' = 0$, in which case we can prove that $\Phi_{f,\varphi,y,z}$ is an embedding for all z. This is because when dim $N' = 0$, there is only a finite number of possible $z \in (N')^d$. We can thus work with a single fixed $z \in (N')^d$, and at the end of the proof simply take the finite intersection of the appropriate residual sets of $(f, \varphi) \in \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M \times N', \mathbb{R})$.

Thus fix $z \in (N')^d$, and by analogy to (5.2) define $\Phi_{f,\varphi,I,z}\colon M \times N^{d-1} \to \mathbb{R}^\alpha$ by

$$\begin{aligned}\Phi_{f,\varphi,I,z}(x, y) &= (\varphi(f^{(j_1)}(x, y), z_{j_1}), \varphi(f^{(j_2)}(x, y), z_{j_2}), \ldots, \varphi(f^{(j_\alpha)}(x, y), z_{j_\alpha}))^\dagger \\ &= (\varphi_1(f^{(j_1)}(x, y)), \varphi_2(f^{(j_2)}(x, y)), \ldots, \varphi_\alpha(f^{(j_\alpha)}(x, y))^\dagger,\end{aligned}$$

where $\varphi_j(x) = \varphi(x, z_j)$ and $\Phi_{f,\varphi,I,y,z}\colon M \to \mathbb{R}^\alpha$ by $\Phi_{f,\varphi,I,y,z}(x) = \Phi_{f,\varphi,I,z}(x, y)$. Then, similarly to (5.4) let $\tau_{I,z}\colon \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M \times N', \mathbb{R}) \to \mathcal{C}^{r-1}(\tilde{T}M \times N^{d-1}, T\mathbb{R}^\alpha)$ by

$$\tau_{I,z}(f, \varphi) = T\Phi_{f,\varphi,I,z},$$

which has the evaluation function $ev_{\tau_{I,z}}(f, \pi, v, y) = \tau_{I,z}(f, \varphi)(v, y) = T_x\Phi_{f,\varphi,I,y,z}(v)$, $x = \tau_M(v)$. As in section 5.3, if $T_x\Phi_{f,\varphi,I,y,z}(v) \neq 0$ for some I, then $T_x\Phi_{f,\varphi,y,z}(v) \neq 0$. Hence to show that $\Phi_{f,\varphi,y,z}$ is immersive, it suffices to prove that for every $v \in \tilde{T}_xM$ there is some I such that $T_x\Phi_{f,\varphi,I,y,z}(v)$ does not lie in L_I. To show this, define $\tau'_{I,z}\colon \mathcal{D}^{2r}(M \times$

$N, M) \times \mathcal{C}^{2r}(M \times N', \mathbb{R}) \to \mathcal{C}^{r-1}(\tilde{T}M \times \tilde{N}_{d-1}, T\mathbb{R}^{\alpha} \times M^d)$ by

$$\tau'_{I,z}(f, \varphi) = (\tau_{I,z}(f, \varphi), \iota_\tau \circ \hat{\rho}(f)),$$

so that, as in section 5.3, $ev_{\tau'_{I,z}}$ is $\mathcal{C}^{r-1}$ and

$$ev_{\tau'_{I,z}}(f, \varphi, v, y) = (T_x\Phi_{f,\varphi,I,y,z}(v), (f^{(0)}(x, y), f^{(1)}(x, y), \ldots, f^{(d-1)}(x, y))).$$

The proof that $ev_{\tau'_{I,z}}$ is transversal to $L_I \times \Delta_I$ is identical to that for $ev_{\tau'_I}$ given in Propositions 5.4. Let $\Xi = (f, \varphi, v, y)$ and suppose that $ev_{\tau'_{I,z}}(\Xi) \in L_I \times \Delta_I$. The second component of $ev_{\tau'_{I,z}}$ is the same as the second component of $ev_{\tau'_I}$ and hence is surjective. Thus given any $u = (u_0, u_1, \ldots, u_{d-1}) \in T_{\bar{\omega}}\Delta_I$, there exists $\eta \in T_f\mathcal{D}^{2r}(M \times N, M)$ and $w \in T_v(TM)$ such that $T_{(f,v,y)}ev_{\hat{\rho}_\tau}(\eta, w, 0_y) = u$, where $ev_{\hat{\rho}_\tau}$ is the evaluation map of $\iota_\tau \circ \hat{\rho}$. The first component is given by $T_\Xi ev_{\tau_{I,z}}(0_f, \xi, 0_x, 0_y) = (\bar{\omega}(T_{x_1}\xi_{j_1}(v_{j_1})), \bar{\omega}(T_{x_2}\xi_{j_2}(v_{j_2})), \ldots, \bar{\omega}(T_{x_\alpha}\xi_{j_\alpha}(v_{j_\alpha})))^\dagger$ where now $\xi \in T_\varphi\mathcal{C}^{2r}(M \times N', \mathbb{R})$ and $\xi_j(x) = \xi(x, z_j)$. Since $\hat{\rho}(f)(x, y) \in \Delta_I$, the points $x_{j_1}, x_{j_2}, \ldots, x_{j_\alpha}$ are all distinct. Thus for any $\bar{u} \in T_0(T\mathbb{R}^\alpha)$ there exists a $\xi \in T_\varphi\mathcal{C}^{2r}(M, \mathbb{R})$ such that $T_\Xi ev_{\tau_{I,z}}(0_f, \xi, 0_x, 0_y) = \bar{u} - T_\Xi ev_{\tau_{I,z}}(\eta, 0_j, w, 0_y)$ and hence $T_\Xi ev_{\tau'_{I,z}}(\eta, \xi, w, 0_y) = (\bar{u}, u)$. Thus $T_\Xi ev_{\tau_{I,z}}$ is surjective so that $T_\Xi ev_{\tau_{I,z}}$ is transversal to $L_I \times \Delta_I$.

By the Parametric Transversality Theorem there is a residual set of $(f, \varphi) \in \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M \times N', \mathbb{R})$ for which $\tau'_{I,z}(f, \varphi)$ is transversal to $L_I \times \Delta_I$. Fix any (f, φ) in this set, and define $\tau'_{f,\varphi,I,z}$: $\tilde{N}_{d-1} \to \mathcal{C}^{r-1}(\tilde{T}M, T\mathbb{R}^\alpha \times M^d)$ by

$$\tau'_{f,\varphi,I,z}(y)(v) = \tau'_{I,z}(f, \varphi)(v, y),$$

so that $ev_{\tau'_{f,\varphi,I,z}}$ is transversal to $L_I \times \Delta_I$. Therefore, by the Measure Theoretic Finite Dimensional Parametric Transversality Theorem (Appendix A), $\tau'_{f,\varphi,I,z}(y)$ is transversal to $L_I \times \Delta_I$ for a residual set of full Lebesgue measure of y in $\tilde{N}_{d-1}$. The same dimension counting argument as in section 5.3 shows that this implies that the image of $\tau'_{f,\varphi,I,z}(y)$ cannot intersect $L_I \times \Delta_I$, so that either $T_x\Phi_{f,\varphi,I,y,z}(v) \neq 0$ or $\hat{\rho}(f)(x, y) \notin \Delta_I$. But for every $(x, y) \in M \times \tilde{N}_{d-1}$ we have $\hat{\rho}(f)(x, y) \in \Delta_I$ for some I, and thus $T_x\Phi_{f,\varphi,I,y,z}(v) \neq 0$ for some I. For this choice of I we must thus have $T_x\Phi_{f,\varphi,I,y}(v) \notin L_I$. Thus $T_x\Phi_{f,\varphi,y,z}(v) \neq 0$ for all $v \in \tilde{T}M$, so that $\Phi_{f,\varphi,y,z}$ is immersive.

To prove injectivity, we define $\Phi_{f,\varphi,I,R,z}$: $M \times N^{d-1} \to \mathbb{R}^\gamma$ as in section 5.4 by

$$\begin{aligned}\Phi_{f,\varphi,I,R,z}(x, y) &= (\varphi(f^{(j'_1)}(x, y), z_{j'_1}), \varphi(f^{(j'_2)}(x, y), z_{j'_2}), \ldots, \varphi(f^{(j'_\gamma)}(x, y), z_{j'_\alpha}))^\dagger \\ &= (\varphi_{j'_1}(x_{j'_1}), \varphi_{j'_2}(x_{j'_2}), \ldots, \varphi_{j'_\gamma}(x_{j'_\gamma}))^\dagger,\end{aligned}$$

and let $\rho_{I,R,z}$: $\mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M \times N', \mathbb{R}) \to \mathcal{C}^r((M \times M)\backslash\Delta \times \tilde{N}_{d-1}, \mathbb{R}^\gamma \times \mathbb{R}^\gamma)$ by

$$\rho_{I,R,z}(f, \varphi)(x, x', y) = (\Phi_{f,\varphi,I,R,z}(x, y), \Phi_{f,\varphi,I,R,z}(x', y)),$$

and $\rho'_{I,R,z}$: $\mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M \times N', \mathbb{R}) \to \mathcal{C}^r((M \times M)\backslash\Delta \times \tilde{N}_{d-1}, \mathbb{R}^\gamma \times \mathbb{R}^\gamma \times M^d \times M^d)$ with

$$\rho'_{I,R,z}(f, \varphi) = (\rho_{I,R,z}(f, \varphi), \hat{\rho}'(f)).$$

We thus need to show that for every $(x, x') \in (M \times M)\backslash\Delta$, there is a pair I, R such that $\rho_{I,R,z}(f,\varphi)(x,x',y) \notin \Delta_\gamma$. To do this, we first prove that $ev_{\rho'_{I,R,z}}$ is transversal to $\Delta_\gamma \times \Delta_{I,R}$, as in Proposition 5.5. Let $\Xi = (f, \varphi, x, x', y)$ and suppose that $ev_{\rho'_{I,R,z}}(\Xi) \in \Delta_\gamma \times \Delta_{I,R}$. The second component $ev_{\hat{\rho}'}$ of $ev_{\rho'_{I,R,z}}$ is identical to that of $ev_{\rho'_{I,R}}$ and hence $T_{(f,x,x',y)}ev_{\hat{\rho}'}$ is surjective, and in particular, given any $(u_0, \ldots, u_{d-1}) \in T_{x_0}M \times \ldots \times T_{x_{d-1}}M$ and $(u'_0, \ldots, u'_{d-1}) \in T_{x'_0}M \times \ldots \times T_{x'_{d-1}}M$, we can find a $\eta \in T_f\mathcal{C}^{2r}(M \times N, M)$ such that $T_{(f,x,x',y)}ev_{\hat{\rho}'}(\eta, u_0, u'_0, 0_y) = ((u_0, \ldots, u_{d-1}), (u'_0, \ldots, u'_{d-1}))$. If $\gamma = 0$, this is sufficient to ensure the surjectivity of $ev_{\rho'_{I,R,z}}$. Otherwise, let $T_\Xi ev_{\rho_{I,R,z}}(\eta, 0_\xi, u_0, u'_0, 0_y) = (\tilde{u}, \tilde{u}')$. Since $\hat{\rho}'(f)(x, x', y) \in \Delta_{I,R}$, the points $x_{j_1}, x_{j_2}, \ldots, x_{j_\alpha}$ are all distinct and $\{x_{j_1}, x_{j_2}, \ldots, x_{j_\alpha}\} \cap \{x'_{j'_1}, x'_{j'_2}, \ldots, x'_{j'_\gamma}\} = \emptyset$. The first component of $T_\Xi ev_{\rho'_{I,R,z}}$ is given by $T_\Xi ev_{\rho_{I,R,z}}(0_\eta, \xi, 0_x, 0_{x'}, 0_y) = ((\xi_{j'_1}(x_{j'_1}), \xi_{j'_2}(x_{j'_2}), \ldots, \xi_{j'_\gamma}(x_{j'_\gamma})), (\xi_{j'_1}(x'_{j'_1}), \xi_{j'_2}(x'_{j'_2}), \ldots, \xi_{j'_\gamma}(x'_{j'_\gamma})))$, and thus, given any $(\bar{u}, \bar{u}') \in T_{\bar{\omega}}\mathbb{R}^\gamma \times T_{\bar{\omega}}\mathbb{R}^\gamma$, there exists a $\xi \in T_\varphi\mathcal{C}^{2r}(M \times N', \mathbb{R})$ such that $T_\Xi v_{\rho_{I,R,z}}(0_\eta, \xi, 0_x, 0_{x'}, 0_y) = (\bar{u} - \bar{u}', 0_{\bar{\omega}}) - (\tilde{u} - \tilde{u}', 0_{\bar{\omega}})$. Then $T_\Xi ev_{\rho_{I,R,z}}(\eta, \xi, u_0, u'_0, 0_y) = (\bar{u} - \bar{u}' + \tilde{u}', \tilde{u}')$. Since both $(\bar{u}', \bar{u}') \in T_{(\bar{\omega},\bar{\omega})}\Delta_\gamma$ and $(\tilde{u}', \tilde{u}') \in T_{(\bar{\omega})}\Delta_\gamma$, we have $(\bar{u}, \bar{u}') \in$ Image $T_\Xi ev_{\rho_{I,R,z}} + T_{(\bar{\omega},\bar{\omega})}\Delta_\gamma$. Hence $ev_{\rho'_{I,R,z}}$ is transversal to $\Delta_\gamma \times \Delta_{I,R}$.

By the Parametric Transversality Theorem there is a residual set of $(f, \varphi) \in \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M \times N', \mathbb{R})$ for which $\rho_{I,R,z}(f, \varphi)$ is transversal to $\Delta_\gamma \times \Delta_{I,R}$. Fix any (f, φ) in this set, and define $\rho'_{f,\varphi,I,R,z}\colon \tilde{N}_{d-1} \to \mathcal{C}^r((M \times M)\backslash\Delta, \mathbb{R}^\gamma \times \mathbb{R}^\gamma \times M^d \times M^d)$ by

$$\rho'_{f,\varphi,I,R,z}(y)(x, x') = \rho'_{I,R,z}(f, \varphi)(x, x', y).$$

Thus $ev_{\rho'_{f,\varphi,I,R,z}}(x, x', y) = \rho'_{I,R,z}(f, \varphi)(x, x', y)$ and so $ev_{\rho'_{f,\varphi,I,R,z}}$ is transversal to $\Delta_\gamma \times \Delta_{I,R}$. Thus by the Measure Theoretic Finite Dimensional Parametric Transversality Theorem (Appendix A), $\rho'_{f,\varphi,I,R,z}(y)$ is transversal to $\Delta_\gamma \times \Delta_{I,R}$ for a residual set of full Lebesgue measure of y in $\tilde{N}_{d-1}$. The same dimension counting argument as in section 5.4 shows that therefore the image of $\rho'_{f,\varphi,I,R,z}(y)$ cannot intersect $\Delta_\gamma \times \Delta_{I,R}$. Thus for every $(x, x') \in (M \times M)\backslash\Delta$ we have $\Phi_{f,\varphi,I,R,z}(x, y) \neq \Phi_{f,\varphi,I,R,z}(x', y)$ for some I, R and thus $\Phi_{f,\varphi,y,z}(x) \neq \Phi_{f,\varphi,y,z}(x')$. Hence $\Phi_{f,\varphi,y,z}$ is injective.

7.4. *Proof for* $\dim N = 0$, $\dim N' = 0$

This is similar in spirit to the previous sections, except that we follow sections 6.3 and 6.4 instead of sections 5.3 and 5.4. We thus start by recalling the definition of $\tilde{\mathcal{P}}^{(2d-1)}$ for a given $f \in \tilde{\mathcal{B}}^{(2d)}$. Then, as in Proposition 6.24, we construct a dense open set of $\varphi \in \mathcal{C}^{2r}(M \times N', \mathbb{R})$ for which $\varphi(x, z_0) \neq \varphi(x', z_0)$ for all $x, x' \in \tilde{\mathcal{P}}^{(2d-1)}$ such that $x \neq x'$ and all $z_0 \in N'$. This is possible since $\tilde{\mathcal{P}}^{(2d-1)}$ and N' consist of a finite number of points. Next, as in Proposition 6.25, we construct an open and dense set of $\varphi \in \mathcal{C}^{2r}(M \times N', \mathbb{R})$ for which if $x \in \tilde{P}_{y_0 \ldots y_{d-2}}$ for some $y \in N^{d-1}$, then $T_x\Phi_{f,\varphi,y,z}$ has rank m for all $z \in (N')^d$. We let $\mathcal{A}_f$ be the intersection of the resulting sets of φ and define $\mathcal{E}^{2r}$ as before to be the set of (f, φ) such that $f \in \tilde{\mathcal{B}}^{(2d)}$ and $\varphi \in \mathcal{A}_f$. This is open dense in $\mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M \times N', \mathbb{R})$. Now, fix both $y \in N^{d-1}$ and $z \in (N')^d$, and by analogy to (6.7), define $\tau_{y,z}\colon \mathcal{E}^{2r} \to \mathcal{C}^{r-1}(\tilde{T}M, T\mathbb{R}^d)$ by

$$\tau_{y,z}(f, \varphi) = T\Phi_{f,\varphi,y,z}.$$

Then $ev_{\tau_{y,z}}$ is $\mathcal{C}^{r-1}$ and $T_{(f,\varphi,v)}ev_{\tau_{y,z}}(0_f, \xi, 0_x) = (\bar{\omega}(T_{x_0}\xi_{z_0}(v_0)), \bar{\omega}(T_{x_1}\xi_{z_1}(v_1)), \ldots, \bar{\omega}(T_{x_{d-1}}\xi_{z_{d-1}}(v_{d-1})))^\dagger$. If $ev_{\tau_{y,z}}(f, \varphi, v) \in L$, then rank $T_x\Phi_{f,\varphi,y,z} < m$, and since $\varphi \in \mathcal{A}_f$, we cannot have $x \in \tilde{\mathcal{P}}_{y_0\ldots y_{d-2}}$. Thus the points $\{x_0, \ldots, x_{d-1}\}$ are distinct and so $T_{(f,\varphi,v)}ev_{\tau_{y,z}}$ is surjective, and hence $ev_{\tau_{y,z}}$ is transversal to L. Thus by the Parametric Transversality Theorem, $\tau_{y,z}(f, \varphi)$ is transversal to L for an open dense set of $(f, \varphi) \in \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M \times N', \mathbb{R})$. For any such (f, φ) the same dimension calculation as in section 6.3 shows that transversality implies nonintersection. Hence $\tau_{y,z}(f, \varphi)(v) = T_x\Phi_{f,\varphi,y,z}(v) \notin L$ for all $v \in \tilde{T}M$, so that $\Phi_{f,\varphi,y,z}$ is immersive.

To prove injectivity, we define $\rho_{y,z}\colon \mathcal{E}^{2r} \to \mathcal{C}^r((M \times M)\backslash\Delta, \mathbb{R}^d \times \mathbb{R}^d)$ by

$$\rho_{y,z}(f, \varphi)(x, x') = (\Phi_{f,\varphi,y,z}(x), \Phi_{f,\varphi,y,z}(x')).$$

If $ev_{\rho_{y,z}}(f, \varphi, x, x') \in \Delta_d$, we have $\Phi_{f,\varphi,y,z}(x) = \Phi_{f,\varphi,y,z}(x')$ and hence in particular $\varphi_{z_0}(x) = \varphi_{z_0}(x')$. Since $\varphi \in \mathcal{A}_f$, at least one of x or x' cannot be in $\tilde{\mathcal{P}}^{(2d-1)}$, and as before we assume without loss of generality that this is x. Then $ev_{\rho_{y,z}}$ is $\mathcal{C}^r$ and we show that $ev_{\rho_{y,z}}$ is transversal to Δ_d exactly as in the proof of Proposition 6.30. Thus by the Parametric Transversality Theorem, $\rho_{y,z}(f, \varphi)$ is transversal to Δ_d for a residual set of (f, φ) in $\mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M \times N', \mathbb{R})$. For any such (f, φ) the same dimension calculation as in section 6.4 shows that transversality implies nonintersection. Hence for a residual set of (f, φ) we have $\Phi_{f,\varphi,y,z}(x) \neq \Phi_{f,\varphi,y,z}(x')$ for all $x \neq x'$, i.e. $\Phi_{f,\varphi,y,z}$ is injective.

Taking the intersection over all $y \in N^{d-1}$ and $z \in (N')^d$, we obtain a residual set of (f, φ) for which $\Phi_{f,\varphi,y,z}$ is an injective immersion and hence an embedding, since M is compact. Finally, embeddings are open in $\mathcal{C}^r(M, \mathbb{R}^d)$ and $\Phi_{f,\varphi,y,z}$ depends continuously on f and φ, and hence the set of (f, φ) for which $\Phi_{f,\varphi,y,z}$ is an embedding is open and dense.

8. Proof of Theorem 2.6

We now turn to the case where both f and φ depend on the same process, so that after reduction to finite dimensions we want to show that

$$\Phi'_{f,\varphi,y}(x) = (\varphi(f^{(0)}(x, y), y_0), \varphi(f^{(1)}(x, y), y_1), \ldots, \varphi(f^{(d-1)}(x, y), y_{d-1}))^\dagger$$

is an embedding. As with both Theorem 2.3 and Theorem 2.5 we have a different approach depending on the dimension of N. The case of $\dim N > 0$ is very similar to that in section 7.2 above, whilst that for $\dim N = 0$ parallels section 7.4. We briefly sketch details in the next two sections respectively.

8.1. $\dim N > 0$

When $\dim N > 0$, we can use a very similar approach to that in section 7.2 above to obtain a slightly stronger version of Theorem 2.6:

Theorem 8.1. *Let M and N be compact manifolds, with* $m = \dim M > 0$ *and* $\dim N > 0$. *Suppose that* $d \geq 2m + 1$ *and* $r \geq 1$. *Given any* $f \in \mathcal{D}^r(M \times N, M)$*, there exists* *a residual set of $\varphi \in \mathcal{C}^r(M \times N, \mathbb{R})$ such that for any φ in this set there is an open dense set $\Sigma_{f,\varphi} \subset \Sigma$ such that $\tilde{\Phi}_{f,\varphi,\omega}$ is an embedding for all $\omega \in \Sigma_{f,\varphi}$. If μ_Σ is an invariant probability measure on $\Sigma = N^{\mathbb{Z}}$ with μ_{d-1} absolutely continuous with respect to Lebesgue measure on N^{d-1}, then we can choose $\Sigma_{f,\varphi}$ such that $\mu_\Sigma(\Sigma_{f,\varphi}) = 1$.*

Using the usual reduction to N^{d-1}, as in section 3.1, this follows from

Theorem 8.2. *Let M and N be compact manifolds, with $m = \dim M > 0$ and $\dim N > 0$. Suppose that $d \geq 2m + 1$ and $r \geq 1$, and let ν_{d-1} be normalized Lebesgue measure on N^{d-1}. Then given any $f \in \mathcal{D}^r(M \times N, M)$, there exists a residual set of $\varphi \in \mathcal{C}^r(M \times N, \mathbb{R})$ such that for any φ in this set there is an open dense set $U_{f,\varphi} \subset N^{d-1}$ such that $\nu_{d-1}(U_{f,\varphi}) = 1$ and the delay map $\tilde{\Phi}_{f,\varphi,y}$ is an embedding for all $y \in U_{f,\varphi}$.*

Sketch Proof. This is identical to the proof of Theorem 7.2, apart from appropriate changes to relevant definitions. The dependence of φ on y causes us no difficulties since in establishing transversality we never need make perturbations in y. Thus fix $f \in \mathcal{D}^{2r}(M \times N, M)$ and define $\tau'_f\colon \mathcal{C}^{2r}(M \times N, \mathbb{R}) \to \mathcal{C}^{r-1}(\tilde{T}M \times \tilde{N}_d, T\mathbb{R}^d)$ by

$$\tau'_f(\varphi)(v, y) = T_x \Phi'_{f,\varphi,y}(v).$$

Then $ev_{\tau'_f}$ is $\mathcal{C}^{r-1}$ and $T_{(\varphi,v,y)} ev_{\tau'_f}(\xi, 0_v, 0_y) = (\bar{\omega}(T_{x_0}\xi(v_0, y_0)), \bar{\omega}(T_{x_1}\xi(v_1, y_1)), \ldots, \bar{\omega}(T_{x_{d-1}}\xi(v_{d-1}, y_{d-1})))^\dagger$. Since the points $\{y_0, \ldots, y_{d-1}\}$ are distinct, we can choose each $T_{x_i}\xi(v_i, y_i)$ independently, and hence $T_{(\varphi,x,y)} ev_{\tau'_f}$ is surjective. Thus $ev_{\tau'_f}$ is transversal to the zero section L in $T\mathbb{R}^d$ and by the Parametric Transversality Theorem, therefore, there is a residual set of $\varphi \in \mathcal{C}^{2r}(M \times N, \mathbb{R})$ such that $\tau'_f(\varphi)$ is transversal to L. Fix any such φ and define $\tau'_{f,\varphi}\colon \tilde{N}_d \to \mathcal{C}^{r-1}(\tilde{T}M, T\mathbb{R}^d)$ by $\tau'_{f,\varphi}(y) = T\Phi'_{f,\varphi,y}$. Then $ev_{\tau'_{f,\varphi}}$ is $\mathcal{C}^{r-1}$ and $ev_{\tau'_{f,\varphi}}(y, v) = \tau'_f(\varphi)(v, y)$, and hence $ev_{\tau'_{f,\varphi}}$ is transversal to L. Applying the Measure Theoretic Finite Dimensional Parametric Transversality Theorem (Appendix A), we obtain an open dense set of y of full μ_d measure for which $T\Phi'_{f,\varphi,y}$ is transversal to L. The same dimension calculation as in section 7.2 then shows that, as usual, transversality implies that $T\Phi'_{f,\varphi,y}$ does not intersect L and hence $\Phi'_{f,\varphi,y}$ is immersive.

To prove injectivity, we define $\rho'_f\colon \mathcal{C}^{2r}(M \times N, \mathbb{R}) \to \mathcal{C}^r((M \times M)\backslash\Delta \times \tilde{N}_d, \mathbb{R}^d \times \mathbb{R}^d)$ by

$$\rho'_f(\varphi)(x, x', y) = (\Phi'_{f,\varphi,y}(x), \Phi'_{f,\varphi,y}(x')).$$

Thus $ev_{\rho'_f}$ is $\mathcal{C}^r$ and $T_{(\varphi,x,x',y)} ev_{\rho'_f}(\xi, 0_x, 0_{x'}, 0_y) = ((\xi(x, y_0), \xi(x_1, y_1), \ldots, \xi(x_{d-1}, y_{d-1})), (\xi(x', y_0), \xi(x'_1, y_1), \ldots, \xi(x'_{d-1}, y_{d-1})))^\dagger$. Since for $y \in \tilde{N}_d$, the points $\{y_0, \ldots, y_{d-1}\}$ are distinct, and $x_i \neq x'_i$ for all i, the points $\{(x_0, y_0), (x_1, y_1), \ldots, (x_{d-1}, y_{d-1}), (x'_0, y_0), (x'_1, y_1), \ldots, (x'_{d-1}, y_{d-1})\}$ are all distinct. Therefore $T_{(\varphi,x,x',y)} ev_{\rho'_f}$ is surjective, so that $ev_{\rho'_f}$ is transversal to the diagonal Δ_d in $\mathbb{R}^d \times \mathbb{R}^d$. By the Parametric Transversality Theorem, there is a residual set of $\varphi \in \mathcal{C}^{2r}(M \times N, \mathbb{R})$ such that $\rho'_f(\varphi)$ is transversal to Δ_d. Fix any such φ and define $\rho'_{f,\varphi}\colon \tilde{N}_d \to \mathcal{C}^r((M \times M)\backslash\Delta, \mathbb{R}^d)$ by $\rho'_{f,\varphi}(y)(x, x') = \rho'_f(\varphi)(x, x', y)$. Then $ev_{\rho'_{f,\varphi}}$ is $\mathcal{C}^r$ and $ev_{\rho'_{f,\varphi}}(y, x, x') = \rho'_f(\varphi)(x, x', y)$ and hence $ev_{\rho'_{f,\varphi}}$ is transversal to Δ_d. Applying the Measure Theoretic Finite Dimensional Parametric Transversality Theorem (Appendix A), we obtain an open dense set of

y of full μ_d measure for which $\Phi'_{f,\varphi,y} \times \Phi'_{f,\varphi,y}$ is transversal to Δ_d. The same dimension calculation as in section 7.2 shows that transversality implies that Image $\Phi'_{f,\varphi,y} \times \Phi'_{f,\varphi,y}$ does not intersect Δ_d and hence $\Phi_{f,\varphi,y,z}(x) \neq \Phi_{f,\varphi,y,z}(x')$ for all $x \neq x'$, as required. □

8.2. **dim $N = 0$**

In this case, as with Theorem 2.4 and Theorem 7.5 we can strengthen Theorem 2.6 to

Theorem 8.3. *Let M and N be compact manifolds, with $m = \dim M > 0$ and dim $N = 0$. Suppose that $d \geq 2m + 1$ and $r \geq 1$. Then there exists an open dense set of $(f, \varphi) \in \mathcal{D}^r(M \times N, M)\mathcal{C}^r(M \times N, \mathbb{R})$ such that for any (f, φ) in this set $\Phi'_{f,\varphi,\omega}$ is an embedding for all $\omega \in \Sigma$.*

This is follows from the finite dimensional version:

Theorem 8.4. *Let M and N be compact manifolds, with $m = \dim M > 0$ and dim $N = 0$. Suppose that $d \geq 2m + 1$ and $r \geq 1$. Then there exists an open dense set of $(f, \varphi) \in \mathcal{D}^r(M \times N, M) \times \mathcal{C}^r(M \times N, \mathbb{R})$ such that for any (f, φ) in this set $\Phi'_{f,\varphi,y}$ is an embedding for all $y \in N^d$.*

Sketch Proof. This closely follows the proof of Theorem 7.6, which in turn was based on sections 6.3 and 6.4. As in Proposition 6.24, we construct a dense open set of $\varphi \in \mathcal{C}^{2r}(M \times N, \mathbb{R})$ for which $\varphi(x, y_0) \neq \varphi(x', y_0)$ for all $x, x' \in \tilde{\mathcal{P}}^{(2d-1)}$ such that $x \neq x'$ and all $y_0 \in N$. Then, as in Proposition 6.25, we construct an open and dense set of $\varphi \in \mathcal{C}^{2r}(M \times N, \mathbb{R})$ for which if $x \in \tilde{\mathcal{P}}_{y_0 \ldots y_{d-2}}$ for some $y \in N^d$, then $T_x\Phi'_{f,\varphi,y}$ has rank m for all $y \in N^d$. We let $\mathcal{A}_f$ be the intersection of the resulting sets of φ and define $\mathcal{E}^{2r}$ as before to be the set of (f, φ) such that $f \in \tilde{\mathcal{B}}^{(2d)}$ and $\varphi \in \mathcal{A}_f$. This is open dense in $\mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M \times N, \mathbb{R})$. Now fix $y \in N^d$ and define $\tau_y\colon \mathcal{E}^{2r} \to \mathcal{C}^{r-1}(\tilde{T}M, T\mathbb{R}^d)$ by

$$\tau'_y(f, \varphi) = T_x\Phi'_{f,\varphi,y}.$$

Then $ev_{\tau'_y}$ is $\mathcal{C}^{r-1}$ and $T_{(f,\varphi,v)}ev_{\tau'_y}(0_f, \xi, 0_x) = (\bar{\omega}(T_{x_0}\xi_{y_0}(v_0)), \bar{\omega}(T_{x_1}\xi_{y_1}(v_1)), \ldots, \bar{\omega}(T_{x_{d-1}}\xi_{y_{d-1}}(v_{d-1})))^\dagger$. If $ev_{\tau'_y}(f, \varphi, v) \in L$, then rank $T_x\Phi'_{f,\varphi,y} < m$, and since $\varphi \in \mathcal{A}_f$, we cannot have $x \in \tilde{\mathcal{P}}_{y_0 \ldots y_{d-2}}$. Thus the points $\{x_0, \ldots, x_{d-1}\}$ are distinct and so $T_{(f,\varphi,v)}ev_{\tau'_y}$ is surjective, and hence $ev_{\tau'_y}$ is transversal to L. Thus by the Parametric Transversality Theorem, $\tau'_y(f, \varphi)$ is transversal to L for an open dense set of $(f, \varphi) \in \mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M \times N, \mathbb{R})$. For any such (f, φ) the same dimension calculation as in section 6.3 shows that transversality implies nonintersection. Therefore $T_x\Phi'_{f,\varphi,y}(v) \notin L$ for all $v \in \tilde{T}M$, so that $\Phi'_{f,\varphi,y}$ is immersive.

To prove injectivity, we define $\rho'_y\colon \mathcal{E}^{2r} \to \mathcal{C}^r((M \times M)\backslash\Delta, \mathbb{R}^d \times \mathbb{R}^d)$ by

$$\rho'_y(f, \varphi)(x, x') = (\Phi'_{f,\varphi,y}(x), \Phi'_{f,\varphi,y}(x')).$$

If $ev_{\rho'_y}(f, \varphi, x, x') \in \Delta_d$, we have $\Phi'_{f,\varphi,y}(x) = \Phi'_{f,\varphi,y}(x')$ and in particular $\varphi_{y_0}(x) = \varphi_{y_0}(x')$. Since $\varphi \in \mathcal{A}_f$, at least one of x or x' cannot be in $\tilde{\mathcal{P}}^{(2d-1)}$ and as usual we assume without loss of generality that this is x. We show that $ev_{\rho'_y}$ is transversal to Δ_d exactly as in the proof of Proposition 6.30. Thus by the Parametric Transversality Theorem, $\rho'_y(f, \varphi)$ is transversal to Δ_d for a residual set of (f, φ) in $\mathcal{D}^{2r}(M \times N, M) \times \mathcal{C}^{2r}(M \times N, \mathbb{R})$. For any such (f, φ) the same dimension calculation as in section 6.4 shows that transversality implies nonintersection and hence that $\Phi'_{f,\varphi,y}(x) \neq \Phi'_{f,\varphi,y}(x')$ for all $x \neq x'$. Hence $\Phi'_{f,\varphi,y}$ is injective for a residual set of (f, φ).

Taking the intersection over all $y \in N^d$, we obtain a residual set of (f, φ) for which $\Phi'_{f,\varphi,y}$ is an injective immersion, and hence an embedding since M is compact. Finally, embeddings are open in $\mathcal{C}^r(M, \mathbb{R}^d)$ and $\Phi'_{f,\varphi,y}$ depends continuously on f and φ, and hence the set of (f, φ) for which $\Phi'_{f,\varphi,y}$ is an embedding is open and dense. $\square$

Appendix A

Measure Theoretic Finite Dimensional Parametric Transversality Theorem. Let $\mathcal{A}$, M, and N be $\mathcal{C}^r$ manifolds and $\rho\colon \mathcal{A} \to \mathcal{C}^r(M, N)$ be a map such that ev_ρ is $\mathcal{C}^r$. Let $L \subset N$ be a submanifold of finite codimension p in N. Suppose that $\mathcal{A}$ and M are finite dimensional, that $r > \max\{0, m - p\}$ where m is the dimension of M, and that ev_ρ is transversal to L. Then the set of a such that $\rho(a)$ is not transversal to L has zero Lebesgue measure in $\mathcal{A}$. Furthermore if L is closed and M is compact, then the set of such a is open.

Sketch Proof. This is identical to the proof of the standard Parametric Transversality Theorem, except that instead of Smale's Density Theorem, we use Sard's Theorem. Recall that this states that the set of regular values of a $\mathcal{C}^r$ map $f\colon M \to N$ has full Lebesgue measure in N if $r \geq \max\{0, m - n\}$ where $m = \dim M$ and $n = \dim N$. A *regular value* is a point $y \in N$ such that $T_x f$ is surjective at any $x \in f^{-1}(y)$ (so that any y such that $f^{-1}(y) = \emptyset$ is automatically regular).

So, suppose that ev_ρ is transversal to L, then $L' = (ev_\rho)^{-1}(L)$ is a $\mathcal{C}^r$ submanifold of $\mathcal{A} \times M$ of codimension p. Since ev_ρ is transversal to L, given any $(a, x) \in L'$ and any $u \in T_y N$, where $y = ev_\rho(a, x) = \rho(a)(x)$, there exists a $\alpha \in T_a\mathcal{A}$, $v \in T_x M$ and $w \in T_y L$ such that $u = T_{(a,x)} ev_\rho(\alpha, v) + w = T_a\rho(\alpha)(x) + T_x\rho(a)(v) + w$.

Let $\pi_{L'}\colon L' \to \mathcal{A}$ be the restriction to L' of the projection $\pi_{\mathcal{A}}\colon \mathcal{A} \times M \to \mathcal{A}$ onto the first factor and suppose that a is a regular value of $\pi_{L'}$. If $(\pi_{L'})^{-1}(a) = \emptyset$, then $\rho(a)(x) \notin L$ for all $x \in M$ and so $\rho(a)$ is trivially transversal to L. On the other hand, if for some $x \in M$ we have $\rho(a)(x) \in L$ then given any $u \in T_y N$, by the above there exists a α, v, and w such that $T_a\rho(\alpha)(x) + T_x\rho(a)(v) + w = u$. Since $T_{(a,x)}\pi_{L'}$ is surjective, there exists a $v' \in T_x M$ such that $(\alpha, v') \in T_{(a,x)} L'$, which implies that $(T_a\rho(\alpha)(x) + T_x\rho(a)(v')) \in T_y L$. Thus by linearity $T_x\rho(a)(v - v') + (T_a\rho(\alpha)(x) + T_x\rho(a)(v') + w) = u$, which implies that $\rho(a)$ is transversal to L. Hence if a is a regular value of $\pi_{L'}$, then $\rho(a)$ is transversal to L (the converse is also true) and so the set of a such that $\rho(a)$ is transversal to L has full Lebesgue measure. $\square$

Acknowledgements

JS would like to thank helpful conversations with Ludwig Arnold over many years. Without his encouragement and generous advice, this paper would never have been written. Some of the preliminary work was carried out at IMPA, Brazil, during the International Conference on Dynamical Systems held there in August 1993, and JS is extremely grateful to IMPA for their hospitality and financial support, and to the Royal Society for proving travel funds to attend this conference. The authors would also like to thank the Basic Research Institute in Mathematical Sciences at Hewlett-Packard in Bristol for its hospitality in the summer of 1994 and the Isaac Newton Institute Part for its hospitality and financial support during the programme "From Finite to Infinite Dimensional Dynamical Systems" in the summer of 1995. Finally, we would like to gratefully acknowledge financial support from the UK Engineering and Physical Sciences Research Council (GR/J97243 and GR/L42513), the Nuffield Foundation (SCI/180/92/365/G), the Royal Commission for the Exhibition of 1851, and the Royal Society (Leverhume Trust Senior Research Fellowship awarded to JS).

References

[1] Abraham R., 1963, Transversality in Manifolds of Mappings, *Bull. Amer. Math. Soc.*, **69**, 470–474.
[2] Abraham R. and Robbin J., 1967, *Transversal Mappings and Flows*, W.D. Benjamin, New York.
[3] Aeyels D., 1981, Generic Observability of Differentiable Systems, *SIAM J. Control and Optimization*, **19**, 595–603.
[4] Arnold L., 1998, *Random Dynamical Systems*, Springer-Verlag, Berlin.
[5] Barnsley M., 1988, *Fractals Everywhere*, Academic Press, Boston.
[6] Broomhead D.S., Huke J.P. and Muldoon M.R., 1999, Oversampling Nonlinear Digital Channels, in preparation.
[7] Casdagli M., 1992, A Dynamical Systems Approach to Modelling Input-Output Systems. In M. Casdagli and S. Eubank, eds., *Nonlinear Modelling and Forecasting*, Addison-Wesley, Reading, MA.
[8] Chen S. and Billings S.A., 1989, Modelling and Analysis of Nonlinear Time Series, *Int. J. Control*, **50**, 2151–2171.
[9] Eells J., 1966, A Setting for Global Analysis, *Bull. Amer. Math. Soc.*, **72**, 751–807.
[10] Foster M.J., 1975, Calculus on Vector Bundles, *J. London Math. Soc.*, **11**, 65–73.
[11] Franks J., 1979, Manifolds of C^r Mappings and Applications to Dynamical Systems, *Adv. in Math. Supplementary Studies*, **4**, 271–290.
[12] Hirsch M., 1976, *Differential Topology*, Springer-Verlag, New York.
[13] Huke J.P., 1993, Embedding Nonlinear Dynamical Systems, A Guide to Takens Theorem, Internal Report, DRA, Malvern, UK.
[14] Kifer Y., 1988, *Random Perturbations of Dynamical Systems*, Birkhäuser, Basel.
[15] Leontaritis I.J. and Billings S.A., 1985, Input-Output Parametric Models for Nonlinear Systems, Part II: Stochastic Nonlinear Systems, *Int. J. Control*, **41**, 329–344.
[16] Martin R., 1998, *Irregularly Sampled Signals: Theories and Techniques for Analysis*, Ph.D. Thesis, UCL.

[17] Norman M.F., 1968, Some Convergence Theorems for Stochastic Learning Operators, *J. Math. Psychology*, **5**, 61–101.
[18] Ott E., Sauer T., and Yorke J.A., 1994, *Coping with Chaos*, Wiley, New York.
[19] Sauer T., Yorke J.A., and Casdagli M., 1991, Embedology, *J. Stat. Phys.*, **65**, 579–616.
[20] Sontag E.D., 1979a, *Polynomial Response Maps*, Springer-Verlag, Berlin.
[21] Sontag E.D., 1979b, Realization Theory of Discrete-Time Nonlinear Systems, Part I: The Bounded Case, *IEEE Trans. Circuits and Systems*, **26**, 342–356.
[22] Sontag E.D., 2002, For Differential Equations with r Parameters, $2r + 1$ Experiments Are Enough for Identification, *J. Nonlinear Sci.*, **12**, 553–583.
[23] Stark J., 1999, Delay Embeddings for Forced Systems, Part I: Deterministic Forcing, *J. Nonlinear Sci.*, **9**, 255–332.
[24] Stark J., 2001, Delay Reconstruction: Dynamics v. Statistics. In A.I. Mees, ed., *Nonlinear Dynamics and Statistics*, Birkhäuser, Basel.
[25] Stark J., Broomhead D.S., Davies M.E., and Huke J., 1997, Takens Embedding Theorems for Forced and Stochastic Systems, *Nonlinear Analysis—Theory Methods & Applications*, **30**, 5303–5314.
[26] Takens F., 1980, Detecting Strange Attractors in Fluid Turbulence. In D.A. Rand and L.-S. Young, eds., *Dynamical Systems and Turbulence, Warwick 1980*, Springer-Verlag, Berlin.
[27] Wang Y. and Sontag E.D., 1995, Orders of Input/Output Differential Equations and State Space Dimensions, *SIAM J. Control Optim.*, **33**, 1102–1126.