A Short Proof of Takens' Theorem

The goal of this note is to present a short version the proof of Takens' theorem [1]. This is heavily based on Noakes' paper [2].

First, we state the main result.

Theorem 1: Takens' theorem

Let M be a compact manifold of dimension m.
Let $\phi \in Diff^1(M)$ and $y \in C^1(M, \mathbb{R})$.
Define the map $f : M \to \mathbb{R}^{2m+1}$ by $f(x) = (y(x), y(\phi(x)), \ldots, y(\phi^{2m}(x)))$.
Then, for pairs (ϕ, y), it is a generic property that f is an embedding.

We will say that a property is *generic* if it holds for a set which is *open and dense* in the C^1 topology.
The dual concept of an open dense set is a *nowhere dense* set (also called a *rare* set). A set is nowhere dense if and only if its closure has empty interior. Equivalently, a set is nowhere dense if and only if it is a subset of the boundary of an open set, or if its closure is the boundary of an open set.

A *meagre* set is a countable union of nowhere dense sets. A *residual* set (also called a *comeagre* set) is the complement of a meagre set. A residual set is a countable intersection of open dense sets. Since $C^1(M, N)$ is a Baire space (a space where countable unions of closed sets with empty interiors also have empty interiors), every residual set of $C^1(M, N)$ is dense. The notion of generic property can be extended to be any property that hold for a set containing a residual set. A generic property is 'typical' in the sense that if a point does not satisfy the property, any small perturbation of the point will land on the residual set. Hence, for example an algorithm can be designed to work only on generic points. Since a open dense set is a residual set, it is sufficient to prove that the property holds on an open dense set to be able to say that it is a generic property. This is the approach followed here.

1 Summary of the proof

In this section, we provide a sketch of the proof in order to expose the intuitions behind Taken's theorem while avoiding technicalities. A more rigorous treatment is provided in the next sections.

Let $\phi \in Diff^1(M)$ and $y \in C^1(M, \mathbb{R})$.

1.1 Step 1: Perturbation of ϕ to obtain a finite number of periodic point with period less or equal to $2m$

We deal first with fixed points of ϕ. Let x be such a fixed point.

- Consider the *graph* of ϕ denoted $\gamma : M \to M \times M$.
- Denote by Δ the diagonal of $M \times M$. Note that Δ has dimension m.
- Hence, $T_{(x,x)}\Delta$ is a vector subspace of $T_{(x,x)}(M \times M)$ with dimension m.
- Similarly, $d\gamma_x(T_x M)$ is a vector subspace of $T_{(x,x)}(M \times M)$ with dimension m.
- A small perturbation of $d\gamma_x$ gives us $d\gamma_x(T_x M) \cap T_{(x,x)}\Delta = \{0\}$.

- The *transversality principle* tells us that a perturbation of ϕ (hence of γ) gives us the desired perturbation of $d\gamma_x$ (see the Thom's transversality theorem).
- The *linearity principle* tells us that there exists a neighbourhood U_x of x in which x is the only fixed point of ϕ. Let P_1 be the set of fixed points of ϕ. We have just shown that P_1 is a discrete set, hence at most countable.
- Now, denote by X the union of the open sets U_x for every $x \in P_1$. For every $p \in (M \setminus X)$, there exists a neighbourhood V_p such that V_p does not contain any fixed point. The collection $\{V_p, p \in (M \setminus X)\}$ together with $\{U_x, x \in P_1\}$ form an open cover of M, with the property that each open set of the cover contains at most one fixed point. Since M is compact, we can extract a finite subcover $\{W_i\}$, where each W_i contains at most one fixed point. Hence, the fixed points of our perturbed ϕ are finite in number.

Now, let x be a periodic point of ϕ with period $p \leq 2m$.

- Then, x is a fixed point of ϕ^p.
- By the same argument, we can perturb ϕ^p (by perturbing ϕ) such that x becomes an isolated point.
- Then, by the compactness argument, we can show that ϕ^p admits a finite number of fixed points. Hence, ϕ admits a finite number of periodic points with period $p \leq 2m$.

We repeat this argument for every $p \leq 2m$. Hence, after a finite number of perturbations of ϕ, the set P_{2m} of periodic points with period less or equal to $2m$ is finite.

1.2 Step 2: Perturbation of ϕ and y such that f be an immersion on P_{2m}

We only deal with the fixed points of ϕ. By the same argument as in the previous section, we can extend our results on any periodic point with period less or equal to $2m$. Let x be a fixed point of ϕ.

- Consider $d\phi_x : T_xM \to T_xM$ and $dy_x : T_xM \to \mathbb{R}$. This is a linear time-invariant control system: $u_{t+1} = d\phi_x(u_t)$ and $v_t = dy_x(u_t)$.
- A small perturbation of $d\phi_x$ and dy_x makes the system observable, which means that the map $F = (F_1, \ldots, F_m) : T_xM \to \mathbb{R}^m$ with $F_i = dy_x \circ d\phi_x^{i-1}$ is full rank (i.e., is an isomorphism).
- By the transversality principle, we can perturb ϕ and y to obtain the desired $d\phi_x$ and dy_x i.e., $df_x : T_xM \to \mathbb{R}^{2m+1}$ is full rank ($= m$).

After a finite number of such perturbations, f is immersive on P_1. Then, we show that f is immersive on P_{2m}.

1.3 Step 3: Obtain an embedding on a neighbourhood of P_{2m}

- By the linearity principle, since f is an immersion on P_{2m}, it is embedding on a neighbourhood of any point in P_{2m} (taken one by one).
- With a small perturbation of f, we obtain an embedding on a neighbourhood W of P_{2m}. By taking a smaller neighbourhood, we can choose W to be closed so that $M \setminus W$ is open. Hence $M \setminus W$ is a submanifold of M on which we can apply Whitney's embedding theorem.

1.4 Step 4: Obtain an embedding on $M \setminus W$

Let $x \in (M \setminus W)$. Then, $x, \phi(x), \ldots, \phi^{2m}(x)$ are distinct.

- We have $f_i = y \circ \phi^{i-1}$ so we can perturb y on disjoint neighbourhoods of the points $x, \phi(x), \ldots, \phi^{2m}(x)$ to independently modify each $f_i(x)$. This is a necessary condition for applying the Whitney's embedding theorem. Otherwise we would have at least two equal components in f so that the codomain of f would have a dimension strictly less than $2m + 1$.

- By Whitney's embedding theorem, we can perturb f to obtain an embedding on $M \setminus W$.

Finally, by perturbing ϕ and y, we have obtain an embedding f on M.

One of the key element lacking from this summary is to explain why it is possible to perturb a map to obtain a given property without destroying the property obtained with previous perturbations. This is addressed in the next sections.

2 Some useful results

Let $Imm(M, \mathbb{R}^{2m+1})$ be the space of immersions of M in $\mathbb{R}^{2m+1}$. Let $Emb(M, \mathbb{R}^{2m+1})$ be the space of embeddings. It is the subset of $Imm(M, \mathbb{R}^{2m+1})$ of injective immersions. We do not require an embedding to be an homeomorphism since by the compactness of M, this is always the case.

Lemma 1: Openness of immersions

$Imm(M, \mathbb{R}^{2m+1})$ is open in $C^1(M, \mathbb{R}^{2m+1})$ equipped with the Whitney C^1 topology.

Lemma 2: Whitney embedding theorem

$Emb(M, \mathbb{R}^{2m+1})$ is dense in $C^1(M, \mathbb{R}^{2m+1})$.

Definiton 1: Good embedding

Let B be a subset of M.
We call $g \in C^1(M, \mathbb{R}^{2m+1})$ a *good embedding* of B when there is an open set $W \subset M$ such that $\mathrm{cl}B \subset W$ and $g|_W$ is an embedding of W in $\mathbb{R}^{2m+1}$.

This definition is introduced to address the difficulty of destroying a previous property by perturbing a map to obtain another property. Suppose f is an embedding over B. We want to perturb f in order to obtain an embedding over an open set V. To avoid destroying our previous progress, we perturb f in order to obtain an embedding over an open set U such that $\mathrm{cl}(B) \subset U$ and $\mathrm{cl}(V) \subset U$. We call V a *shrunken neighbourhood* and U a *large neighbourhood*.

Lemma 3: Openness of good embeddings

The good embeddings of B form an open subset of $C^1(M, \mathbb{R}^{2m+1})$.
In particular, this is true for $B = M$.

Lemma 4: Genericity of finitely many periodic points

Let $q \geq 1$.
There exists an open dense subset $\Phi_q \subset C^1(M, M)$ such that any $\phi \in \Phi_q$ has finitely many periodic points of periods $p \leq q$.
In particular, this is true for $q = 2m$.

Proof: Genericity of finitely many periodic points

It suffices to prove this in the case $q = 1$.
Indeed, a periodic point of ϕ of period p is a fixed point of ϕ^p, and intersections of open dense sets are

open and dense.
Then apply the Thom's transversality theorem (or the multijet transversality theorem?). □

Let $\phi \in \Phi_{2m}$ and P_{2m} its finite set of periodic points of period $p \leq 2m$.

Lemma 5: Genericity of submersions on periodic points

There exists an open dense subset $\mathcal{E} \in C^1(M, \mathbb{R})$ such that every $y \in \mathcal{E}$ is submersive on P_{2m} (i.e., y has no critical point in P_{2m}).

Proof: Genericity of submersions on periodic points

Since $\mathbb{R}$ has dimension 1, then x is a critical point of y if and only if $dy_x = \mathbf{0}$.
Let $x \in P_{2m}$. If $dy_x = \mathbf{0}$, perturb y by adding a small C^1 (or even C^∞) function whose support is in a neighbourhood of x that does not contain other points of P_{2m}, and whose derivative at x is nontrivial (i.e., non identically zero). Such a neighbourhood of x exists since P_{2m} is finite. □

Let $L : V \to V$ be a endomorphism of an m-dimensional vector space V.
Let $P : V \to \mathbb{R}$ be linear and nontrivial.
Let $F : V \to \mathbb{R}^m$ be a linear transformation where $F = (F_1, \ldots, F_m)$ with $F_i = P \circ L^{i-1}$.

Lemma 6: Observability

A small perturbation of L is sufficient to ensure that F has rank $\min(m, n)$.

Proof: Observability

Without loss of generality, assume $n \leq m$.
By induction on n, we show that L can be suitably perturbed.
If $n = 1$, then $F = P$. Since $P \neq 0$, we have that $\text{rank}(F) = \text{rank}(P) = 1$ and the lemma is true.
Let $n > 1$. Assume that $F_1, \ldots, F_{n-1} \in V^*$ are linearly independent (induction hypothesis). If $F_1, \ldots, F_n$ are linearly independent, there is nothing to prove.
Now assume that they are not linearly independent. F_1 is a nontrivial linear form, hence $\dim(\ker(F_1)) = m-1$. Since $F_1, \ldots, F_{n-1}$ are linearly independent, we have that $\dim(\ker(F_1) \cap \ker(F_2)) = m-2$. Then we have that $\dim(\cap_{i=1}^{n-1} \ker((F_i))) = m - (n-1)$. Since $n \leq m$, we obtain $\dim(\cap_{i=1}^{n-1} \ker((F_i))) \geq 1$. Hence, there exists a nonzero $v \in V$ such that all of $F_1(v), \ldots, F_{n-1}(v)$ vanish. In other words, we have that $v, L(v), \ldots, L^{n-2}(V) \in \ker(P)$.
Assume $n \geq 3$. We would like to make $v, \ldots, L^{n-2}(v)$ linearly independent in $\ker(P)$. It is possible since $\dim(\text{span}(v, \ldots, L^{n-2}(v))) \leq n-1 \leq m-1 = \dim(\ker(P))$ because $n \leq m$. So, we perturb L in turn on the vectors $v, L(v), \ldots, L^{n-3}(v)$, so that the vectors $v, L(v), L^2(v) = L(L(v)) \ldots, L^{n-2}(v) = L(L^{n-1}(v))$ remain in $\ker(P)$ and become linearly independent.
Then, we perturb L on $L^{n-2}(v)$ while keeping it fixed on $v, L(v), \ldots, L^{n-3}(v)$ in such a way that $L^{n-1}(v) \notin \ker(P)$.
Then $F_1, \ldots, F_{n-1}$ are unaltered and vanish on v, whereas $F_n(v) \neq 0$.
Let α be such that $\sum_{i \leq n} \alpha_i F_i = 0$. Then $\sum_{i \leq n} \alpha_i F_i(v) = 0 = \alpha_n F_n(v)$ since $F_i(v) = 0$ for every $i < n$. Since $F_n(v) \neq 0$, we have that $\alpha_n = 0$. We are left with $\sum_{i \leq n-1} \alpha_i F_i = 0$. The induction hypothesis tells us that $\alpha_i = 0$ for every i. Therefore, $F_1, F_2, \ldots, F_n$ are linearly independent and F has rank n. □

Lemma 7: Genericity of immersions on periodic points

There exists an open dense subset $\Phi^*_{2m} \subset \Phi_{2m}$ such that if $\phi \in \Phi^*_{2m}$ then df_x has rank $\min(m, n)$ for each $x \in P_{2m}$.

Proof: Genericity of immersions on periodic points

Let $\phi \in \Phi_{2m}, y \in \mathcal{E}$, and $x \in P_{2m}$ with period p.
We have $d(\phi^p)_x : T_xM \to T_xM$ and $dy_x : T_xM \to \mathbb{R}$.
We can apply lemma 6, meaning that we need to perturb $d(\phi^p)_x$. We want to do so by perturbing only ϕ. $\square$

References

[1] F. Takens, "Detecting strange attractors in turbulence," *Dynamical Systems and Turbulence*, vol. 898, pp. 366–381, 1981.

[2] L. Noakes, "The takens embedding theorem," *International Journal of Bifurcation and Chaos*, vol. 1, no. 4, pp. 867–872, 1991.