A Geometric Approach to Evaluation-Transversality Techniques in Generic Bifurcation Theory

By

Søren Karl Aalto

B.Sc., St. Francis Xavier University, 1982

A Thesis Submitted in Partial Fulfilment of

The Requirements for the Degree of

Master of Science

in

The Faculty of Graduate Studies

Mathematics

We accept this thesis as conforming
to the required standard

The University of British Columbia

September, 1987

46

Department of MATHEMATICS

The University of British Columbia
1956 Main Mall
Vancouver, Canada
V6T 1Y3

Date Oct. 15, 1987

Abstract

The study of bifurcations of vectorfields is concerned with changes in qualitative behaviour that can occur when a non-stucturally stable vectorfield is perturbed. In a sense, this is the study of how such a vectorfield fails to be structurally stable. Finding a systematic approach to the study of such questions is a difficult problem.

One approach to bifurcations of vectorfields, known as "generic bifurcation theory," is the subject of much of the work of Sotomayor (Sotomayor [1973a],Sotomayor [1973b],Sotomayor [1974]). This approach attempts to construct generic families of k-parameter vectorfields (usually for k=1), for which all the bifurcations can be described. In Sotomayor [1973a]it is stated that the vectorfields associated with the "generic" bifurcations of individual critical elements for k-parameter vectorfields form submanifolds of codimension $\leq k$ of the Banach space $\mathcal{X}^r(M)$ of vectorfields on a compact manifold M. The bifurcations associated with one of these submanifolds of codimension-k are called *generic codimension-k bifurcations.* In Sotomayor [1974]the construction of these submanifolds and the description of the associated bifurcations of codimension-1 for vectorfields on two dimensional manifolds is presented in detail. The bifurcations that occur are due to the parameterised vectorfield crossing one of these manifolds transversely as the parameter changes.

Abraham and Robbin used transversality results for evaluation maps to prove the Kupka-Smale theorem in Abraham and Robbin [1967]. In this thesis, we shall show how an extension of these evaluation transversality techniques will allow us to construct the submanifolds of $\mathcal{X}^r(M)$ associated with one type of generic bifurcation of critical elements, and we shall consider how this approach might be extended to include the other well known generic bifurcations. For saddle-node type bifurcations of critical points, we will show that the changes in qualitative behaviour are related to geometric properties of the submanifold Σ_0 of $\mathcal{X}^r(M) \times M$, where Σ_0 is the pull-back of the set of zero vectors–or *zero section*–by the evaluation map for vectorfields. We will look at the relationship between the Taylor series of a vectorfield X at a critical point p and the geometry of Σ_0 at the corresponding point (X, p) of $\mathcal{X}^r(M) \times M$. This will motivate the non-degeneracy conditions for the saddle-node bifurcations, and will provide a more general geometric picture of this approach to studying bifurcations of critical points. Finally, we shall consider how this approach might be generalised to include other bifurcations of critical elements.

Table of Contents

1.1 Some Parts of a Dynamical System

The flow of a dynamical system is usually composed of several flow-invariant sets, including critical points, periodic orbits, and the stable and unstable manifolds of the critical points and periodic orbits, as well as other more complicated types of recurrent sets. By isolating each of these elements, we obtain a qualitative picture of the behaviour of a flow. When we study bifurcations, the easiest qualitative changes to consider are associated with changes in these elements because of the description of the dynamics in terms of critical elements is relatively complete and well-understood.

Definition of a Dynamical System

In the most general terms, a dynamical system is the deterministic evolution in time of the states of some state space. We shall restrict ourselves to dynamical systems defined by smooth vectorfields on spaces such as $\mathbf{R}^n$ and other finite-dimensional smooth manifolds. For example, suppose X is a smooth vectorfield defined on an open subset U of $\mathbf{R}^n$ which vanishes off a compact subset of U. Then X is a smooth map $X : U \longrightarrow \mathbf{R}^n$ which defines a differential equation

$$\frac{dx}{dt} = X(x).$$

The solution to this differential equation is a map

$$\Phi : U \times \mathbf{R} \rightarrow U,$$

where $\Phi(x,\cdot)$ is the unique solution to the initial value problem

$$\frac{d}{dt}x(t) = X(x(t)), \quad x(0) = x.$$

The map Φ is called the *flow* of the vectorfield X. Several properties of the flow follow from elementary existence and uniqueness theory for ordinary differential equations—A solution to the above initial value problem exists for suitably small t, and any such solution can be extended to be defined for all real t. Furthermore, this solution is guaranteed to be unique and to depend smoothly (with the same smoothness as X) on initial conditions and time. Thus, it is is easily shown that the flow is a 1-parameter group of diffeomorphisms of U under composition in the following way;

$$\Phi(\cdot, s+t) = \Phi(\cdot, s)\circ\Phi(\cdot, t).$$

The flow of a dynamical system gives the time evolution of states/points/initial conditions in this way. Also, given a flow Φ on a subset U of $\mathbf{R}^n$, we may obtain the associated vectorfield X by differentiating with respect to time, i.e.,

$$X(x) = \frac{\partial}{\partial t}\Phi(0,x).$$

The derivative of the flow in the variable x is given by the *variational equation*

$$\frac{d}{dt}\Big(\frac{\partial}{\partial x}\Phi(x,t)\Big) = \frac{\partial}{\partial x}X(\Phi(x,t))\cdot\frac{\partial}{\partial x}\Phi(x,t).$$

A full accounting of these results can be found in Abraham *et. al.* [1983], sect. 4.1.

In subsequent chapters of this thesis, we will be exclusively concerned with dynamical systems defined on compact smooth manifolds. This restriction is necessitated by the requirements of the evaluation transversality lemmas used in chapters 5 and 6. While there is no difficulty in defining a flow on a compact manifold, it is not immediately obvious what is meant by a vectorfield on a compact manifold, and so not obvious what the result of differentiating a flow with respect to time would be.

We will define vectorfields on manifolds and derivatives of maps between manifolds in sect. 2.1. For the present we will assume that we are working with dynamical systems defined on $\mathbf{R}^n$. Most of the definitions in this section are topological in nature, so that corresponding generalizations to the case of dynamical systems on smooth manifolds is immediate.

Orbits, Trajectories

The *trajectory* $\gamma_x(t)$ of a point $x \in \mathbf{R}^n$ for a vectorfield $X : \mathbf{R}^n \to \mathbf{R}^n$ is the solution to the initial value problem

$$\frac{d}{dt}\gamma_x(t) = X(\gamma_x(t)), \quad \gamma_x(0) = x,$$

or in terms of the flow Φ_X of X;

$$\gamma_x(t) = \Phi_X(x,t).$$

The *orbit* of the point x is the set of points in the (range of the) trajectory of x;

$$\begin{aligned} O(x) &= \{\gamma_x(t) | t \in \mathbf{R}\} \\ &= \{\Phi_X(x,t) | t \in \mathbf{R}\}. \end{aligned}$$

The trajectories of X partition $\mathbf{R}^n$ into orbits—each point of $\mathbf{R}^n$ is in exactly one orbit of X. Also, the orbits are trivially the smallest sets invariant under the flow.

Critical Elements

Certain kinds of orbits are of particular interest. A *critical point* is an orbit consisting of a single point. The trajectory of a critical point $x \in \mathbf{R}^n$ is just the constant solution $\gamma_x(t) = x$. A *periodic orbit* is the orbit of a *periodic point*; i.e. a point x whose trajectory $\gamma_x(t)$ is a *periodic trajectory*– which means there exists $\tau \in \mathbf{R}$ such that $\gamma_x(\tau) = x$. The smallest positive τ such that $\gamma_x(\tau) = x$ is called the *period* of the periodic orbit/point/trajectory. Note that

$$\begin{aligned} \gamma_x(t+\tau) &= \Phi(x,t+\tau) = \Phi(\gamma_x(\tau),t) \\ &= \Phi(x,t) = \gamma_x(t), \end{aligned}$$

so we see that all points of the periodic orbit are periodic with the same period.

The critical points and periodic orbits of a vectorfield X are collectively referred to as the *critical elements* of X and denoted by $\Gamma(X)$. These are the most basic *recurrent sets* of the vectorfield X.

Limit Sets

A trajectory γ_x of X that remains bounded will have an orbit with compact closure. In the case of a dynamical system on a compact set or manifold, all orbits will have compact closure. For such orbits we may define the $\alpha-$ and $\omega-$ limit sets of the trajectory γ_x by

$$\omega(x) = \bigcap_{T>0} \overline{\{\gamma_x(t) | t > T\}},$$

and

$$\alpha(x) = \bigcap_{T<0} \overline{\{\gamma_x(t) | t < T\}}.$$

The set $\alpha(x), (resp. \omega(x))$ are where the trajectory through x ends up when $t \to \infty$ $(resp. t \to -\infty)$. Intuitively, the orbit through the point x is "born" in $\alpha(x)$ and "dies" in $\omega(x)$.

Orbit Structure

Critical elements are their own α- and ω-limit sets, and so exhibit a very strong form of recurrence. The α- and ω-limit sets of any point are invariant under the flow. The orbit through any point x "joins" its $\omega-$ limit set to its $\alpha-$ limit set in the sense that the trajectory γ_x through x as the trajectory tends to the invariant set $\omega(x)$ as $t \to -\infty$ and tends to $\alpha(x)$ as $t \to +\infty$. Of course, these sets $\omega(x), \alpha(x)$ may be equal, such as in the case of critical elements.

We may think of the *orbit structure* of the dynamical system, in qualitative terms, as consisting of recurrent sets that are joined by other "connecting" orbits. If S is a compact flow-invariant set for the flow Φ, then we may define the *inset* (resp. *outset*) of S is the set of all points x such that $\omega(x) \subset S$ (resp. $\alpha(x) \subset S$). We can obtain a great deal of information about the orbit structure by describing the recurrent sets of the flow and indicating which of these recurrent sets are joined by "connecting" orbits, i.e., orbits that are in the inset of one recurrent set and in the outset of another. Of course, the structure of the recurrent sets and the various connections can be very complicated. In the simplest case, the only recurrent sets would be the critical elements of the dynamical system, and the orbit structure of the system would consist of critical points, periodic orbits, and orbits connecting various critical elements. More general cases involve more complicated recurrent sets associated with less strong notions of recurrence. One important type of recurrent point is a *non-wandering* point. *A point x is non-wandering if for any neighborhood U_x of x and time $T > 0$, there exists a $t > T$ such that some of the orbits starting in U_x have come back to U_x*; i.e.,

$$\Phi(U_x, t) \bigcap U_x \neq \emptyset.$$

The set of non-wandering points of a vectorfield X is denoted by $\Omega(X)$. This is a comparatively general notion of recurrence, and difficult to understand well. Below we shall consider a family of dynamical systems where all of the recurrence is in the critical elements.

Stable Manifolds of Critical Points

Since the trajectory of a critical point is constant, it is easy to compute the derivative of the flow with respect to initial conditions at a critical point. We recall that

$$\begin{aligned}\frac{d}{dt}\left(\frac{\partial}{\partial x}\Phi(x,t)\right) &= \frac{\partial}{\partial x}X(\Phi(x,t)) \cdot \frac{\partial}{\partial x}\Phi(x,t) \\ &= \frac{\partial}{\partial x}X(x) \cdot \frac{\partial}{\partial x}\Phi(x,t),\end{aligned}$$

if x is a critical point. Thus, $\frac{\partial}{\partial x}\Phi(x,t) = \exp(t \cdot \frac{\partial}{\partial x}X(x))$. The eigenvalues of the matrix $\frac{\partial}{\partial x}X(x)$ are called the *characteristic exponents* of the critical point x.

The characteristic exponents of a critical point are the various exponential rates of growth and decay for the linearization of the flow at the critical point. One might (correctly) expect that if none of these rates were zero, that the linear behaviour would be dominant near the critical point. For example, if all of the characteristic exponents of the critical point x have negative real part, then the linearization of the flow contracts all perturbations exponentialy and we can show that x is an *asymptotically stable* or *attracting* critical point. This means that there is a neighborhood V_x of x such that for any ϵ-neighborhood U of x there is a $T > 0$ such that $\Phi(V_x, t) \subset U$ for all $t > T$. In particular this means that x is the ω-limit of all points in V_x, and that V_x is in the inset of x.

If no eigenvalues of the critical point x have zero real part, then x is a ***hyperbolic*** critical point. In this case the linearization of the flow contracts perturbations in the subspace E_s corresponding to eigenvalues with negative real parts and expands perturbations in the subspace E_u corresponding to eigenvalues with positive real parts at various exponential rates. The *stable manifold theorem* tells us about the structure of the inset and outset of x;

Theorem (Local Stable Manifold Theorem). *If x is a hyperbolic critical point of the smooth vectorfield X, then there is an ϵ-neighborhood U_ϵ of x such that the subsets $W^s_\epsilon(x), W^u_\epsilon(x)$ of U_ϵ which are characterized by*

$$W^s_\epsilon(x) = \Big\{ y \in U_\epsilon | \gamma_y(t) \to x \text{ as } t \to +\infty \text{ and } \gamma_y(t) \in U_\epsilon, \ \forall t > 0 \Big\},$$

$$W^u_\epsilon(x) = \Big\{ y \in U_\epsilon | \gamma_y(t) \to x \text{ as } t \to -\infty \text{ and } \gamma_y(t) \in U_\epsilon, \ \forall t < 0 \Big\}.$$

are submanifolds of U_ϵ, called the local stable and unstable manifolds of the critical point x. Furthermore, the tangent spaces $T_x W^s_\epsilon(x), T_x W^u_\epsilon(x)$ of the local stable manifolds at the critical point are the subspaces E_s, E_u mentioned above.

A statement of this result with references is given in Guckenheimer and Holmes [1983], p. 13, Theorem 1.3.2, and is also given with proof as Theorem 27.1 of Abraham and Robbin [1967].

Points in $W^s_\epsilon(x)$ are attracted to x along $W^s_\epsilon(x)$ at exponential rates. This manifold is the set of points near x that are attracted to x without first wandering away. Points that are not in $W^s_\epsilon(x)$ may be attracted to x eventually, but will first have to leave the neighborhood U_ϵ. Corresponding statements hold for $W^u_\epsilon(x)$. Clearly, the local stable manifold is invariant under the flow in positive time. If we take the union

$$W^s(x) = \bigcup_{t \in \mathbf{R}} \Phi(W^s_\epsilon(x), t)$$

then $W^s(x)$ is an injectively immersed flow-invariant submanifold, as it is an expanding union of embedded submanifolds since the local stable manifold is invariant under the flow in positive time. $W^s(x)$ is called the *stable manifold* of x, and is in fact the inset of x as the orbit of any point in W^s must eventually end up in the local stable manifold, and so be in the inset of x, whereas points not in the stable manifold will never end up in $W^s_{loc}(x)$, and so must leave the neighborhood U_ϵ of x for arbitrarily large time. This characterizes the set of points that tend asymptotically to the critical point x under the flow. Similarly we may define the unstable manifold of x. Thus, the inset and outset of a hyperbolic critical point have the structure of immersed submanifolds.

Periodic Orbits, Characteristic Multipliers

We would like to extend the above results to periodic orbits. Specifically, we want to know which orbits will be attracted to a periodic orbit $\gamma(t)$. If τ is the period of γ, then and x is a point in the orbit of γ, then $\Phi(x,\tau) = x$. We would expect that orbits near γ would approach γ along directions transverse to the orbit of γ which are contracted by $\frac{\partial}{\partial x}\Phi(\cdot,\tau)$. We need to consider the asymptotic behaviour of perturbations transverse to the orbit of γ. Let S be an $n-1$ dimensional subspace (or submanifold) that intersects the orbit of γ transversely at x ($X(x)$ does not lie in S). Then we may define a diffeomorphism of a neighborhood of x in S, called the *Poincaré map* of the periodic orbit γ. Points y of S are defined by $n \cdot (y - x) = 0$, where n is the normal vector of the subspace S. Then $n \cdot (\Phi(x, \tau) - x) = 0$ and since $\frac{\partial}{\partial t}(n \cdot (\Phi(x, \tau) - x) = n \cdot \frac{\partial}{\partial t}\Phi(x, \tau) = n \cdot X(x) \neq 0$, then by the implicit function theorem, we have that there is a map $\tau(y)$ in a neighborhood U of x such that $n \cdot (\Phi(y, \tau(y)) - x = 0$. This means that $\Phi(y, \tau(y)) \in S$. If we restrict y to S, then we end up with a map from a neighborhood of x in S to a neighborhood of x in S. This is the Poincaré map of the periodic orbit γ which we shall denote by Θ. We have for y in S near x

$$\Theta(y) = \Phi(y, \tau(y)),$$

$$\Theta : S|(S \cap U) \to S.$$

The Poincaré map is smooth (as smooth as the vectorfield X) as it is the composition of smooth maps. Furthermore, Θ is a local diffeomorphism by the inverse function theorem. This follows as we may compute

$$\begin{aligned}\frac{\partial}{\partial y}\Theta(x) \cdot v &= \frac{\partial}{\partial x}\Phi(x, \tau) \cdot v + \frac{\partial}{\partial t}\Phi(x, \tau) \cdot \frac{\partial}{\partial x}\tau(x) \cdot v \\ &= \frac{\partial}{\partial x}\Phi(x, \tau) \cdot v + (\frac{\partial}{\partial x}\tau(x) \cdot v)X(x)\end{aligned}$$

for v in S. Since $\frac{\partial}{\partial x}\Phi(x, \tau)$ is onto $\mathbf{R}^n$, then the kernel of $\frac{\partial}{\partial y}\Theta(x)$ must consist of vectors that are mapped along the direction of $X(x)$ by $\frac{\partial}{\partial x}\Phi(x, \tau)$. But $X(x)$ is invariant under $\frac{\partial}{\partial x}\Phi(x, \tau)$ by the following;

$$\Phi(x, \tau) = \Phi(\Phi(x, 0), \tau),$$

so that differentiating with respect to time gives

$$\begin{aligned}\frac{\partial}{\partial t}\Phi(x, \tau) &= X(\Phi(x, \tau)) = X(x) \\ &= \frac{\partial}{\partial t}\Phi(\Phi(x, 0), \tau) \\ &= \frac{\partial}{\partial x}\Phi(x, \tau) \cdot \frac{\partial}{\partial t}\Phi(x, 0) \\ &= \frac{\partial}{\partial x}\Phi(x, \tau) \cdot X(x).\end{aligned}$$

Thus, the kernel of $\frac{\partial}{\partial y}\Theta(x)$ is along the direction of $X(x)$, which is not in S.

The Poincaré map of a periodic orbit describes the behaviour of nearby orbits in directions transverse to the orbit. The image of a point y in S under Θ is the point of first intersection (in positive time) of the orbit through y with S. The asymptotic behaviour of points under Θ indicates the asymptotic behaviour of the associated orbits. If the successive images of point approach x, then the orbit through this point is asymptotic to the periodic orbit γ. But the main reason for introducing the Poincaré map is that there is a version of the stable manifold theorem for fixed points of diffeomorphisms. We expect that nearby orbits will be attracted along directions that are expanded by the derivative $\frac{\partial}{\partial y}\Theta(x)$ of the Poincaré map, and repelled along directions that are contracted by $\frac{\partial}{\partial y}\Theta(x)$, by analogy with the stable manifold theorem for critical points. We state the theorem

Theorem (Stable Manifold Theorem for Diffeomorphisms). *Let f be a diffeomorphism with fixed point x. If x is a hyperbolic fixed point (Df has no eigenvalues of unit modulus), then there is a neighborhood U of x such that the sets*

$$W^s_{loc}(x) = \Big\{ y \in U | f^n(y) \rightarrow x \text{ as } n \rightarrow +\infty \text{ and } f^n(x) \in U,\ n > 0 \Big\}$$
$$W^u_{loc}(x) = \Big\{ y \in U | f^n(y) \rightarrow x \text{ as } n \rightarrow -\infty \text{ and } f^n(x) \in U,\ n < 0 \Big\}$$

are submanifolds in U, and the tangent spaces $T_xW^s_{loc}(x)$(resp. $T_xW^u_{loc}(x)$) to these manifolds at x are the subspaces corresponding to eigenvalues of $Df(x)$ of modulus greater than 1 (resp. less than 1).

All of the observations we made for the local stable and unstable manifolds of a critical point hold for W^s_{loc}, W^u_{loc}.

The eigenvalues of the Poincaré map are called the ***characteristic multipliers*** or ***Floquet multipliers*** of the periodic orbit. A periodic orbit is hyperbolic if the corresponding fixed point of its Poincaré map is a hyperbolic fixed point. We define the stable and unstable manifolds of a hyperbolic periodic orbit γ as the unions of all orbits passing through the local stable and unstable manifolds of the Poincaré map of the γ. These are injectively immersed submanifolds which are the inset and outset of the periodic orbit.

Morse-Smale Systems

Now we define a family of vectorfields for which we have a fairly complete description of the orbit structure. A vectorfield is a *Morse-Smale* vectorfield if it satisfies:

(1) There are a finite number of critical elements, and each is hyperbolic.

(2) All stable and unstable manifolds of critical elements must intersect transversely, and

(3) The non-wandering set Ω consists only of critical elements.

This defines a family of vectorfields whose orbit structures are relatively simple. Given a Morse-Smale vectorfield X, we may define its *phase diagram* Γ,

Definition. *The phase diagram Γ of a Morse-Smale vectorfield X is the set of critical elements of X with the following partial order: If σ_1, σ_2 are critical elements of X, then $\sigma_1 \prec \sigma_2$ if $W^s(\sigma_1) \cap W^u(\sigma_2) \neq \emptyset$. In other words, $\sigma_1 \prec \sigma_2$ if there is an orbit joining σ_1 to σ_2 that is "born" in σ_2 and dies in σ_1.*

The phase diagram of a Morse-Smale vectorfield gives us a great deal of information about the flow of the vectorfield.

1.2 An Example of a Bifurcation

In this section we will examine the bifurcations that arise from the failure of a critical point of a vectorfield to persist smoothly under perturbations of the vectorfield. The usual approach to bifurcation theory involves dynamical systems depending on parameters and the analysis of qualitative changes that occur in the dynamics as the dynamical system is perturbed by varying the parameters. As an example, let us consider a vectorfield X on a 1-dimensional manifold (R^1 , or its one-point compactification S^1 —if we insist on compactness) depending on a scalar parameter μ . Then X is a function

$$X : \mathbf{R}^1 \times \mathbf{R}^1 \longrightarrow \mathbf{R}^1 : (x, \mu) \longmapsto X(x, \mu).$$

Assume that X is at least C^2 .

Non-Degenerate Critical Points

Now, suppose that x_0 is a critical point of $X(., \mu_0)$, i.e., that $X(x_0, \mu_0) = 0$. We may examine what happens to the critical point x_0 as the parameter μ is varied near μ_0 by looking at the solutions of $X(x, \mu) = 0$ that are near (x_0, μ_0) . For example, if $\frac{\partial}{\partial x} X(x_0, \mu_0) \neq 0$, then by the implicit function theorem of calculus, there is a smooth (at least C^2) function $x(\mu)$, defined near μ_0 such that $X(x(\mu), \mu) = 0$. Furthermore, the implicit function theorem states that the curve $(x(\mu), \mu)$ is the unique solution of $X(x, \mu) = 0$ in some neighborhood of (x_0, μ_0).

Thus, the critical point x_0 varies smoothly as the parameter μ is varied near μ_0 and no new critical points appear near x_0. This is shown in the graph of Fig. 1.

Critical points of this type are said to be *non-degenerate.*

Degenerate Critical Points

However, if the critical point x_0 is *degenerate* in the sense that $\frac{\partial}{\partial x}X(x_0, \mu_0) = 0$ then it may fail to persist as the parameter μ is varied. If $\frac{\partial}{\partial \mu}X(x_0, \mu_0) \neq 0$ then the implicit function theorem may be applied as before to conclude there is a smooth function $\mu(x)$ such that $X(x, \mu(x)) = 0$ near x_0. Furthermore, $\mu'(x_0) = 0$ as $\frac{\partial}{\partial x}X(x, \mu(x)) + \frac{\partial}{\partial \mu}X(x, \mu(x))\mu'(x) = 0$ by implicit differentiation. The graph of such a curve $(x, \mu(x))$ is shown in Fig. 2.

We can obtain some qualitative information about the bifurcation occurring at (x_0, μ_0) from the graph of Fig. 2. When $\mu < \mu_0$ we see that the vectorfield $X(\cdot, \mu)$ has two distinct critical points near x_0. At the *bifurcation value* of the parameter, μ_0, there is only one critical point of $X(\cdot, \mu_0)$ indicated by the graph, x_0. For parameter values above μ_0 there are no critical points of $X(\cdot, \mu)$ near x_0. There is an obvious change in qualitative behaviour as the parameter μ is varied through the bifurcation value μ_0 as the number of critical points of the vectorfield near x_0 changes. This change in behaviour is caused by two distinct critical points of the vectorfield coalescing and annihilating each other.

Taylor Series Conditions

Of course, the above qualitative analysis depends on a qualitative property of the graph of Fig. 2—that it is concave at x_0. A sufficient condition for this concavity is $\mu''(x_0) \neq 0$. In terms of the Taylor series of the vectorfield X at (x_0, μ_0) the

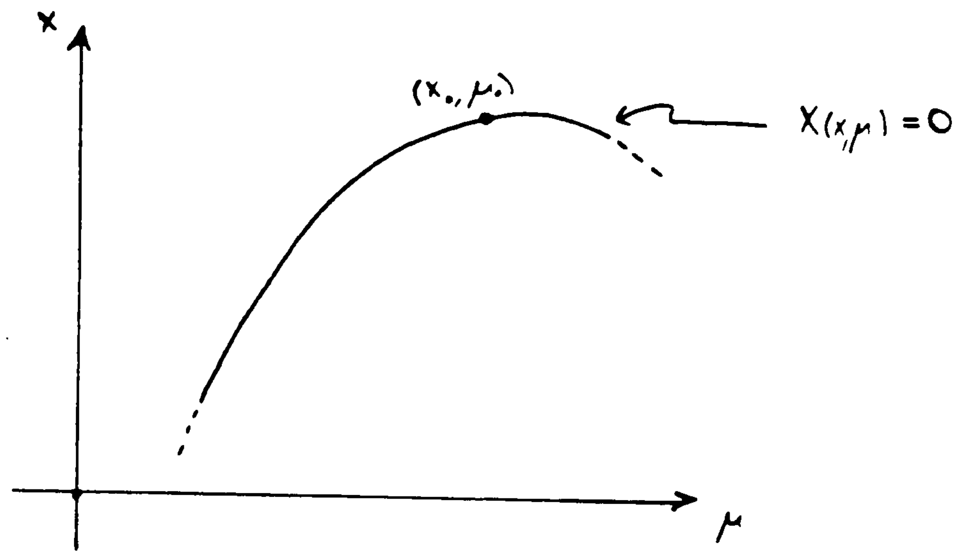

Figure 1.2.1

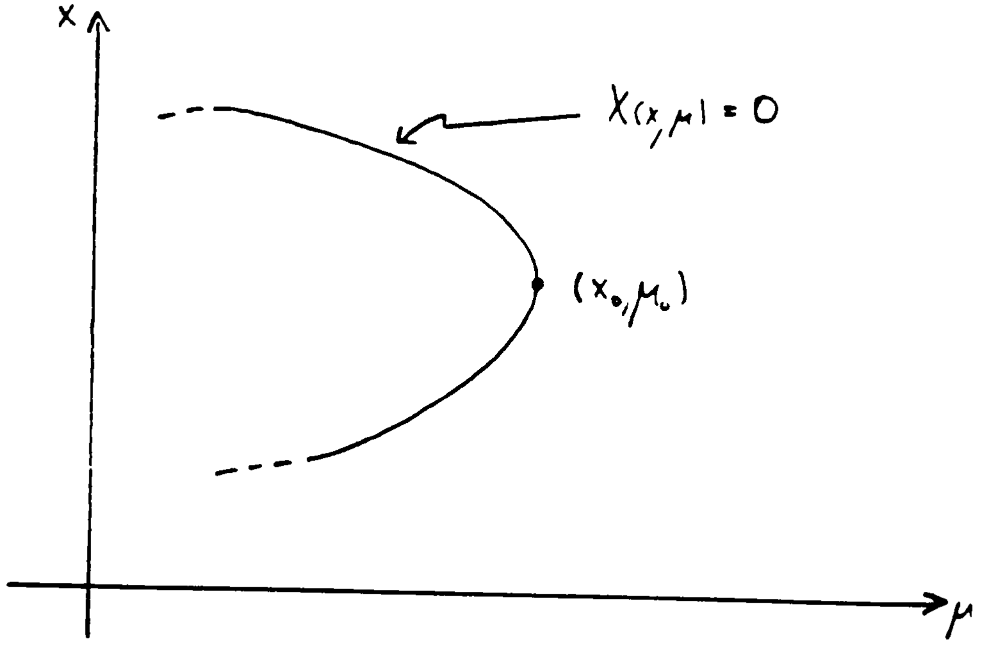

Figure 1.2.2

sufficient condition above is equivalent to the following:

$$X(x_0,\mu_0) = 0, \quad \frac{\partial}{\partial x}X(x_0,\mu_0) = 0,$$
$$\frac{\partial}{\partial \mu}X(x_0,\mu_0) \neq 0,$$
$$\text{and} \quad \frac{\partial^2}{\partial x^2}X(x_0,\mu_0) \neq 0.$$

Indeed, $X(x,\mu(x)) = 0$, so that

$$\frac{\partial}{\partial x}X(x,\mu(x)) + \frac{\partial}{\partial \mu}X(x,\mu(x))\mu'(x) = 0,$$

and also

$$\frac{\partial^2}{\partial x^2}X(x,\mu(x)) + 2\frac{\partial^2}{\partial x \partial \mu}2X(x,\mu(x))\mu'(x) +$$
$$\frac{\partial^2}{\partial \mu^2}X(x,\mu(x))(\mu'(x))^2 + \frac{\partial}{\partial \mu}X(x,\mu(x))\mu''(x) = 0.$$

by implicit differentiation. We have shown that $\mu'(x_0) = 0$ as $\frac{\partial}{\partial x}X(x_0,\mu(x_0)) = \frac{\partial}{\partial x}X(x_0,\mu_0) = 0$. Then, from the last relation,

$$\mu''(x_0) = -\frac{\frac{\partial^2}{\partial x^2}X(x_0,\mu_0)}{\frac{\partial}{\partial \mu}X(x_0,\mu_0)}.$$

so that $\mu''(x_0) \neq 0$ if and only if $\frac{\partial^2}{\partial x^2}X(x_0,\mu_0) \neq 0$.

Now consider the graph of the function $X(\cdot,\mu) : \mathbf{R}^1 \to \mathbf{R}^1 : x \mapsto X(x,\mu)$ for fixed μ. For $\mu = \mu_0$ we have that $\frac{\partial}{\partial x}X(x_0,\mu_0) = 0$ and also $\frac{\partial^2}{\partial x^2}X(x_0,\mu_0) \neq 0$ so that the function $X(.,\mu_0)$ has a local extreme value at $x = x_0$. Also, $X(x_0,\mu_0) = 0$ so that this local extreme value is zero. From the implicit function theorem, there is a C^2 function $\tilde{x}(\mu)$ defined for μ near μ_0 such that $\frac{\partial}{\partial x}X(\tilde{x}(\mu),\mu) = 0$ and $\tilde{x}(\mu_0) = x_0$ as we have assumed that $\frac{\partial^2}{\partial x^2}X(x_0,\mu_0) \neq 0$. We can see that the point $\tilde{x}(\mu)$ must be a local extreme point for the graph of $X(\cdot,\mu)$ as $\frac{\partial^2}{\partial x^2}X \neq 0$ in some neighborhood of (x_0,μ_0). Also, since $X(x_0,\mu_0) = X(\tilde{x}(\mu),\mu) = 0$ we know that the local extreme value for the graph of $X(\cdot,\mu_0)$ is zero. Let us compute the derivative of the extreme value as the parameter μ is changed, $\frac{d}{d\mu}X(\tilde{x}(\mu),\mu)$, at $\mu = \mu_0$.

$$\frac{d}{d\mu}X(\tilde{x}(\mu),\mu) = \frac{\partial}{\partial x}X(\tilde{x}(\mu),\mu)\tilde{x}'(\mu) + \frac{\partial}{\partial \mu}X(\tilde{x}(\mu),\mu).$$

But, at $\mu = \mu_0$, we have $\tilde{x}(\mu_0) = x_0$, and also $\frac{\partial}{\partial x}X(x_0, \mu_0) = 0$ so that

$$\frac{d}{d\mu}X(\tilde{x}(\mu_0), \mu_0) = \frac{\partial}{\partial\mu}X(x_0, \mu_0) \neq 0,$$

so that the value of the local extreme point $\tilde{x}(\mu)$ of $X(\cdot, \mu)$ changes sign as μ is varied through the bifurcation value μ_0. This shows us that the graphs of the $X(\cdot, \mu)$ are qualitatively the same as what is shown in Fig. 3.

In Fig. 3. we see that as the parameter μ is increased through μ_0 the local extreme value of the graph of $X(\cdot, \mu)$ (which in this case is a local minimum) changes from negative to positive. This induces a change in the qualitative behaviour of our system near x_0. Because the graph of $X(\cdot, \mu)$ is concave up near x_0, then when the local minimum of this graph is negative ($\mu < \mu_0$) then there must be two distinct zeroes of the graph of $X(\cdot, \mu)$, which will correspond to two distinct critical points of the vectorfield $X(\cdot, \mu)$. As the parameter μ approaches the bifurcation value μ_0, these two zeroes approach each other, until, at the bifurcation value μ_0 the the two zeroes of the graph meet, and the graph has a double zero corresponding to a quadratic tangency. Above the bifurcation value μ_0 the graph of $X(\cdot, \mu)$ has no roots near x_0. It is worth noting that this qualitative analysis depends only on the conditions we have given on the Taylor series of X at (x_0, μ_0), which state that X has a root at (x_0, μ_0) which is non-degenerate in the μ-direction (i.e. $\frac{\partial}{\partial\mu}X(x_0, \mu_0) \neq 0$) and quadratically tangent in the x-direction (i.e. $\frac{\partial}{\partial x}X(x_0, \mu_0) = 0, \frac{\partial^2}{\partial x^2}X(x_0, \mu_0) \neq 0$).

Orbit Structure

In the 1-dimensional case it is easy to obtain the *phase diagram* of a vectorfield from its graph. The critical points of the vectorfield are the zeroes of the graph of the function. A zero such that the derivative $\frac{\partial}{\partial x}X$ is positive is a source, and if $\frac{\partial}{\partial x}X$ is negative then it is a sink. A double zero is neither a source nor a sink, but is instead the coalescence of a source and a sink as shown in the middle phase portrait

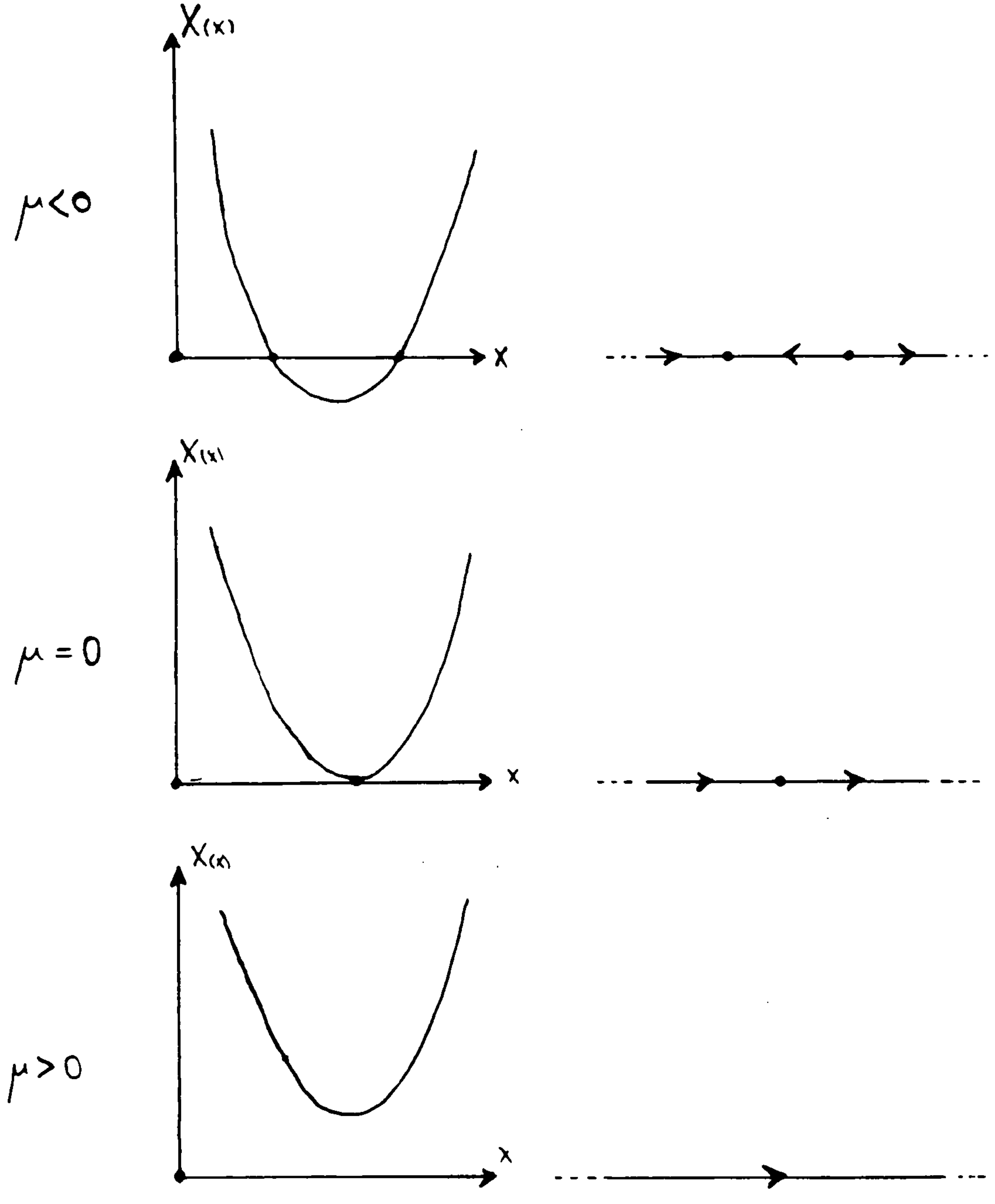

Figure 1.2.3

of Fig. 3. The critical points of the vectorfield are joined by orbits between them, the direction of these orbits being determined by the sign of the function in between the two zeroes associated with the critical points. For adjacent pairs of non-degenerate critical points the orbits joining them will start in a source and end in a sink. From this viewpoint the bifurcation shown in Fig. 3 is caused by two critical points, one source and one sink, coalescing and annihilating one another.

Generic Conditions for Vectorfields

In the above analyses, we have prescribed various conditions on the Taylor series of the vectorfield X about a critical point (x_0, μ_0). It seems reasonable to ask why we should choose these particular conditions instead of some others, or to ask if are "likely" to hold. Consider a vectorfield X on R^1 with a critical point x_0 which is non-degenerate in the sense that $\frac{\partial}{\partial x} X(x_0) \neq 0$. We know that such a critical point will persist smoothly under small perturbations of the vectorfield X, indeed, if $\tilde{X}(x, \epsilon) = X(x) + \epsilon\xi(x)$, then we have already seen that for small values of ϵ the vectorfield $\tilde{X}$ has a critical point $x(\epsilon)$ near x_0 where $x(\epsilon)$ is a C^2 function. It is easy to see that this new critical point $x(\epsilon)$ will also be non-degenerate as

$$\frac{\partial}{\partial x}\tilde{X}(x(\epsilon), \epsilon) = \frac{\partial}{\partial x}X(x(\epsilon)) + \epsilon\frac{\partial}{\partial x}\xi(x)$$

which will be non-zero for ϵ sufficiently small. This shows that non-degenerate critical points persist and remain non-degenerate under small perturbations. On any compact subset of $\mathbf{R}^1$ (or on a compact manifold such as S^1), there will be only a finite number of non-degenerate critical points, as non-degenerate critical points are separated from other critical points by finite distance. So, in the case of a vectorfield defined on a compact set or manifold, the property of a vectorfield having all of its critical points be non-degenerate will persist under suitably small perturbations of the vectorfield.

In the analysis of the bifurcation above, the vectorfield $X(., \mu)$ had a degenerate

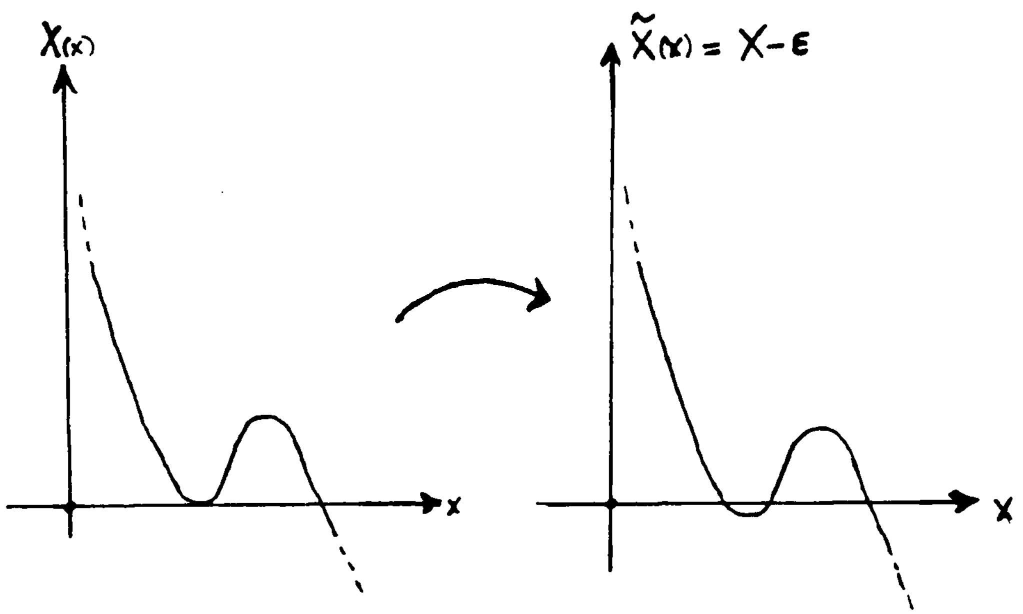

Figure 1.2.4. Perturbing a vectorfield to make 0 a regular value

critical point x_0 for one value μ_0 of the parameter, but this degenerate critical point either vanished or became two distinct non-degenerate critical points if the value of the parameter μ was changed. Most critical points of vectorfields are non-degenerate in the sense that a vectorfield with some degenerate critical points can, through arbitrarily small perturbations, be made a vectorfield whose critical points are all non-degenerate. Indeed, Sard's theorem states that for a C^1 map from $\mathbf{R}^1$ to $\mathbf{R}^1$ that the set of critical values of the map has measure zero and hence is nowhere dense. If the vectorfield $X : \mathbf{R}^1 \to \mathbf{R}^1$ has a degenerate critical point, then 0 is not a regular value of X. But we can find an arbitrarily small $\epsilon \in \mathbf{R}^1$ that is a regular value for X (unless we have chosen $X \equiv 0$, which is silly). Then the perturbed vectorfield $\tilde{X} = X - \epsilon$ has zero as a regular value and all the critical points of $\tilde{X}$ are non-degenerate as an Fig. 4.

Thus, we see that for such vectorfields on a compact space the property of having all critical points non-degenerate is "open" in the sense of persisting under perturbation and "dense" in that any vectorfield can be approximated by one with only non-degenerate critical points. This property is then said to be *generic* for vectorfields (on 1-dimensional compact manifolds).

Generic Conditions for 1-Parameter Vectorfields

For vectorfields depending on a parameter, we still expect that the zeroes of the function $X : (x,\mu) \mapsto X(x,\mu)$ will be generically non-degenerate, but now the non-degeneracy means something different for the critical points of the various vectorfields $X(.,\mu)$ from the non-parameterized case. Since X is now a map from $\mathbf{R}^2$ to $\mathbf{R}^1$, the derivative of X at a point (x_0,μ_0) is a linear map

$$DX(x_0,\mu_0) : \mathbf{R}^2 \longrightarrow \mathbf{R}^1 : (v,w) \longmapsto \frac{\partial}{\partial x}X(x_0,\mu_0)v + \frac{\partial}{\partial \mu}X(x_0,\mu_0)w.$$

Even if (x_0,μ_0) is a regular point of X, there will be some **direction** (v_0,w_0) such that $DX(x_0,\mu_0)(v_0,w_0) = 0$. If (x_0,μ_0) is a zero of X, then there is a curve tangent to (v_0,w_0) at (x_0,μ_0) along which the value of X is zero. This follows from a corollary of the implicit function theorem. Thus, one could still have $\frac{\partial}{\partial x}X(x_0,\mu_0) = 0$ if (x_0,μ_0) is a regular point of X, but only if $\frac{\partial}{\partial \mu}X(x_0,\mu_0) \neq 0$.

If a vectorfield $X(x,\mu)$ has a critical point (x_0,μ_0) such that $\frac{\partial}{\partial x}X(x_0,\mu_0) = 0$, and $\frac{\partial}{\partial \mu}X(x_0,\mu_0) \neq 0$, then, does this type of critical point persist under perturbation? Under the non-degeneracy condition we have assumed for this bifurcation, $\frac{\partial^2}{\partial x^2}X(x_0,\mu_0) \neq 0$, we may apply the implicit function theorem to the map

$$J^1X_\epsilon : (x,\mu,\epsilon) \longmapsto \left(X_\epsilon(x,\mu), \frac{\partial}{\partial x}X_\epsilon(x,\mu)\right)$$

where X_ϵ is the perturbed vectorfield

$$X_\epsilon(x,\mu) = X(x,\mu) + \epsilon\xi(x,\mu).$$

Indeed, we compute

$$\frac{\partial}{\partial x}J^1X_\epsilon(x,\mu,\epsilon) = \left(\frac{\partial}{\partial x}X(x,\mu) + \epsilon\frac{\partial}{\partial x}\xi(x,\mu), \frac{\partial^2}{\partial x^2}X(x,\mu) + \epsilon\frac{\partial^2}{\partial x^2}\xi(x,\mu)\right)$$

and

$$\frac{\partial}{\partial \mu} J^1 X_\epsilon(x,\mu,\epsilon) = (\frac{\partial}{\partial \mu} X(x,\mu) + \epsilon \frac{\partial}{\partial \mu}\xi(x,\mu), \frac{\partial^2}{\partial x \partial \mu} X_\epsilon(x,\mu)).$$

when $(x,\mu,\epsilon) = (x_0,\mu_0,0)$, this becomes

$$\begin{pmatrix} \frac{\partial}{\partial x} J^1 X_\epsilon(x_0,\mu_0,0) \\ \frac{\partial}{\partial \mu} J^1 X_\epsilon(x_0,\mu_0,0) \end{pmatrix} = \begin{pmatrix} 0 & \frac{\partial^2}{\partial x^2} X(x_0,\mu_0) \\ \frac{\partial}{\partial \mu} X(x_0,\mu_0), & \frac{\partial^2}{\partial \mu \partial x} X(x_0,\mu_0) \end{pmatrix}$$

as $\frac{\partial}{\partial x} X(x_0,\mu_0) = 0$. This is an invertable matrix as both of the terms $\frac{\partial^2}{\partial x^2} X(x_0,\mu_0)$ and $\frac{\partial}{\partial \mu} X(x_0,\mu_0)$ are nonzero. Thus, it follows from the implicit function theorem that there are points $(x_0(\epsilon),\mu_0(\epsilon))$ depending smoothly on ϵ such that

$$X_\epsilon(x_0(\epsilon),\mu_0(\epsilon)) = 0$$

and

$$\frac{\partial}{\partial x} X_\epsilon(x_0(\epsilon),\mu_0(\epsilon)) = 0.$$

Thus, the degenerate critical point x_0 persists *under perturbations of the whole 1-parameter family of vectorfields* $X(x,\mu)$. Since a 1-parameter family of vectorfields is the smallest family which may contain this kind of degenerate critical point in a persistent way (recall that it wasn't persistent for vectorfields with no parameters), we call this type of critical point a *codimension-1* critical point.

We may also use Sard's theorem to obtain an approximation result for 1-parameter families of vectorfields. Consider the map

$$J^1 X : (x,\mu) \longmapsto (X(x,\mu), \frac{\partial}{\partial x} X(x,\mu)).$$

Since we have assumed that X was at least C^2 , then the map $J^1 X$ is a C^1 map. Sard's theorem gives us the existence of a regular value (ϵ_0,ϵ_1) of $J^1 X$, arbitrarily close to $(0,0)$. Then, if the range of X contains a neighborhood of 0 , we have that for the perturbed vectorfield

$$\tilde{X}(x,\mu) = X(x,\mu) - \epsilon_0 - \epsilon_1 x,$$

(0,0) is a regular value for the associated map $J^1\tilde{X}$. Then, if $J^1\tilde{X}(x_0,\mu_0) = 0$, or equivalently $\tilde{X}(x_0,\mu_0) = 0$ and $\frac{\partial}{\partial x}\tilde{X}(x_0,\mu_0) = 0$, we know that the derivative map $DJ^1\tilde{X}(x_0,\mu_0)$ is a surjection onto $\mathbf{R}^2$. We compute

$$\begin{aligned} DJ^1\tilde{X}(x_0,\mu_0)\begin{pmatrix} v \\ w \end{pmatrix} &= \begin{pmatrix} \frac{\partial}{\partial x}\tilde{X}(x_0,\mu_0)v + \frac{\partial}{\partial \mu}\tilde{X}(x_0,\mu_0)w \\ \frac{\partial^2}{\partial x^2}\tilde{X}(x_0,\mu_0)v + \frac{\partial^2}{\partial x \partial \mu}\tilde{X}(x_0,\mu_0)w) \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial}{\partial \mu}\tilde{X}(x_0,\mu_0)w \\ \frac{\partial^2}{\partial x^2}\tilde{X}(x_0,\mu_0)v + \frac{\partial^2}{\partial x \partial \mu}\tilde{X}(x_0,\mu_0)w) \end{pmatrix} \end{aligned}$$

as $\frac{\partial}{\partial x}\tilde{X}(x_0,\mu_0) = 0$. This map can only be a surjection onto $\mathbf{R}^2$ if both of the non-degeneracy conditions $\frac{\partial}{\partial \mu}\tilde{X}(x_0,\mu_0) \neq 0$ and $\frac{\partial^2}{\partial x^2}\tilde{X}(x_0,\mu_0) \neq 0$ are satisfied. Therefore, degenerate critical points of the vectorfield $\tilde{X}(.,\mu)$ must satisfy these non-degeneracy conditions. This shows that this type of codimension-1 degenerate critical points are generic for 1-parameter families of vectorfields.

2.1 Tangent Bundle, Vectorfields

In the study of dynamical systems we regard vectorfields as differential equations whose flows define dynamical systems. In this chapter we are mostly concerned with the set of critical points of a given vectorfield and as such shall consider a vectorfield as a differentiable map between manifolds. While there is no difficulty in defining a dynamical system by its flow on a compact mainfold, it is not immediately clear how to define the associated vectorfield. Specifically, a vectorfield X on M is a smooth map from M to what manifold?

<u>Local Vectorfields</u>

In the case of a flow Φ defined on an open subset U of $\mathbf{R}^n$, a vectorfield X is given on U by the differential equation

$$\frac{d}{dt}\Phi(x,t) = X(\Phi(x,t))$$

or

$$X(x) = \frac{d}{dt}\Phi(x,t)|_{t=0}.$$

Then, X is a smooth function $X : U \to \mathbf{R}^n$ which we call a *local vectorfield.* Note that the value of X at a point x depends only on the tangency of the integral curve $\gamma_x(t) = \Phi(x,t)$ at $t = 0$. Indeed, one possible approach to extending the notion of tangent vectors and vectorfields to manifolds that is adoped by many texts (Abraham and Marsden [1978], Abraham *et. al.* [1983], Chillingworth [1976]) is to define tangent vectors to a manifold by the equivalence class of curves on the manifold that are tangent at a point. We shall follow a slightly different approach here in order to motivate the important concept of a *vector bundle.*

Transformation Rule for Tangent Vectors

First, we consider what happens to vectorfields under a change of coordinate system. We require that our vectorfield in the new coordinates has the same flow as the old vectorfield after changing coordinates. Precisely, let $X : U \to \mathbf{R^n}$ be a local vectorfield and $\varphi : U \to V$ be the diffeomorphism that gives our change of coordinates. Then, we require that our new vectorfield X' defined on V has the flow $\varphi \circ \Phi(x,t)$ where Φ is the flow of X on U. Thus

$$\begin{aligned} X'(\varphi \circ \Phi(x,t)) &= \frac{d}{dt}\varphi \circ \Phi(x,t), \\ &= D\varphi(\Phi(x,t)) \cdot \frac{d}{dt}\Phi(x,t), \\ &= D\varphi(\Phi(x,t)) \cdot X(\Phi(x,t)), \end{aligned}$$

or equivalently

$$X'(y) = D\varphi(\varphi^{-1}(y)) \cdot X(\varphi^{-1}(y)).$$

This is the transformation rule for vectorfields (indeed even vectors) under changes in coordinates.

Now suppose we have a flow Φ on a compact n-manifold M which has an atlas of charts $\{(U_\alpha, \varphi_\alpha)\}$. Via the charts, the flow Φ defines local flows Φ_α on open subsets of $\mathbf{R^n}$ by

$$\Phi_\alpha(x,t) = \varphi_\alpha \circ \Phi(\varphi_\alpha^{-1}(x),t).$$

As we have already seen, these local flows give rise to local vectorfields

$$X_\alpha : \mathbf{R}^n | \varphi_\alpha(U_\alpha) \to \mathbf{R}^n.$$

Now we will apply the "globalization process" which will "patch together" these local objects X_α into global object which will be a vectorfield on M.

By our transformation rule for vectorfields, we know that

$$X_\beta = D\varphi_{\beta\alpha}(\varphi_{\alpha\beta}(x)) \cdot X_\alpha(\varphi_{\alpha\beta}(x)),$$

where $\varphi_{\alpha\beta} = \varphi_\alpha \circ \varphi_\beta^{-1}$, $\varphi_{\beta\alpha} = \varphi_\beta \circ \varphi_\alpha^{-1}$ are the chart transition maps for our atlas of M.

Equivalence Relation

Consider the disjoint union

$$S = \bigcup_\alpha \{\alpha\} \times \varphi_\alpha(U_\alpha) \times \mathbf{R}^n.$$

Our transformation rule for vectors motivates the following equivalence relation on S;

$$\begin{aligned} &(\alpha, x, v) \sim (\beta, y, w) \\ \text{iff} \quad & x = \varphi_{\alpha\beta}(y) \\ \text{and} \quad & v = D\varphi_{\alpha\beta}(\varphi_{\beta\alpha}(x)) \cdot w, \end{aligned}$$

or equivalently

$$x = \varphi_{\alpha\beta}(y) \qquad \text{and} \qquad v = D\varphi_{\alpha\beta}(y) \cdot w.$$

We see that this does define an equivalence relation from the chain rule. Indeed, reflexivity follows as

$$x = \varphi_{\alpha\beta}(y) \iff y = \varphi_{\beta\alpha}(x)$$

and so

$$v = D\varphi_{\alpha\beta}(y) \cdot w \iff w = D\varphi_{\beta\alpha}(x) \cdot v$$

since

$$\begin{aligned} D\varphi_{\alpha\beta}(y) \cdot D\varphi_{\beta\alpha}(x) &= D\varphi_{\alpha\beta}(\varphi_{\beta\alpha}(x)) \cdot D\varphi_{\beta\alpha}(x) \\ &= D(\varphi_{\alpha\beta} \circ \varphi_{\beta\alpha})(x) \\ &= D(id) = id \end{aligned}$$

by the chain rule. Transitivity of the relation also follows from the chain rule. We have $\varphi_{\gamma\alpha} = \varphi_{\gamma\beta} \circ \varphi_{\beta\alpha}$, so that

$$D\varphi_{\gamma\alpha}(x) = D\varphi_{\gamma\beta}(\varphi_{\beta\alpha}(x)) \cdot D\varphi_{\beta\alpha}(x)$$

$\varphi_{\gamma\alpha} = \varphi_{\gamma\beta} \circ \varphi_{\beta\alpha}$. by the chain rule.

Tangent Bundle

We define the *tangent bundle* TM of the manifold M by

$$TM = S/\sim .$$

If M is a C^r manifold, then the tangent bundle has a C^{r-1} manifold structure given by an atlas on TM that is inherited from our atlas $\{(U_\alpha, \varphi_\alpha)\}$ on M as follows: Letting $[(\alpha, x, v)]$ denote the equivalence class of (α, x, v) under the relation $\sim$, the charts of this atlas are $(TU_\alpha, T\varphi_\alpha)$, where

$$TU_\alpha = \{[(\alpha, \varphi_\alpha(p), v)] : p \in U_\alpha, v \in \mathbf{R}^n\},$$

and

$$T\varphi_\alpha : [(\alpha, x, v)] \longmapsto (x, v).$$

The charts $(TU_\alpha, T\varphi_\alpha)$ of this atlas are called *vector bundle charts* for TM. The transition maps for this atlas are given by

$$\begin{aligned} T\varphi_{\alpha\beta} &: \mathbf{R}^n|\varphi_\beta(U_\beta) \times \mathbf{R}^n \longrightarrow \mathbf{R}^n|\varphi_\alpha(U_\alpha) \times \mathbf{R}^n \\ &: (x, v) \longmapsto (\varphi_{\alpha\beta}(x), D\varphi_{\alpha\beta}(x) \cdot v). \end{aligned}$$

It is easy to verify that these transition maps are C^{r-1} diffeomorphisms and that the so-called cocycle condition

$$T\varphi_{\alpha\gamma} = T\varphi_{\alpha\beta} \circ T\varphi_{\beta\gamma}$$

follows from the transitivity of our equivalence relation $\sim$. It is not at all suprising that the transition maps for this atlas on TM came directly from the definition of the equivalence relation $\sim$ when one considers that the manifold M can itself be defined as the disjoint union $\bigcup_\alpha\{\alpha\} \times \varphi_\alpha(U_\alpha)$ partitioned by the equivalence relation given by the chart transition maps; i.e. $(\alpha, x) \sim (\beta, y) \iff x = \varphi_{\alpha\beta}(y)$. This is essentially what the globalisation process is—the creation of a global object from its component local representations in different local coordinates, and noting that such a global object is defined by any set of local objects that is consistent with a transformation rule for the object in question.

Vector Bundles

The tangent bundle TM is the prototypical example of a *vector bundle.* That TM is a vector bundle means:

1. TM has a *local product structure* given by the vector bundle charts as

$$T\varphi_\alpha(TU_\alpha) = \mathbf{R}^n|\varphi_\alpha(U_\alpha) \times \mathbf{R}^n,$$

   so that TM is locally diffeomorphic to the product of a neighbourhood in M (U_α) and a linear vectorspace ($\mathbf{R}^n$).

2. The maps $T\varphi_\alpha$ are "inherited" from the chart maps φ_α, in that the first component of the map $T\varphi_\alpha$ is the chart map φ_α; i.e.,

$$T\varphi_\alpha([(\alpha, \varphi_\alpha(p), v)]) = (\varphi_\alpha(p), v),$$

   or, for any tangent vector $v_p \in T_pM$, (see below) we have

$$T\varphi_\alpha(v_p) = (\varphi_\alpha(p), \text{something} \in \mathbf{R}^n).$$

3. The chart overlap maps in the atlas of vector bundle charts are linear isomorphisms on the second factor; i.e. for fixed x,

$$T\varphi_{\alpha\beta}(x, \xi) = (\varphi_{\alpha\beta}(x), D\varphi_{\alpha\beta}(x) \cdot \xi)$$

is a linear isomorphism in the variable $\xi \in \mathbf{R}^n$ from the vectorspace $\{x\} \times \mathbf{R}^n$ into the vectorspace $\{\varphi_{\alpha\beta}(x)\} \times \mathbf{R}^n$. Also, the chart overlap map $\varphi_{\alpha\beta}$ is the first component of the overlap map $T\varphi_{\alpha\beta}$ as was hinted at in (2) above.

We see that the vector bundle charts preserve the vector space structure of the set

$$T_pM = \{[(\alpha, \varphi_\alpha(p), v] : v \in \mathbf{R}^n\}.$$

This set is called the *fiber of* TM *over* p. Clearly T_pM is isomorphic to $\mathbf{R}^n$. We may also define T_pM as

$$T_pM = \pi^{-1}(p),$$

where $\pi : TM \to M$ is the natural projection map of the tangent bundle given by

$$\pi([(\alpha, \varphi_\alpha(p), v)]) = p.$$

Items (2)—(3) above state that the vector bundle chart maps $T\varphi_\alpha$ satisfy

$$T\varphi_\alpha | T_pM = \varphi_\alpha(p),$$

and that this restricted map is an isomorphism from T_pM and $\mathbf{R}^n$.

Vectorfields, Sections

The set T_pM is also called the *tangent space to* M *at* p and can be thoughtof as the space of all *tangent vectors* to M which are based at p. A *vectorfield* X on M is a map which takes a point $p \in M$ to a tangent vector to M at p; i.e., $X(p) \in T_pM$. Equivalently, a vectorfield is any smooth map $X : M \to TM$ which is a *section* of the projection map π; i.e. that satisfies

$$(\pi \circ X)(p) = p, \quad \forall p \in M.$$

The space of all C^r-vectorfields on M, or sections of TM /sections of π is a linear space which we shall denote by $\mathcal{X}^r(M)$. In section 2.4 we shall define a topology of

$\mathcal{X}^r(M)$ that makes it a Banach space. Given that $\mathcal{X}^r(M)$ is a Banach space, we will be able to establish properties of the *evaluation map*

$$ev : \mathcal{X}^r(M) \times M \to TM : (X, p) \mapsto X(p) \in TM$$

that we shall use in our analysis of the bifurcations of the critical points of vector-fields on M.

2.2 Jet Bundles

In our study of bifurcations of critical points of vectorfields on 1-dimensional manifolds in section 1.3, we looked at the effect of various assumptions about the Taylor series of a vectorfield X about one of its critical points p on the changes that could occur in the set of critical points of the vectorfield and also in the local dynamics of the vectorfield under perturbation. In generalising this approach to consider the case of vectorfields on compact manifolds, we realize that the Taylor series of a vectorfield depends on the coordinate system in which it is expressed. In order to make the most consistent use of the tools of differential theory, we shall consider another example of a vector bundle, the *bundle of k-jets of vector fields on a compact manifold M*. In the same way that a tangent vector $v(p) \in T_pM$ is a coordinate independent object that is represented by its local representatives in a coordinate independant way, the k-jet of a vectorfield X at a point $p \in M$ is in essence a coordinate independant notion of the k-th order Taylor polynomial of any local representative X_α at the corresponding point $\varphi_\alpha(p) \in \mathbf{R}^n$. The bundle of k-jets of vectorfields arises as the globalisation of local (coordinate) definition of the k-th order Taylor polynomial of a local vectorfield at a point.

Transformation Rule for Taylor Polynomials

Recall that a vectorfield $X \in \chi^r(M)$ is a smooth map $X : M \to TM$. If $(TU_\alpha, T\varphi_\alpha)$ is a vectorbundle chart for TM, then

$$X|TU_\alpha : U_\alpha \longrightarrow TU_\alpha,$$

and the *local representative* X_α is defined using the chart maps $\varphi_\alpha, T\varphi_\alpha$ by

$$T\varphi_\alpha \circ X \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha) \subset \mathbf{R}^n \longrightarrow T\varphi_\alpha(TU_\alpha) = \varphi_\alpha(U_\alpha) \times \mathbf{R}^n,$$

and X_α is the second component of this map; i.e.,

$$T\varphi_\alpha \circ X \circ \varphi_\alpha^{-1}(x) = (x, X_\alpha(x)).$$

For a point $p \in U_\alpha$, the local vectorfield X_α has a k-th order Taylor polynomial at $x = \varphi_\alpha(p)$,

$$P^k X_\alpha(x,h) = X_\alpha(x) + DX_\alpha(x)\cdot h + \frac{1}{2!}D^2X_\alpha(x)\cdot(h,h) + \ldots + \frac{1}{k!}D^kX_\alpha(x)\cdot \underbrace{(h\ldots h)}_{\text{k times}}.$$

The *coefficients* of this polynomial map $P^kX_\alpha(x)$ are

$$X_\alpha(x), DX_\alpha(x), \frac{1}{2!}D^2X_\alpha(x), \ldots, \frac{1}{k!}D^kX_\alpha(x).$$

These coefficients lie in the vector space $P_S^k(\mathbf{R}^n)$ of symmetric k-th order polynomials on $\mathbf{R}^n$;

$$\begin{aligned}(X_\alpha(x), DX_\alpha(x), \ldots, \frac{1}{k!}D^kX_\alpha(x)) &\in \mathbf{R}^n \times L(\mathbf{R}^n) \times \ldots \times L_S^k(\mathbf{R}^n) \\ &\equiv P^k(\mathbf{R}^n),\end{aligned}$$

where $L_S^j(\mathbf{R}^n)$ denotes the space of symmetric j-fold multilinear maps from $\mathbf{R}^n$ to $\mathbf{R}^n$.

In order to discover the transformation rule for the Taylor polynomials P^kX_α under changes of coordinates, recall that

$$X_\beta(x) = D\varphi_{\beta\alpha}(\varphi_{\alpha\beta}(x)) \cdot X_\alpha(\varphi_{\alpha\beta}(x)).$$

To obtain the coefficients of $P^kX_\beta(x)$ in terms of the coefficients of $P^kX_\alpha(\varphi_{\alpha\beta}(x))$, we differentiate the transformation rule for vectors;

$$D^qX_\beta(x) = D^q\Big[D\varphi_{\beta\alpha}(\varphi_{\alpha\beta}(x)) \cdot X_\alpha(\varphi_{\alpha\beta}(x))\Big].$$

The composition $D\varphi_{\beta\alpha} \cdot X_\alpha$ is bilinear; Hence, we may apply Leibniz' Rule for bilinear maps, Abraham and Robbin [1967], p.3, which states that

$$D^q(\alpha \cdot \beta) = \sum_{0 \leq l \leq q} \binom{q}{l} D^l \alpha \cdot D^{q-l} \beta$$

whence

$$D^q X_\beta(x) = \sum_{0 \leq l \leq q} \binom{q}{l} D^l \Big[D\varphi_{\beta\alpha}(\varphi_{\alpha\beta(x)}) \Big] \cdot D^{q-l} \Big[X_\alpha(\varphi_{\alpha\beta}(x)) \Big].$$

In order to differentiate the terms $D\varphi_{\beta\alpha}(\varphi_{\alpha\beta}(x)), X_\alpha(\varphi_{\alpha\beta}(x)$, we employ the composite function rule, Abraham and Robbin [1967], p. 3, which states that

$$D^s(\alpha \circ \beta)(x) = \sum_{1 \leq j \leq s} \sum_{|i|=s} \sigma_s(i_1, \ldots, i_j) D^j \alpha(\beta(x)) \cdot \Big(D^{i_1}\beta(x), \ldots, D^{i_j}\beta(x) \Big),$$

where the $\sigma_s(i_1, \ldots, i_j)$ are constants obtained inductively in the proof of the result. Then we have

$$D^l \Big[D\varphi_{\beta\alpha}(\varphi_{\alpha\beta}(x)) \Big]$$

$$= \sum_{1 \leq j \leq l} \sum_{|i|=l} \sigma_l(i_1, \ldots, i_j) D^{j+1} \varphi_{\beta\alpha}(\varphi_{\alpha\beta}(x)) \cdot \Big(D^{i_1}\varphi_{\alpha\beta}(x), \ldots, D^{i_j}\varphi_{\alpha\beta}(x) \Big),$$

and

$$D^{q-l}\Big(X_\alpha(\varphi_{\alpha\beta}(x)) \Big)$$

$$\sum_{1 \leq m \leq q-l} \sum_{|n|=q-l} \sigma_{q-l}(n_1, \ldots, n_m) D^m X_\alpha(\varphi_{\alpha\beta}(x)) \cdot \Big(D^{n_1}\varphi_{\alpha\beta}(x), \ldots, D^{n_m}\varphi_{\alpha\beta}(x) \Big),$$

so that the full change of coordinate formula becomes:

$$\begin{aligned} D^q X_\beta(x) = & \\ & \sum_{0 \leq l \leq q} \sum_{1 \leq j \leq l} \sum_{1 \leq m \leq q-l} \sum_{|i|=l} \sum_{|n|=q-l} \binom{q}{l} \sigma_l(i_1, \ldots, i_j), \sigma_{q-l}(n_1, \ldots, n_m) \\ & \cdot D^{j+1}\varphi_{\beta\alpha}(\varphi_{\alpha\beta}(x)) \cdot \Big(D^{i_1}\varphi_{\alpha\beta}(x), \ldots, D^{i_j}\varphi_{\alpha\beta}(x) \Big) \\ & \cdot D^m X_\alpha(\varphi_{\alpha\beta}(x)) \cdot \Big(D^{n_1}\varphi_{\alpha\beta}(x), \ldots, D^{n_m}\varphi_{\alpha\beta}(x) \Big). \end{aligned}$$

This transformation rule expresses the derivative $D^q X_\beta(x)$, for $0 \leq q \leq k$, in terms of $X_\alpha(\varphi_{\alpha\beta}(x))$, the old coordinate representation of X and its first q derivatives $DX_\alpha(\varphi_{\alpha\beta}(x), \ldots, D^q X_\alpha(\varphi_{\alpha\beta}(x))$, at the point $\varphi_{\alpha\beta}(x)$ that corresponds to x in the old coordinates φ_α. It is interesting to note that the transformation rule for the $k-th$ derivative of a vectorfield depends on the first $k+1$ derivatives of the chart transition maps. This is because the trasition maps for vectorfields involve the derivative of the chart transition maps, so that in order to consider C^r vectorfields, the manifold M must be at least C^{r+1}.

Bundle of k-Jets

Given the change of coordinate formulas for the derivatives of a vectorfield, we know how the Taylor series transforms under changes of coordinates. Unfortunately, the transformation rules are rather unwieldy, and as such we will not proceed as we did in defining the tangent bundle. Instead, we use the following definition;

Definition. *Let $X, Y \in \mathcal{X}(M)$, $p \in U_\alpha \subset M$, where $(U_\alpha, \varphi_\alpha)$ is a chart on M. We say that X and Y have the same k-jet at p if the local representatives X_α and Y_α have the same k-th order Taylor polynomial at the point $\varphi_\alpha(p)$; i.e.,*

$$\text{if and only if } \; P^k X_\alpha(\varphi_\alpha(p)) = P^k Y_\alpha(\varphi_\alpha(p)).$$

The k-jet of a vectorfield at a point p is the eqivalence class of vectorfields having the same k-jet at p.

From our observations about the transformation rule for the Taylor series of a vectorfield, it follows that this definition does not depend on the choice of chart $(U_\alpha, \varphi_\alpha)$. Indeed, if local representatives X_α and Y_α have identical k-th order Taylor polynomials at a point $\varphi_\alpha(p)$, then in new coordinates φ_β, the coefficients of the k-th order Taylor polynomials of the new local representatives X_β and Y_β at the point $\varphi_\beta(p)$ may be expressed in terms of the coefficients of the k-th order Taylor polynomials of X_α and Y_α and so are also equal.

The k-jet of the vectorfield X at the point p is denoted by $j^kX(p)$. The set of all k-jets of vectorfields at a given point p forms a vector space which we denote by $J_p^k(TM)$. Indeed, the vector space structure is given in the way we would expect, with $\lambda \cdot j^kX(p) = j^k(\lambda X)(p)$, and $j^kX(p) + j^kY(p) = j^k(X+Y)(p)$. Letting

$$J^k(TM) = \cup_{p\in M} J_p^k(TM),$$

we may define a vector bundle structure over M as follows: Given a chart $(U_\alpha, \varphi_\alpha)$ for M, the associated vector bundle chart $(J^kU_\alpha, J^k\varphi_\alpha)$ is defined by:

$$J^kU_\alpha = \pi_k^{-1}(U_\alpha)$$

where $\pi_k : J^k(TM) \to M$ is the natural projection map given by $\pi(J_p^k(TM)) = p$. Then J^kU_α is the set of k-jets of vectorfields at points $p \in U_\alpha$. We define the chart maps $J^k\varphi_\alpha$ for the bundle by

$$\begin{aligned} J^k\varphi_\alpha : J^k(TM)|J^kU_\alpha &\longrightarrow \mathbf{R}^n \times P^k(\mathbf{R}^n) \\ : j^kX(P) &\longmapsto (\varphi_\alpha(p), P^kX_\alpha(\varphi_\alpha(p))). \end{aligned}$$

In other words, the chart map assigns to a k-jet $j^kX(p)$ the k-th order Taylor polynomial of the local representative of the vectorfield X (or any vectorfield with the k-jet $j^kX(p)$ at p).

In order to show that the charts $(J^kU_\alpha, J^k\varphi_\alpha)$ define a vector bundle structure on $J^k(TM)$ it suffices to note:

1. $J^k(TM)$ has a local product structure given by the chart maps. This is evident as

   $$J^k\varphi_\alpha(J^kU_\alpha) = \mathbf{R}^n|\varphi_\alpha(U_\alpha) \times P_S^k(\mathbf{R}^n).$$

   This map is surjective, for if $Q_k \in P^k(\mathbf{R}^n)$ is an arbitrary k-th order symmentric polynomial, we may define a vectorfield X' such that its local representative X'_α has Q_k for its k-th order Taylor polynomial at a point $\varphi_\alpha(p)$.

2. The induced chart transition maps $J^k\varphi_{\alpha\beta} = J^k\varphi_\alpha \circ J^k\varphi_\beta^{-1}$,

$$J^k\varphi_{\alpha\beta} : \mathbf{R}^n|\varphi_\beta(U_\beta) \times P_S^k(\mathbf{R}^n) \longrightarrow \mathbf{R}^n|\varphi_\alpha(U_\alpha) \times P_S^k(\mathbf{R}^n),$$

are linear isomorphisms of the space $P_S^k(\mathbf{R}^n)$ for fixed $x \in \varphi_\beta(U_\beta)$. This follows from looking at the transformation rule we obtained for the first k derivatives of a vectorfield under changes of coordinates. The maps we obtained were linear in the derivatives of the local representatives $X_\alpha(\varphi_{\alpha\beta}(x)), \ldots, D^k X_\alpha(\varphi_{\alpha\beta}(x)))$. These maps must also be isomorphisms as the map $J^k\varphi_{\alpha\beta}$ has an inverse given by $J^k\varphi_{\beta\alpha}$.

As we have already noted, the chart transistion map $J^k\varphi_{\alpha\beta}$ depends on the first k+1 derivatives of the chart transistion maps $\varphi_{\alpha\beta}$ of M. Thus, for a C^r manifold M, we have that the bundle of k-jets of vectorfields is a C^{r-k-1} vector bundle.

3.1 The Banach Space $\mathcal{X}^r(M)$.

In the example of a bifurcation in Chapter 1, the vectorfield X depended on a scalar parameter μ and so was a map from $\mathbf{R}^1 \times \mathbf{R}^1$ to $\mathbf{R}^1$. We analysed the bifurcation associated with a given critical point (x_0, μ_0) of X by looking at the geometry of the the set of critical points of X near (x_0, μ_0). This analysis depended heavily on the fact that X was a differentiable map, and as such we could apply several results from differential theory. In generalising this approach to bifurcation, we do not use a particular parameterised family; instead we think of a bifurcation as being associated with a particular vectorfield and wish to consider the possible changes in the set of critical points (or other parts of the dynamics) that occur when the system is perturbed.

The spaces $B^r(\varphi_\alpha(U_\alpha); \mathbf{R}^n)$

Consider a finite collection of charts $\{(U_\alpha, \varphi_\alpha)\}_{\alpha=1}^N$ that cover M. For a fixed α, we have the map that takes a vectorfield $X \in \mathcal{X}^r(M)$ to its local representation

$$X_\alpha : \mathbf{R}^n | \varphi_\alpha(U_\alpha) \longrightarrow \mathbf{R}^n,$$

defined through the vector bundle charts $(TU_\alpha, T\varphi_\alpha)$ by

$$X_\alpha(\varphi_\alpha(p)) = \text{second component of } T\varphi_\alpha(X(p)),$$

since we have

$$T\varphi_\alpha(X(p)) = (\varphi_\alpha(p), X_\alpha(\varphi_\alpha(p)).$$

This map is a surjection

$$\mathcal{X}^r(M) \longrightarrow B^r(\varphi_\alpha(U_\alpha);\mathbf{R}^n),$$

where $B^r(\varphi_\alpha(U_\alpha);\mathbf{R}^n)$ is the space of C^r maps from the open subset (with compact closure) $\varphi_\alpha(U_\alpha)$ of $\mathbf{R}^n$ into $\mathbf{R}^n$ which are bounded in the C^r-norm

$$\|X_\alpha\|_r = \sup_{x\in\varphi_\alpha(U_\alpha)} \left\{|X_\alpha(x)| + |DX_\alpha(x)| + \ldots + |D^r X_\alpha(x)|\right\}.$$

We shall make use of the following well known result.

Lemma. *Let $U \subset \mathbf{R}^n$ be an open set. Then the space $B^r(U,\mathbf{R}^n)$ with the norm $\|\cdot\|_r$ above is a Banach space.*

Proof. Clearly $B^r(U;\mathbf{R}^n)$ is a vector space under pointwise addition and scalar multiplication of fuctions, and $\|\cdot\|_r$ is a norm on $B^r(U;\mathbf{R}^n)$. We must show that $B^r(U;\mathbf{R}^n)$ is complete in this norm.

Let $\{X_n\} \subset B^r(U;\mathbf{R}^n)$ be a Cauchy sequence in the norm $\|\cdot\|_r$. Then, since the convergence is uniform,

$$X_n \to \overline{X}$$

$$\text{and} \qquad D^q X_n \to \overline{X}^q$$

for some continuous functions $\overline{X}, \overline{X}^1, \ldots, \overline{X}^r$ on U. Clearly, the $\overline{X}, \overline{X}^q$ are all uniformly bounded on U. It remains for us to show that $D^q\overline{X}(x) = \overline{X}^q(x)$ for $q = 1, 2, \ldots, r$ and for $x \in U$. Let us first show that $D\overline{X}(x)\cdot v = \overline{X}^1(x)\cdot v$. This entails

$$\lim_{t\to 0}\left|\frac{\overline{X}(x+tv) - \overline{X}(x) - \overline{X}^1(x)\cdot tv}{t}\right| = 0.$$

But $X_n \to \overline{X}$ so that this limit becomes

$$\lim_{t\to 0}\lim_{n\to\infty}\left|\frac{X_n(x+tv) - X_n(x) - \overline{X}^1(x)\cdot tv}{t}\right|.$$

But, by the mean value theorem

$$\left|\frac{X_n(x+tv) - X_n(x) - \overline{X}^1(x)\cdot tv}{t}\right| = |DX_n(x+\xi v)\cdot v - \overline{X}^1(x)\cdot v|$$

$$= \left| DX_n(x+\xi v)\cdot v - \overline{X}^1(x+\xi v)\cdot v + \overline{X}(x+\xi v)\cdot v - \overline{X}^1(x)\cdot v \right|$$
$$\leq \left| DX_n(x+\xi v)\cdot v - \overline{X}^1(x+\xi v)\cdot v \right| + \left| \overline{X}^1(x+\xi v)\cdot v - \overline{X}^1(x)\cdot v \right|$$

for some $\xi in (0,t)$. In the last expression the first term goes to zero in n uniformly in t, and the second term goes to zero in t as $\overline{X}^1$ is continuous. This proves that $D\overline{X} = \overline{X}^1$. By induction, we see that $D^q\overline{X} = \overline{X}^q$ for $q = 1,\ldots,r$ and the lemma is proved.

Constructing $\mathcal{X}^r(M)$

Now, consider the direct sum of these Banach spaces

$$\mathcal{X} = \oplus_{\alpha=1}^{N} B^r(\varphi_\alpha(U_\alpha);\mathbf{R}^n),$$

with the usual norm that makes $\mathcal{X}$ into a Banach space;

$$\|X_1 \oplus \ldots \oplus X_N\|_r = \|X_1\|_r + \ldots + \|X_N\|_r.$$

We shall show that $\mathcal{X}^r(M)$ is isomorphic to a closed subspace of $\mathcal{X}$, whence $\mathcal{X}^r(M)$ is itself a Banach space.

Clearly the map

$$\mathcal{X}^r(M) \longrightarrow \mathcal{X} : X \longmapsto (X_1,\ldots,X_N)$$

is an injection (here the X_α are the local representatives of X in the α-th coordinate chart on M.). The image of this map is the subspace of $\mathcal{X}$ defined by

$$\left\{(X_1,\ldots,X_N) \in \mathcal{X} : \tilde{X}_\alpha = T\varphi_{\alpha\beta} \circ \tilde{X}_\beta \right\},$$

where $T\varphi_{\alpha\beta}$ are the vector bundle chart transition maps for TM and the $\tilde{X}_\alpha = id_{U_\alpha} \times X_\alpha$, etc... In other words, the image of $\mathcal{X}^r(M)$ under this map is the set of collections of local representatives that are consistent with the transformation rule for vectorfields. This is exactly what we have already seen in section 2.1. To see that this is really a subspace, it suffices to note that the transition maps are linear in the X_β. Also, this is a closed subspace. Indeed, if $(X_1^m, \ldots, X_N^m)_{m=1}^{\infty}$ is a Cauchy sequence in $\mathcal{X}$ satisfying

$$X_\alpha^m(x) = D\varphi_{\alpha\beta}(\varphi_{\beta\alpha}(x)) \cdot X_\beta^m(\varphi_{\beta\alpha}(x)),$$

then

$$\begin{aligned}\lim_{m\to\infty} X_\alpha^m &= \lim_{m\to\infty} D\varphi_{\alpha\beta}(\varphi_{\beta\alpha}(x)) \cdot X_\beta^m(\varphi_{\beta\alpha}(x)) \\ &= D\varphi_{\alpha\beta}(\varphi_{\beta\alpha}(x)) \cdot \lim_{m\to\infty} X_\beta^m(\varphi_{\beta\alpha}(x)).\end{aligned}$$

Thus, $\mathcal{X}^r(M) \subset \mathcal{X}$ has closed range and so $\mathcal{X}^r(M)$ is isomorphic to a closed subspace of the Banach space $\mathcal{X}$. Then $\mathcal{X}^r(M)$ inherits a topology from this embedding that makes it a Banach space. This topology on $\mathcal{X}^r(M)$ is the ***topology of uniform C^r-convergence on compacta, which is the same as the topology of uniform C^r-convergence since M is compact.***

3.2 Differentiability of the Evaluation Map

In this section we shall show that the evaluation map for vectorfields,

$$ev : \mathcal{X}^r(M) \times M \longrightarrow TM,$$

$$(X, p) \longmapsto X(p),$$

is a C^r – map and we will compute a formula for the derivatives of this map. In addition, we shall show that the derivative $Dev(X,p)$ at the point (X,p) is *split-surjective*; that is, it is surjective and its kernel $ker(Dev(X,p))$ splits in the Banach space $T_{(X,p)}(\mathcal{X}^r(M) \times M)$.

The purpose of showing these properties of the evaluation map, is that we may then apply the results for differentiable maps that we will obtain in the next chapter, which will yield some results that are at the core of generic bifurcation theory. The results that obtain from the study of the evaluation map itself are primarily useful for the consideration of bifurcations that involve only the critical points of a vectorfield. However, it is possible to consider the more general relatives of the evaluation map in this same framework, and obtain similar results for periodic orbits and the like. The power of these so-called evaluation-transversality techniques is in reducing different kinds of bifurcation questions to questions about the geometry of certain submanifolds of $\mathcal{X}^r(M) \times M$. In chapter 5, we shall consider the example of the saddle-node bifurcation in detail in terms of this framework, as well as indicating how we might generalise the approach taken there so that it would include other bifurcations of critical elements and connections of critical elements.

The Derivatives of *ev*

Let us begin by formally differentiating the evaluation map. We have

$$ev : \mathcal{X}^r(M) \times M \to TM$$

$$ev : (X, p) \longmapsto X(p) \in T_pM.$$

The derivative of this map will be a globalisation of the derivative of the map ev_α induced by local coordinates φ_α, so it suffices to consider the derivative of this map. We have

$$ev_\alpha : \mathcal{X}^r(M) \times \mathbf{R}^n|\varphi_\alpha(U_\alpha) \longrightarrow \varphi_\alpha(U_\alpha) \times \mathbf{R}^n = T\varphi_\alpha(TM|U_\alpha),$$

where this is defined by

$$ev_\alpha = T\varphi_\alpha \circ ev \circ (id_{\mathcal{X}^r(M)} \times \varphi_\alpha)^{-1}.$$

Now, $ev_\alpha(X, x) = X_\alpha(x)$, where X_α is the induced local representative of X in local coordinates φ_α. In order to differentiate this map formally, we consider that

$$Dev_\alpha(X,p) \cdot (\xi, v) = Dev_\alpha(X,p) \cdot (\xi, 0) + Dev_\alpha(X,p) \cdot (0, v).$$

Each of these partial derivatives is easy to calculate. First, since for a fixed x we have that ev_α is the linear functional $X \mapsto X_\alpha(x)$ in the vectorfield X, we have that $Dev_\alpha(X,x) \cdot (\xi, 0) = \xi_\alpha(x)$. For the other partial derivative, we notice that

$$\begin{aligned} Dev_\alpha(X,x)(0,v) &\sim X_\alpha(x+v) - X_\alpha(x) \\ &\sim DX_\alpha(x) \cdot v, \end{aligned}$$

so that we would guess that the derivative of ev_α is

$$Dev_\alpha(X,x) \cdot (\xi, v) = \xi_\alpha(x) + DX_\alpha(x) \cdot v.$$

That this is in fact the derivative is trivial to verify. We consider

$$\begin{aligned} ev_\alpha((X,x) + (\xi, v)) - ev_\alpha(X,x) &= X_\alpha(x+v) + \xi_\alpha(x+v) - X_\alpha(x) \\ &= \xi_\alpha(x+v) + X_\alpha(x+v) - X_\alpha(x). \end{aligned}$$

Now

$$|\xi_\alpha(x+v) + X_\alpha(x+v) - X_\alpha(x) - (\xi_\alpha(x) + DX_\alpha(x)\cdot v)| \leq$$

$$|\xi_\alpha(x+v) - \xi_\alpha(x)| + |X_\alpha(x+v) - X_\alpha(x) - DX_\alpha(x)\cdot v|,$$

and the second term clearly goes to zero faster than $\|\xi\|_r + |v|$ as X is differentiable. The first term goes to zero as ξ_α is continuous. Furthermore, by the mean value theorem, we have that

$$|\xi_\alpha(x+v) - \xi_\alpha(x)| = |D\xi_\alpha(c)\cdot v|$$

for some c between x and $x+v$. So, since ξ_α is bounded in the C^r-norm, we have

$$|\xi_\alpha(x+v) - \xi_\alpha(x)| \leq \|\xi_\alpha\|_r \cdot |v|$$

and we know that this goes to zero like $(\|\xi_\alpha\|_r + |v|)^2$ from the inequality $|xy| \leq \frac{1}{2}(|x| + |y|)^2$. Thus the above function is Dev_α by the definition of derivative.

Now, let us consider the higher derivatives of ev_α. We consider the map

$$(X, x) \longmapsto \xi^1_\alpha(x) + DX_\alpha(x)\cdot v^1.$$

Then, $D^2 ev_\alpha(X,x)\cdot((\xi^1, v^1), (\xi^2, v^2))$ is just the derivative of the above map in the direction (ξ^2, v^2). Again, we compute the partial derivatives of this map. We see that

$$\text{derivative in } (0, v^2) \text{ direction} = D\xi^1_\alpha(x)\cdot v^2 + D^2X_\alpha(x)\cdot(v^1, v^2),$$

and that

$$\text{derivative in } (\xi^2, 0) \text{ direction} = D\xi^2_\alpha(x)\cdot v^1,$$

since the the part of the map that depends on X is a linear functional in X. So, we would guess that the formula for $D^2 ev_\alpha(x)$ is given by

$$D^2 ev_\alpha(x)\cdot(\xi^1, v^1, \xi^2, v^2) = D\xi^1_\alpha(x)\cdot v^2 + D\xi^2_\alpha(x)\cdot v^1 + D^2X_\alpha(x)\cdot(v^1, v^2).$$

Continuing in this fashion, we arrive at the following formula for the first r derivatives of ev_α,

$$D^p ev_\alpha(x)\cdot(\xi^1,v^1,\ldots,\xi^p,v^p) = D^pX_\alpha(x)\cdot(v^1,\ldots,v^p) + \sum_{i=1}^{p} D^{p-1}\xi_\alpha^i(x)\cdot(v^1,\ldots,\tilde{v}^i,\ldots,v^p),$$

where the notation $(v^1,\ldots,\tilde{v}^i,\ldots,v^p)$ is used for the (p-1)-tuple that does not contain v^i.

We proceed to prove this formula by induction. Let us compute the (p+1)st derivative of ev_α by taking the derivative of the above formula in the direction (ξ^{p+1},v^{p+1}). We have

$$\begin{aligned}
&D^{p+1}ev_\alpha(X,x)\cdot(\xi^1,v^1,\ldots,\xi^{p+1},v^{p+1})\\
&\qquad = D\Big(D^p ev_\alpha(X,x)\cdot(\xi^1,v^1,\ldots,\xi^p,v^p)\Big)\cdot(\xi^{p+1},v^{p+1})\\
&\qquad = D(\Phi^p(X,x))\cdot(\xi^p,v^p),
\end{aligned}$$

where Φ^p denotes our formula for the p-th derivative,

$$\Phi^p(X,x) = D^pX_\alpha(x)\cdot(v^1,\ldots,v^p) + \sum_{i=1}^{p} D^{p-1}\xi_\alpha(x)\cdot(v^1,\ldots,\tilde{v}^i,\ldots,v^p)\cdot(\xi^{p+1},v^{p+1}).$$

As before, we compute,

$$\begin{aligned}
&\left|\Phi(X+\xi^{p+1},x+v^{p+1})-\Phi(X,x)\right| \le\\
&\quad \left|\Phi(X+\xi^{p+1},x+v^{p+1})-\Phi(X,x+v^{p+1})\right| +\\
&\quad \left|\Phi(X,x+v^{p+1})-\Phi(X,x)\right|
\end{aligned}$$

and

$$\begin{aligned}
&\le \left|(D^p\xi_\alpha^p(x)\cdot(v^1,\ldots,v^p)\right| + \left|\Big(D^pX_\alpha(x+v^{p+1})-D^pX_\alpha(x)\Big)\cdot(v^1,\ldots,v^p)\right| +\\
&\quad \left|\sum_{i=1}^{p}\Big(D^{p-1}\xi_\alpha(x+v^{p+1})-D^{p-1}\xi_\alpha(x)\Big)\cdot(v^1,\ldots,\tilde{v}^i,\ldots,v^p)\right|,
\end{aligned}$$

Considering each of the terms in the second part above separately, we see that $(D^pX_\alpha(x+v^{p+1})-D^pX_\alpha(x))$, and each of the terms $(D^{p-1}\xi_\alpha(x+v^{p+1})-D^{p-1}\xi_\alpha(x))$ are approximated by their derivatives, $D^{p+1}X_\alpha(x)\cdot v^{p+1}$ and $D^p\xi_\alpha(x)\cdot v^p$. As long as $p+1 \leq r$, these derivatives will exist as X and the ξ^i's are all C^r. We know that

$$\|D^p X_\alpha(x+v^{p+1}) - D^p X_\alpha(x) - D^{p+1} X_\alpha(x) \cdot v^{p+1}\| \to 0$$

faster than $|v^{p+1}|$, and hence faster than $\|\xi_\alpha\|_r + |v^{p+1}|$, by the definition of differentiability and the fact that X is C^r. Also, each of the terms

$$|D^{p-1}\xi_\alpha(x+v^{p+1}) - D^{p-1}\xi_\alpha(x) - D^p\xi_\alpha(x) \cdot v^{p+1}|$$

goes to zero faster than $\sup_x \|D^{p+1}\xi_\alpha(x)\| \cdot |v^{p+1}| \leq \|\xi_\alpha\|_r \cdot |v^{p+1}|$ by the mean value theorem, since the norm here is uniform. This shows that the formula for the p-th derivative of the evaluation map is

$$D^p ev_\alpha(X,x)\cdot(\xi^1, v^1, \ldots, \xi^p, v^p) =$$

$$D^p X_\alpha(x) \cdot (v^1, \ldots, v^p) + \sum_{i=1}^{p} D^{p-1}\xi_\alpha(x) \cdot (v^1, \ldots, \tilde{v}^i, \ldots, v^p).$$

Since all of the derivatives in this expression exist and are continuous for $p \leq r$, we know that the local representative ev_α of the evaluation map ev is C^r and hence the evaluation map itself is a C^r map from the Banach manifold $\mathcal{X}^r(M) \times M$ to TM.

Split Surjectivity of ev

The implicit function theorem from advanced calculus is usually stated for a C^1-function $f : \mathbf{R}^m \to \mathbf{R}^n$, such that the derivative in the first n-coordinates $D_1 f(x_0, y_0)$ has maximal rank n. Then there is a unique implicit function h such that $f(h(y), y) = f(x_0, y_0)$ for y near y_0. To generalize this theorem to the case of a function between Banach spaces, $f : \mathbf{E} \to \mathbf{F}$, we must replace the assumption of maximal rank with an appropriate generalisation, namely that $Df(x_0)$ is surjective at the point x_0. Additionally, we must assume that the kernel of $Df(x_0)$ *splits* in $\mathbf{E}$, that is, that there is a direct sum of closed subspaces $\mathbf{E} = ker(Df(x_0)) \oplus K'$. This is necessary for the decomposition of $\mathbf{E}$ into a direct sum of two components so that an implicit function can be expressed as a map from one component to the other.

In chapter 4, we will look at the implicit function theorem and some of its global generalisations in the Banach space/Banach manifold setting with the intention of applying these results to the evaluation map. For this reason, the remainder of this section is devoted to showing that the derivative of the evaluation map is surjective and kernel-splitting.

Consider the local representative ev_α of the evaluation map, and its derivative,

$$Dev_\alpha(X,x)\cdot(\xi,v) = \xi_\alpha(x) + DX_\alpha(x)\cdot v.$$

clearly this map is surjective onto $\mathbf{R}^n$ as we may have an arbitrary value for $\xi_\alpha(x)$. To show that the kernel of this map splits in $\mathcal{X}^r(M)\times\mathbf{R}^n$, we consider the subspaces

$$\begin{aligned}
\mathbf{K}_1 &= \{\xi : \xi_\alpha(x) = 0\} \times ker(DX_\alpha(x)),\\
\mathbf{K}_2 &= \{(\xi,v) : \xi_\alpha(x) \neq 0 \text{ and } \xi_\alpha(x) + DX_\alpha(x)\cdot v = 0\},\\
\mathbf{K}_3 &= \{\xi : \xi_\alpha(x) = 0\} \times \mathbf{K}',\\
\mathbf{K}_4 &= \{\xi : \xi_\alpha(x) \neq 0\} \times ker(DX_\alpha(x)),
\end{aligned}$$

where $\mathbf{K}'$ is a complement of $ker(DX_\alpha(x))$. The $\mathbf{K}_i$ are all closed subspaces of $\mathcal{X}^r(M)\times\mathbf{R}^n$ and it is easy to see that

$$\mathcal{X}^r(M)\times\mathbf{R}^n = \mathbf{K}_1 \oplus \mathbf{K}_2 \oplus \mathbf{K}_3 \oplus \mathbf{K}_4$$

and that $ker(Dev_\alpha(X,x)) = \mathbf{K}_1 \oplus \mathbf{K}_2$, and so is complemented. Thus ev is a kernel-splitting submersion.

4.1 The Implicit Function Theorem and Transversality

In this section we consider the implicit function theorem for smooth maps of Banach spaces and Banach submanifolds. Viewed geometrically, the implicit function theorem gives us conditions under which the inverse image, or pull back, of a point under a smooth map is locally a smooth submanifold of the domain. Introducing the definition of transversality of maps to submanifolds allows us to extend these results to the pull-backs of embedded submanifolds. In the next chapter, these results will be applied to the evaluation maps of Chapter 3 for the purpose of studying the relationships between the dependence of critical points of a vectorfield on perturbations of the vectorfield and the jets of the vectorfield at its critical points.

Implicit Function Theorem

We state a version of the implicit function theorem for C^r- maps of Banach spaces. The statement and proof of this theorem is found in [Abraham *et. al.* [1983],p.107]; However, the statement there omits a necessary condition for the uniqueness of the implicit function.

Theorem (implicit function theorem). *Let* $U \subset \mathbf{E}, V \subset \mathbf{F}$ *be open subsets of the Banach spaces* $\mathbf{E}, \mathbf{F}$, *and let* $f : U \times V \to \mathbf{G}$ *be* $C^r, (r \geq 1)$, *into the Banach space* $\mathbf{G}$. *For some* $(x_0, y_0) \in U \times V$ *assume that* $f(x_0, y_0) = w_0$, *and that*

$$D_2 f(x_0, y_0) : \mathbf{F} \longrightarrow \mathbf{G}$$

is an isomorphism. Then, there exist neighborhoods U_0 of x_o and W_0 of w_0 and a unique C^r- map $g : U_0 \times W_0 \to V$ such that

$$(i) \quad g(x_0, w_0) = y_0$$

$$(ii) \quad f(x, g(x, w)) = w \text{ for all } x, w \in U_0 \times W_0.$$

The content of this theorem is essentially geometrical. It states that the part of the inverse image $f^{-1}(w_0)$ that passes through the point (x_0, y_0) is locally given as the graph of a C^r- function $g_{w_0}(x) = g(x, w_0)$. This means that near (x_0, y_0), the *pull-back* $f^{-1}(w_0)$ of w_0 is a submanifold of $\mathbf{E} \times \mathbf{F}$. Furthermore, we can compute the tangent space of this submanifold at (x_0, y_0) by implicit differentiation. Indeed, since $f(x, g_{w_0}(x)) = w_0$, then

$$D_1 f(x_0, y_0) + D_2 f(x_0, y_0) \cdot D_1 g_{w_0}(x_0) = 0$$

which implies

$$D_1 g_{w_0}(x_0) = -\Big(D_2 f(x_0, y_0)\Big)^{-1} \cdot \Big(D_1 f(x_0, y_0)\Big)$$

since $D_2 f(x_0, y_0)$ is an isomorphism. Thus the tangent space $T_{(x_0,y_0)}\Big(f^{-1}(w_0)\Big)$ is of the form $\Big\{(\xi, D_1 g_{w_0}(x_0) \cdot \xi) : \xi \in \mathbf{E_1}\Big\}$.

Kernel Splitting Submersions and Regular Values

Often we are interested in the preimage/pull-back of a point p by a map f where $f : U \subset \mathbf{E} \to \mathbf{F}$ is a C^r-map from an open set U in a Banach space $\mathbf{E}$ into another Banach space $\mathbf{F}$. We may reduce this to the case of the implicit function theorem setting if f is locally a *kernel splitting submersion.* We see this in the following corollary;

Corollary. *Let $f : U \subset \mathbf{E} \to \mathbf{F}$ be $C^r, (r \geq 1)$, defined on the open set U. Assume that for some $u_0 \in U$, we have $f(u_0) = w_0$ and that $Df(u_0)$ is surjective and $\mathbf{E}_1 = ker(Df(u_0))$ splits in $\mathbf{E}$. Then $\mathbf{E} = \mathbf{E}_1 \oplus \mathbf{E}_2$ and there exist neighborhoods*

U_1, U_2 in $\mathbf{E}_1, \mathbf{E}_2$ with $U_1 \oplus U_2 \subset U$ and such that $f^{-1}(w_0) \cap (U_1 \oplus U_2)$ is a submanifold given by the graph of a C^r- function $g : \mathbf{E}_2 \to \mathbf{E}_1$. Furthermore $f^{-1}(p)$ is tangent to $ker(Df(u_0))$ at u_0.

Proof. Since $\mathbf{E}_1 = ker(Df(u_0))$ splits, then there exists a closed complement $\mathbf{E}_2$ to $\mathbf{E}_1$ in $\mathbf{E}$, whence $\mathbf{E} = \mathbf{E}_1 \oplus \mathbf{E}_2$. As $Df(u_0)$ is surjective, then $Df(u_0)|\mathbf{E}_2$ is an isomorphism from $\mathbf{E}_2$ to $\mathbf{F}$. Thus, the conditions of the implicit function theorem are satisfied for the function $\tilde{f}(x,y) = f(x+y)$ on $\mathbf{E}_1 \times \mathbf{E}_2$, since $D_2\tilde{f}(x_0, y_0) = Df(u_0)|\mathbf{E}_2$ where $u_0 = x_0 + y_0$. We can then infer the existence of a unique C^r- function $g : \mathbf{E}_2 \to \mathbf{E}_1$ such that $g(y_0) = x_0$, and $\tilde{f}(g(y), y) = w_0$ for y in some neighborhood of y_0. Also, as shown previously, $D_2g(y_0) = -\left(D_2\tilde{f}(x_0, y_0)\right)^{-1} \cdot D_1\tilde{f}(x_0, y_0)$, which is zero since $D_1\tilde{f}(x_0, y_0) = Df(u_0)|\mathbf{E}_1 = 0$ as $\mathbf{E} = ker(Df(u_0))$.

The generalization of this result to maps between Banach manifolds is immediate;

Corollary. *Let $f : M \to N$ be a C^r-map of Banach manifolds with $f(p) = q$. Assume that f is a kernel splitting submersion at p. Then there exists a neighborhood U_p of p such that $f^{-1}(q)$ is a C^r-submanifold tangent to $ker(Df(p))$ at p.*

Proof. Introducing local coordinates φ_α at p and ψ_β at q gives a local representative f_α^β that satisfies the hypotheses of the previous corollary.

A point $q \in N$ is a ***regular value*** for a C^r-map $f : M \to N$ if for each point $p \in f^{-1}(q)$, f is a kernel-splitting submersion at p. It is evident from the previous corollary that the inverse image/pull- back of a regular value is a C^r-manifold.

Pull-Backs of Submanifolds, Transversality

A point q of a manifold N is a particular case of a submanifold of N. We now consider the pull-back of a submanifold S of a Banach manifold N via a C^r-map.

Recall that S is an *embedded submanifold* of the C^r-manifold N if for each point $q \in S$ there is a chart (V_β, ψ_β) about q *in the atlas of* N that has the *submanifold property*

$$\varphi_\beta : V_\beta \to \mathbf{E},$$

$$\text{and } \varphi_\beta(V_\beta \cap S) = \mathbf{E_1} \times \{0\} \subset \mathbf{E},$$

where $\mathbf{E_1}$ is a subspace that splits in $\mathbf{E}$. Then S inherits a manifold structure from N with chart maps taking values in $\mathbf{E}_1$. If $\mathbf{E}_1$ is a finite dimensional subspace of dimension k, then S is clearly an n-dimensional (sub)manifold. However, if $\mathbf{E}_1$ has a closed complement $\mathbf{E}_2$ of finite dimension k, then we say that S is a *submanifold of codimension-k.*

It is clear from the above that any submanifold of codimension-k can be locally expressed as $S \cap W = \lambda^{-1}(0)$ for some neighborhood W in N where $\lambda : W \to \mathbf{R}^k$ is a submersion, since we may take λ to be the projection onto $\mathbf{E}_2$ of ψ_β above. This provides some motivation for the following definition.

Definition. *Let $S \subset N$ be a codimension-k submanifold and let $f : M \to N$ be a C^r-map. We say that the map f is transverse to S at the point $p \in M$ if either (i) $f(p) \notin S$, or (ii) $f(p) \in S$, $Df(p)$ is kernel-splitting and*

$$Df(p) \cdot T_pM + T_{f(p)}S = T_{f(p)}N.$$

The notation $f \pitchfork_p S$ means that the map f is transverse to the submanifold S at the point p in the domain of f. If f is transverse to S at all points in some set W, we write $f \pitchfork_W S$, or simply $f \pitchfork S$ to mean that f is transverse to S at all points in its domain.

We can now easily obtain the following result.

Theorem (pull-back via transversal maps). *Let $f : M \to N$ be a C^r-map of Banach manifolds, $S \subset N$ be a C^r-submanifold, and assume that $f \pitchfork S$. Then $f^{-1}(S)$ is an immersed submanifold of M, and is an embedded submanifold if S is compact. Furthermore, if S has finite codimension k in N, then $f^{-1}(S)$ has codimension k in M.*

Proof. First consider a small neighborhood V of a point q of S. As we have noted, $S \cap V = \lambda^{-1}(0)$ for some surjection $\lambda : V \to \mathbf{R}^k$. We show that 0 is a regular value of $\lambda \circ f$. First, let $p \in (\lambda \circ f)^{-1}(0)$. Then $p \in f^{-1}(S \cap V)$. Since $f \pitchfork S$, then we have that

$$Df(p) \cdot T_pM + T_{f(p)}S = T_{f(p)}N.$$

Applying $D\lambda(f(p))$ to both sides,

$$\begin{aligned} D\lambda(f(p)) \cdot Df(p) \cdot T_pM + D\lambda(f(p)) \cdot T_{f(p)}S &= D\lambda(f(p)) \cdot T_{f(p)}N, \\ D(\lambda \circ f)(p) \cdot T_pM &= \mathbf{R}^k, \end{aligned}$$

by the chain rule, and noting that λ is a submersion and $T_{f(p)}S$ is the kernel of $D\lambda(f(p))$. This shows that $\lambda \circ f$ is a submersion. Furthermore, we know that $(D(\lambda \circ f)(p))^{-1}(0) = (Df(p))^{-1} \cdot T_{f(p)}S$ is a subspace of T_pM which is complemented and whose complement is isomorphic to the complement of $T_{f(p)}S$ in $T_{f(p)}N$. Indeed, for any linear surjection $A : \mathbf{E} \to \mathbf{F}$, we have that the induced map $\tilde{A} : \mathbf{E}/A^{-1}(\mathbf{F}') \to \mathbf{F}/\mathbf{F}'$ is an isomorphism. Thus $ker(D(\lambda \circ f)(p))$ has closed complement so that $\lambda \circ f$ is kernel-splitting. Thus, $\lambda \circ f$ has 0 as a regular value and so $(\lambda \circ f)^{-1}(0) = f^{-1}(S \cap V)$ is an embedded submanifold of M. Taking a union of neighborhoods that cover S, we see that $f^{-1}(S)$ is a union of embedded submanifolds, which will be an immersed submanifold. In the case that S is compact, the above union can be made finite, so that $f^{-1}(S)$ is still an embedded submanifold.

Also, if S has codimension-k, then we know that the complement of $T_{f(p)}S$

in $T_{f(p)}N$ is isomorphic to $\mathbf{R}^k$. By our observation, the complement of $T_p(f^{-1}(S)$ would be isomorphic to $\mathbf{R}^k$, so that $f^{-1}(S)$ is also a submanifold of codimension-k.

The following corollary to this result is essentially a direct extension of the implicit function theorem.

Corollary. *Let $f : M \to N, S \subset N, f \pitchfork S$ as above. If for $p \in f^{-1}(S)$ we have that $T_pM = \mathbf{E}_p^1 \oplus \mathbf{E}_p^2$ such that the transversality condition holds with the sum being direct when T_pM is replaced by $\mathbf{E}_p^1$, i.e.,*

$$Df(p) \cdot \mathbf{E}_p^1 \oplus T_{f(p)}S = T_{f(p)}N,$$

then in a neighborhood U_p of p, we have that for any local coordinates $\varphi : U_p \to T_pM$ such that $\varphi = (\varphi^1, \varphi^2), \varphi(p) = (0,0)$ with $D\varphi^i(p) \cdot T_pM = \mathbf{E}_p^i$, we know that the component of $f^{-1}(S) \cap U_p$ through p is the graph of a C^r-function from $\mathbf{E}_1$ to $\mathbf{E}_2$.

Proof. Consider the function $F = \lambda \circ f \circ \varphi^{-1} : \mathbf{E}_p^1 \times \mathbf{E}_p^2 \to \mathbf{F}'$ where λ is as above. Then

$$\begin{aligned} D_1F(0,0) \cdot T_pM &= DF(0,0) \cdot \mathbf{E}_p^1 \\ &= D(\lambda \circ f)(p) \cdot \mathbf{E}_p^1 \\ &= D\lambda(f(p)) \cdot Df(p) \cdot \mathbf{E}_p^1. \end{aligned}$$

But this equals

$$D\lambda(f(p)) \cdot (Df(p) \cdot \mathbf{E}_p^1 + T_{f(p)}S) = D\lambda(f(p)) \cdot T_{f(p)}N$$

since $T_{f(p)}S = ker(D\lambda(f(p))$. Since λ is a submersion, we have that $D_1F(0,0)$ is surjective.

Also, we have that $ker(D_1F(0,0))$ is trivial. Otherwise, there would be a $v \in \mathbf{E}_p^1$ with $Df(p) \cdot v \in ker(D\lambda(f(p))) = T_{f(p)}S$, which cannot happen as the sum in the statement was direct. Thus, $D_1F(0,0)$ is an isomorphism, and the result follows from the implicit function theorem.

We will use this result in section 5.2 to obtain a parameterisation of the submanifold of critical points of the evaluation map at a point corresponding to a bifurcating critical point. We will use this to compute the relationship between qualitative properties resulting from the geometry of this manifold and the jets of the vectorfield at the critical point.

5.1 Critical Points of the Evaluation Map

In this section we will look at *critical points* of the evaluation map

$$ev : \mathcal{X}^r(M) \times M \longrightarrow TM.$$

The evaluation map $ev(X,p) = X(p)$ can be thought of as a parameterised vectorfield on M where the parameter is the vectorfield $X \in \mathcal{X}^r(M)$. We are primarily interested in the critical points of individual vectorfields and families of vectorfields in the study of bifurcations of critical points—and these are related to the critical points of the evaluation map: A point p is a critical point for a vectorfield X iff (X,p) is a critical point for ev. However, the critical points of ev are especially useful for studying bifurcations as the local geometry of the set Σ_0 of critcal points of ev depends on the relationship between changes in a vectorfield (perturbations) and changes in critical points. Exploiting the properties of ev that were developed in the last chapter, we shall examine the relationship between the k-jets of a vectorfield X at a critical point p and the local geometry of Σ_0 at (X,p).

The advantage of this approach is that it allows us to take a particularly geometric view of parameterised families of vectorfields. If X_μ is a family of C^r – vectorfields depending on a parameter μ, where the parameter is in some compact manifold Λ, possibly with boundary, then the family (if it is at all reasonable) X_μ defines an embedding

$$\Lambda \longrightarrow \mathcal{X}^r(M).$$

The image of this embedding will be a submanifold of (the Banach space) $\mathcal{X}^r(M)$. Thus, a parameterised family of vectorfields can be regarded as a submanifold of $\mathcal{X}^r(M)$ or a smooth embedding $\Lambda \to \mathcal{X}^r(M)$. This geometric point of view makes it much easier to see the mechanism behind certain bifurcations, and will provide us with a coherent approach to the whole study of bifurcation theory. In a later section, this geometric viewpoint is used in conjunction with *transversality theory* to obtain genericity results for vectorfields and families of vectorfields.

Critical Points, Zero Section

In order to define critical points for vectorfields on compact manifolds, we will use the definition of a critical point for the local representatives of such a vectorfield and then extend the definition in the obvious way.

Definition. *A point p is a critical point of the vectorfield $X \in \mathcal{X}^r(M)$ iff for some (and hence for any) chart $(U_\alpha, \varphi_\alpha)$ with $p \in U_\alpha$, the induced local representative X_α defined by*

$$T\varphi_\alpha \circ X \circ \varphi_\alpha^{-1}(x) = (x, X_\alpha(x)) \in \mathbf{R}^n | \varphi_\alpha(U_\alpha) \times \mathbf{R}^n,$$

has the corresponding point $x = \varphi_\alpha(p)$ as a critical point; i.e., the value $X_\alpha(\varphi_\alpha(p))$ of the local representative at the corresponding point is zero. Equivalently, p is a critical point iff

$$T\varphi_\alpha(X(p)) = (\varphi_\alpha(p), 0).$$

It is obvious that this definition is independant of the choice of chart $(U_\alpha, \varphi_\alpha)$.

While this may seem like an awfully formal definition for such a straightforward concept, this definition does motivate us to define the *zero section* of the tangent bundle TM.

The zero section 0_{TM} of the tangent bundle TM is simply the set of all vectors in that are zero vectors in the sense that their local representatives are zero vectors.

Thus $0_{TM} \subset TM$. It is easy to see that 0_{TM} is a submanifold of TM as

$$T\varphi_\alpha \circ 0_{TM}|TU_\alpha = \mathbf{R}^n \times \{0\},$$

for any vector bundle chart map $T\varphi_\alpha$. Clearly 0_{TM} has codimension n in TM. The zero section can also be thought of as a vectorfield on M; For each point $p \in M$, there is a zero vector $0_{TM}(p)$ at p. Thus, the zero section is a section of the tangent bundle, and is also a vectorfield as the map $0_{TM} : M \to TM$ is smooth.

The Manifold Σ_0

The set Σ_0 of critical points of the evaluation map is the pull back of the zero section by the evaluation map;

$$\Sigma_0 = ev^{-1}(0_{TM}).$$

In section 3.2 it was shown that ev was a kernel-splitting submersion. Thus, applying the results of sect 4.1, Σ_0 must be a codimension-n submanifold of $\mathcal{X}^r(M) \times M$ as 0_{TM} is a codimension-n submanifold of TM.

If (X, p) is a point in Σ_0, then the tangent space $T_{(X,p)}\Sigma_0$ indicates the relationships between perturbations in X and changes in the critical point p. For example, if $(\xi, v) \in T_{(X,p)}\Sigma_0$, then perturbing the vectorfield X in the direction ξ will move the critical point p in the direction v; More precisely, we can say that for the 1-parameter family $X_\epsilon = X + \epsilon\xi$, there is a corresponding 1-parameter family of critical points $p(\epsilon)$ for small ϵ such that

$$p(0) = p, \; X_\epsilon(p(\epsilon)) = 0_{TM}(p(\epsilon)), \text{ and } p'(0) = v.$$

In order to compute the tangent space $T_{(X,p)}\Sigma_0$, recall that since

$$ev(X, p) = 0_{TM}(p)$$

for all $(X,p) \in \Sigma_0$, then differentiating along Σ_0 gives us that

$$T_{(X,p)}ev(\xi,v) = T_p 0_{TM}(v)$$

for all $(\xi,v) \in T_{(X,p)}\Sigma_0$. Putting this expression into coordinate form,

$$T^2\varphi_\alpha \circ T_{(X,p)}ev(\xi,v) = T^2\varphi_\alpha \circ T_p 0_{TM}(T\varphi_\alpha^{-1}(v_\alpha)),$$

or

$$\begin{aligned} T_{(X,x)}ev_\alpha(\xi,v_\alpha) &= T_x(T\varphi_\alpha \circ 0_{TM}\varphi_\alpha^{-1})(v_\alpha), \\ &= (x,0,v_\alpha,0), \end{aligned}$$

as $T\varphi_\alpha \circ 0_{TM}(p) = (\varphi_\alpha(p),0)$. Recall that the local representative of the evaluation map ev_α was defined for a chart $(U_\alpha,\varphi_\alpha)$ on M by

$$ev_\alpha(X,x) = T\varphi_\alpha \circ ev(X,\varphi_\alpha^{-1}(x)),$$

or

$$ev_\alpha(X,x) = T(id_{\mathcal{X}^r(M)} \times \varphi_\alpha) \circ ev \circ (id_{\mathcal{X}^r(M)} \times \varphi_\alpha)^{-1}$$

so that

$$Tev_\alpha = T^2\tilde{\varphi}_\alpha \circ Tev \circ (T\tilde{\varphi}_\alpha)^{-1}$$

where $\tilde{\varphi}_\alpha = (id_{\mathcal{X}^r(M)} \times \varphi_\alpha)$. Finally we have

$$ev_\alpha(X,x) = (x, X_\alpha(x))$$

so that

$$Tev_\alpha(X_\alpha,x,\xi_\alpha,v_\alpha) = \Big(x, X_\alpha(x), v_\alpha, \xi_\alpha(x) + DX_\alpha(x)\cdot v_\alpha\Big),$$

where $x = \varphi_\alpha(p)$, and the subscripted quantities are the local representatives in the φ_α coordinates. This shows that

$$(\xi,v) \in T_{(X,p)}\Sigma_0 \text{ iff } \xi_\alpha(x) + DX_\alpha(x)\cdot v_\alpha = 0.$$

This tells us some things about the geometry of Σ_0 at (X,p). First, in the case where the linearisation DX_α is non-singular, we see that v_α can be expressed in terms of the perturbation $\xi_\alpha(x)$ of the vectorfield X_α at the critical point $x = \varphi_\alpha(p)$. This means that an arbitrary perturbation ξ of the vectorfield X will cause the critical point to move in the local coordinates in the direction $v_\alpha = -(DX_\alpha(x) \cdot \xi_\alpha(x))$ to first order. In the case where the linearisation $DX_\alpha(x)$ has non-trivial kernel, then for $v_0 \in KerDX(p)$, we have $(\xi_0, v_0) \in T_{(X,p)}\Sigma_0$, for any perturbation ξ such that $\xi(p) = 0$ $(i.e.\xi(p) = 0_{TM(p)}.)$ For other perturbations ξ of X we may have several directions v in which critical points may move, or no directions. We shall see what the interpretations of these results are in the following analysis.

Non-Degenerate Critical Points

A critical point p of a vectorfield X is *non-degenerate* if $X \pitchfork_p 0_{TM}$. This means that

$$DX(p) \cdot T_pM + T_{0_{TM}(p)}0_{TM} = T_{0_{TM}(p)}TM.$$

In terms of local coordinates, this means that $DX_\alpha(x)$ is a submersion, as $T^2\varphi_alpha \circ T_{0_{TM}(p)}0_{TM} = (x, 0, \mathbf{R}^n, 0)$. Indeed,

$$DX(p) : T_pM \to T_{X(p)}TM,$$

whence

$$\begin{aligned} T^2\varphi_\alpha \circ DX(p) \circ (T\varphi_\alpha)^{-1}(x, v_\alpha) &= T(T\varphi_\alpha \circ X \circ \varphi_\alpha^{-1}) \cdot (x, v_\alpha) \\ &= (x, X_\alpha(x), v_\alpha, DX_\alpha(x) \cdot v_\alpha) \end{aligned}$$

which is surjective at $(x, X_\alpha(x))$ if $DX_\alpha(x)$ is surjective.

We may apply the implicit function theorem in the case of a non-degenerate critical point. Consider $ev : \mathcal{X}^r(M) \times M \to TM$. We know that ev is surjective, but in the case where $DX_\alpha(x)$ is surjective, we have that

$$Dev(X, p) \cdot T_pM \oplus T_{0_{TM}(p)}0_{TM} = T_{0_{TM}(p)}TM,$$

so that we may apply the extension to the implicit function theorem in section 4.1 to conclude that there is a unique implicit function $\Phi : \mathcal{X}^r(M) \to M$ such that

$$ev(\tilde{X}, \Phi(\tilde{X})) \in 0_{TM}.$$

For vectorfields $\tilde{X}$ sufficiently near X in $\mathcal{X}^r(M)$. In fact

$$ev(X, \Phi(X)) = 0_{TM}(\Phi(X)),$$

so that $\Phi(X) \in M$ is a critical point for the vectorfield X. As we have already seen in Sect. 1.3, this shows that the critical point $\Phi(X)$ is a smooth function of the perturbation $\tilde{X}$ of X and is a locally unique critical point in a neighbourhood of the point p.

Degenerate Critical Points

A critical point $p \in M$ of the vectorfield $X \in \mathcal{X}^r(M)$ is a *degenerate* critical point if the derivative of X at p has non-trivial kernel;i.e,, the map

$$DX(p) : T_pM \to T_{X(p)}TM$$

vanishes on a non-trivial subspace of T_pM. In local coordinates, we can see what this means for the linearisation $DX_\alpha(x)$ of $X_\alpha(x)$, where $x = \varphi_\alpha(x)$. We have that

$$T\varphi_\alpha \circ X \varphi_\alpha^{-1} : x \longmapsto (x, X_\alpha(x)),$$

so that

$$T^2\varphi_\alpha \circ TX \circ T\varphi_\alpha^{-1}(x, v) = (x, X_\alpha(x), v, DX_\alpha(x) \cdot v).$$

Since $T\varphi_\alpha(T_pM) = \{x\} \times \mathbf{R}^n$, and $T^2\varphi_\alpha(T_{X(p)}TM) = (x, X_\alpha(x)) \times \mathbf{R}^n \times \mathbf{R}^n$, then we see the that $T\varphi_\alpha(ker(DX(p)) = \{x\} \times ker(DX_\alpha(x))$, which is what we would expect. In the next section we shall see how may apply the extension to the implicit function theorem in section 4.1 to a non- degenerate critical point in order to obtain a well known bifurcation result in a very geometric way.

5.2 An Example—The Saddle-Node Again

In the previous section we saw what the main distinction between non-degenerate and degenerate critical points of vectorfields. For a non-degenerate critical point p of a vectorfield X, the critical point varies smoothly under small perturbations of X. This is due to the implicit function theorem; More precisely, we can express the critical point p as a smooth function $p(X)$,

$$p : \mathcal{X}^r(M)|N_X \longrightarrow M,$$

with $p(X) = p$, defined in a neighbourhood N_X of $X \in \mathcal{X}^r(M)$. This is in turn due to the local geometry of Σ_0, specifically that the tangent space $T_{(X,p)}\Sigma_0$ is not "vertically tangent" in any direction. However, in the case of a degenerate critical point p, we have already seen that $T_{(X,p)}\Sigma_0$ is tangent to $\{X\} \times M$ along directions that are in $KerDX(p)$.

In order to consider the analysis of degenerate critical points, let us first consider the simplest case in which Σ_0 is tangent to $\{X\} \times M$ along only one direction $v^0 \in T_pM$. This corresponds to a critical point p for which the linearisation $DX(p)$ has 1-dimensional kernel spanned by v^0. Furthermore, let us assume that Σ_0 is quadratically tangent to $\{X\} \times M$ in this direction. The graph of Σ_0 is shown in Figure 5.2.1.

From the graph of this function, we can see what kind of qualitative change will occur when the vectorfield passes-through X. On one side of the graph, there is no

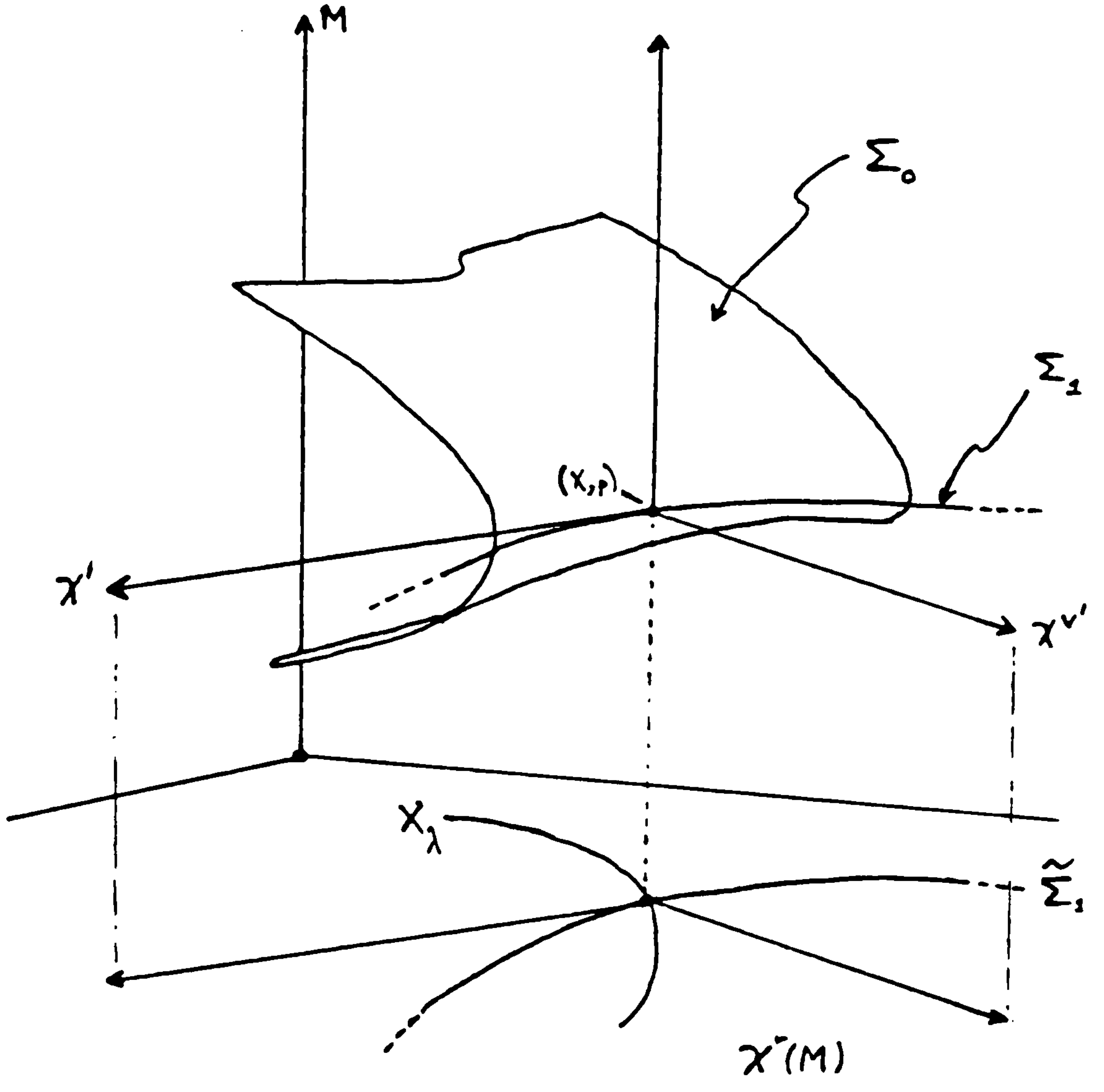

Figure 5.2.1

critical point near p, then as we cross the fold, the critical point p appears and divides into two separate points that grow apart at quadratic rates. This is exactly what happened with the saddle-node example in section 1.2.

We need to consider a parameterisation of Σ_0 near (X,p); i.e., we need to express Σ_0 as the graph of a smooth function. In order to find suitable variables for such a parameterisation, let us try and find a subspace $S \subset T_{(X,p)}(\mathcal{X}^r(M) \times M)$ such that

$$Dev(X,p) \cdot S \oplus T_{0_{TM}(p)} 0_{TM} = T_{0_{TM}(p)} TM.$$

Then, the subspace S will satisfy the conditions of the generalised implicit function theorem of section 4.1 and Σ_0 will be locally diffeomorphic to the graph of a function $\sigma_0 : S' \to S$, where S' complements S in $T_{(X,p)}(\mathcal{X}^r(M) \times M)$. Invoking local coordinates, we have that

$$\begin{aligned} T^2\varphi_\alpha \circ T(TM)_0 &= TT\varphi_\alpha((TM)_0) \\ &= T(\mathbf{R}^n \times \{0\}) \\ &= \mathbf{R}^n \times \{0\} \times \mathbf{R}^n \times \{0\} \end{aligned}$$

and

$$T^2\varphi_\alpha(T_{0_{TM}(p)}TM) = \{\varphi_\alpha(p)\} \times \{0\} \times \mathbf{R}^n \times \mathbf{R}^n.$$

Since

$$Tev_\alpha(X, x, \xi, v_\alpha) = (x, X_\alpha(x), v_\alpha, \xi_\alpha(x) + DX_\alpha(x) \cdot v_\alpha)$$

from before, we can see what vectors would comprise a suitable subspace S. Indeed, we need to find a set of (ξ, v) such that $\xi_\alpha(x) + DX_\alpha(x) \cdot v_\alpha$ will span $\mathbf{R}^n$. Since the range of $DX_\alpha(x)$ is $n-1$ dimensional, we can choose vectors of the form $(0, DX(p) \cdot v)$ to span an $n-1$ dimensional subspace of S. Finally, adding a vector of the form $(\xi, 0)$ where $\xi_\alpha(x)$ complements the range of $DX_\alpha(x)$ would give an n-dimensional subspace satisfying our requirements.

Now, let π' be a projection onto a complement of range $DX_\alpha(x)$ in $\mathbf{R}^n$. Let v' be a vectorfield such that $\pi'(v'_\alpha(x)) \neq 0$. Clearly v' defines a 1-dimensional subspace $\mathcal{X}^{v'}$ of $\mathcal{X}^r(M)$ which is complemented by the subspace

$$\mathcal{X}' = \{\xi \in \mathcal{X}^r(M) | \pi'(\xi_\alpha(x)) = 0\}.$$

Then the subspaces $S = \mathcal{X}^{v'} \times span(v'(p))$ and $S' = \mathcal{X}' \times ker(DX(p))$ are as previously required.

We can see that near to (X,p), the submanifold Σ_0 will admit a parameterisation

$$\sigma_0 : \mathcal{X}' \times ker(DX(p)) \longrightarrow \Sigma_0.$$

We can use this parameterisation for the computation of certain quantities; Specifically, we are interested in obtaining conditions on the k-jets of X at p that are equivalent to the non-degeneracy conditions of the quadradic tangency of Σ_0 to $\{X\} \times M$ in the direction v^0. In terms of the parameterisation σ_0, this non-degeneracy condition is

$$\pi' \circ D_{22}\sigma_0^1(0,0) \neq 0$$

where σ_0^1 is the vectorfield part of σ_0, and the D_{22} means differentiating twice along the second component (along the kernel of $DX(p)$.)

Recalling that $ev(X,p) = 0_{TM}(p)$ for any (X,p) in Σ_0, we have the parameterisation

$$ev(\sigma_0^1(X',v), \sigma_0^2(X',v)) = 0_{TM}(\sigma_0^2(X',v)),$$

for (X',v) in $\mathcal{X}' \times ker(DX(p))$ where σ_0^1 and σ_0^2 are the vectorfield and manifold components of σ_0, respectively. We may take tangents of both sides along $\{0\} \times ker(DX(p))$ using the composite function rule:

$$\begin{aligned} Tev \circ (\sigma_0^1 \times \sigma_0^2)(0,0,0,w) &= Tev \circ (T\sigma_0^1(0,0,0,w) \times T\sigma_0^2(0,0,0,w)) \\ &= T(0_{TM} \circ \sigma_0^2)(0,0,0,w) \\ &= T0_{TM} \circ T\sigma_0^2(0,0,0,w). \end{aligned}$$

Taking another tangent along $ker(DX(p))$ by the composite function rule, we have

$$T^2ev \circ \Big((T^2\sigma_0^1 \times T^2\sigma_0^2)(0,0,0,w,0,0,0,u)\Big) = T^2 0_{TM} \circ T^2\sigma_0^2(0,0,0,w,0,0,0,u).$$

Now, we introduce local coordinate maps so that we can see what this equation means in terms of the Taylor series of X_α at the point p_α. So,

$$T^2ev : T^2\mathcal{X}^r(M) \times T^2M \longrightarrow T^3M,$$

whence

$$\begin{aligned} T^3\varphi_\alpha \circ T^2ev \circ (T^2(id_{\mathcal{X}^r(M)} \times \varphi_\alpha))^{-1} & \\ \circ \Big(T^2\sigma_0^1 \times (T^2\varphi_\alpha \circ T^2\sigma_0^2)\Big)&(0,0,0,w,0,0,0,u) \\ = T^3\varphi_\alpha \circ T^2 0_{TM} \circ (T^2\varphi_\alpha)^{-1} \circ T^2\varphi_\alpha \circ T^2\sigma_0^2&(0,0,0,w,0,0,0,u). \end{aligned}$$

But

$$T^3\varphi_\alpha \circ T^2ev \circ T^2(id_{\mathcal{X}^r(M)} \times \varphi_\alpha)^{-1} = T^2ev_\alpha.$$

We may compute an expression for T^2ev_α from

$$ev_\alpha(X,x) = (x, X_\alpha(x)),$$

$$Tev_\alpha(X,x,\xi,v) = (x, X_\alpha(x), v, \xi_\alpha(x) + DX_\alpha(x)\cdot v),$$

whence

$$\begin{aligned} T^2ev_\alpha&(X,x,\xi,v,\eta,u,\varsigma,w) \\ &= \Big(x, X_\alpha(x), v, \xi_\alpha(x) + DX_\alpha(x)\cdot v, u, \eta_\alpha(x) + DX_\alpha(x)\cdot u, w, \\ &\qquad D\xi_\alpha(x)\cdot u + D\eta_\alpha(x)\cdot v + D^2X_\alpha(x)\cdot(u,v) + \varsigma_\alpha(x) + DX_\alpha(x)\cdot w\Big). \end{aligned}$$

Now, we can compute expressions for the quantities $\xi, v, \eta, u, \varsigma, w$ from our parameterisation σ_0. We have

$$(\sigma_0^1 \times \sigma_0^2)(0,0) = (X,p),$$

whence

$$T(\sigma_0^1 \times \sigma_0^2)(0,0,0,a) = (X, p, D_2\sigma_0^1(0,0)\cdot a, D_2\sigma_0^2(0,0)\cdot a),$$

and, differentiating once more along $kerDX(p)$, we have

$$T^2(\sigma_0^1 \times \sigma_0^2)(0,0,0,a,0,0,0,b)$$
$$= (X, p, D_2\sigma_0^1 \cdot a, D_2\sigma_0^2 \cdot a, D_2\sigma_0^1 \cdot b, D_2\sigma_0^2 \cdot b, D_{22}\sigma_0^1 \cdot (a,b), D_{22}\sigma_0^1 \cdot (a,b)),$$

so that

$$\xi = D_2\sigma_0^1(0,0) \cdot a \qquad v = D_2\sigma_0^2 \cdot a$$
$$\eta = D_2\sigma_0^1(0,0) \cdot b \qquad u = D_2\sigma_0^2 \cdot b$$
$$\varsigma = D_{22}\sigma_0^1(0,0) \cdot (a,b) \qquad v = D_{22}\sigma_0^2 \cdot (a,b).$$

Now, from the quadratic tangency conditions,the parameterisation σ_0 satisfies $D_2\sigma_0^1(0,0) = 0$, whence ξ, η above are zero. Also, $D_2\sigma_0^2(0,0)$ will be along $ker(DX(p))$. Substituting this into the eighth (last) component of T^2ev_α, we have that

$$D^2X_\alpha(x) \cdot (u,v) + T\varphi_\alpha \circ D_{22}\sigma_0^1(0,0) \cdot (a,b) + DX_\alpha(x) \cdot w = 0,$$

by equating this with the last component of $T^2\varphi_\alpha \circ T0_{TM}$. Applying the projection π' onto the complement of $range(DX(p))$, we have that

$$\pi'_\alpha \cdot D^2X_\alpha(x) \cdot (u,v) = -T\varphi_\alpha \circ \left(\pi' \cdot D_{22}\sigma_0^1(0,0) \cdot (a,b))\right).$$

Recalling the non-degeneracy condition for the quadratic tangency, that

$$\pi' \cdot D_{22}\sigma_0^1(0,0) \neq 0,$$

we have that the equivalent condition in terms of the jets of X at p will be

$$\pi' \cdot D^2X(p) \cdot (v,v) \neq 0$$

for $v \in ker(DX(p))$.

The Manifold Σ_1

The "fold" in Σ_0 appears to be a submanifold of $\mathcal{X}^r(M) \times M$ of codimension-$n+1$, as it appears to be of codimension-1 in Σ_0. Knowing that this fold is the set of vectorfield-point pairs (X,p) such that X satisfies our saddle-node conditions at p-that $X(p) = 0$, the derivative $DX(p)$ has rank $n-1$, and the second derivative satisfies the quadratic non-degeneracy condition we derived above. It is easy to show that this defines a codimension-$n+1$ submanifold when we consider that this is the pull-back of a codimension-$n+1$ submanifold of the bundle of 2-jets of vectorfields. Consider the submanifold of $J^1(TM)$ defined by our conditions on $X(p)$ and $DX(p)$. In natural vectorbundle coordinates induced by a chart φ_α, these conditions will become

$$X_\alpha(x) = 0$$

and

$$DX_\alpha(x) \text{ has rank } n-1 .$$

The set of $n \times n$ matrices that has rank $n-1$ is a codimension-1 submanifold of $L(\mathbf{R}^n)$ by the implicit function theorem. Indeed, consider $A \in L(\mathbf{R}^n)$ having rank $n-1$. Then, there is a neighborhood of A such that all matrices have rank at least $n-1$. The determinant map $det : L(\mathbf{R}^n) \to \mathbf{R}$ is a submersion, and the set of rank $n-1$ matrices is $det^{-1}(0)$ restricted to the set of matrices of rank at least $n-1$, which is an open set in $L(\mathbf{R}^n)$. Thus, the set of rank n-1 matrices is a submanifold of codimension equal to the codimension of $\{0\}$ in $\mathbf{R}^1$.

The quadratic non-degeneracy condition is open in that it persists under small perturbations in the 2-jet, so that our saddle-node conditions do define a codimension-2 submanifold of $J^2(TM)$. The manifold Σ_1 indicated in Figure 3.1 is the pull-back of this submanifold by the evaluation map for 2-jets, and so is also a codimension-$n+1$ submanifold.

The important observation to make about Σ_1 is that locally it will project to a codimension-1 submanifold $\tilde{\Sigma}_1$ of $\mathcal{X}^r(M)$. For a vectorfield $X \in \tilde{\Sigma}_1$, we see that $T_X\Sigma_1$ is complemented by the direction $\mathcal{X}^{v'}$ shown in figure 3.1. This direction is the direction in which $v' \cdot \xi$ changes for perturbing vectorfields ξ.

Geometrically we see what causes the saddle-node bifurcation. If we cross through the submanifold Σ_1 transversely, then we are crossing the fold in the graph of Σ_0, and this will cause the appearance of pair of critical points associated with the fold. For a one parameter vectorfield X_λ crossing Σ_1 at the parameter value λ_0, the condition that we cross Σ_1 transversely is that $\frac{d}{d\lambda}(\pi' \cdot X_\lambda)(p) \neq 0$. This motivates the saddle-node bifurcation theorem, which I have taken from Guckenheimer and Holmes [1983], Theorem 3.4.1.

Theorem. *Let $\dot{x} = f_\mu(x)$ be a differential equation depending on the single parameter μ. When $\mu = \mu_0$, assume that there is a an equilibrium p for which the following hypotheses are satisfied:*

(SN1) $D_x f_{\mu_0}$ has a simple eigenvalue 0 with right eigenvector v and left eigenvector w.

(SN2) $w \cdot D_\mu f(p, \mu_0) \neq 0$.

(SN3) $w(D_x^2 f_{\mu_0}(p) \cdot (v, v) \neq 0$. Then there is a smooth curve of equilibria in $\mathbf{R}^n \times \mathbf{R}$ passing through (p, μ_0), tangent to the hyperplane $\mathbf{R}^n \times \{\mu_0\}$. For μ on one side of μ_0 there are no equilibria of f_μ near p, while for μ on the other side of μ_0 there are two distinct equilibria of f_μ near p

We can see that the conditions (SN1) and (SN3) are the same as our conditions that $ker(DX(p))$ is 1-dimensional and that $\pi' \cdot D^2 X_\alpha(p) \cdot (v, v) \neq 0$ for $v \in ker(DX(p))$. If we think of a one-parameter family of vectorfields as a one-dimensional arc in $\mathcal{X}^r(M)$, then the condition (SN2) is really a requirement that the arc X_λ crosses through the point $X \in \mathcal{X}^r(M)$ transversal to projection of the "fold" in figure 5.2.1 onto $\mathcal{X}^r(M)$. This projected fold is a codimension-1 submanifold $\Sigma_1 \subset \mathcal{X}^r(M)$, which we shall discuss in the next section.

5.3 Concluding Remarks.

We have seen an example of how me may study the local geometry of an evaluation map to get a result for saddle-node bifurcations of critical points of vectorfields. However, it remains to be seen how this approach might be extended to obtain other results from *generic bifurcation theory*, or even what the connection between the material in the preceding sections is related to generic bifurcations. In this section, I will make some comments (of a somewhat speculative nature) on how this approach can be extended to include the other types of bifurcations that are encountered in generic bifurcation theory and also on the connection between this approach and that of generic bifurcation theory.

Transversality and Genericity

We have the definition of transversality of a map $f : M \to N$ to a submanifold S of N. We have considered transversality for two particular kinds of maps; vectorfields, which are maps $M \to TM$, and parameterised vectorfields, which are maps from a parameter space into the space $\mathcal{X}^r(M)$ of all vectorfields. It is the transversality of these maps to various submanifolds that give rise to genericity results for vectorfields and parameterised vectorfields by the following well-known results.

Theorem (Openness of Transversal Maps). *Let M, N, S be C^r-Banach manifolds, and $f : M \to N$ be transverse to the closed submanifold $S \subset N$. Then there* *is a neighborhood U_f of f in $C^r(M,N)$ such that $g \in U_f$ implies $g \pitchfork S$. Thus, the set of maps transverse to S is open in $C^k(M,N)$.*

The statement and proof of this result is found in Palis and de Melo (p. 24), or Abraham, Marsden and Ratiu (p. 179). Of course, the statement belies the definition of the topology of the space $C^r(M,N)$. In the case where M is compact, this is the C^r-compact–open topology. We saw an example of this in topology section 3.1 for the case of vectorfields on a compact manifold, where the topology on the space $\mathcal{X}^r(M)$ of sections was defined. As in most of the results in section 4.1, the proof relies on locally replacing transversality of maps with the equivalent surjectivity conditions. Specifically, consider $p \in f^{-1}(S)$. Then, there exists neighborhoods $U_p \subset M, V_{f(p)} \subset N$, and a submersion $\lambda : V_f(p) \to \mathbf{F}'$ such that $f(U_p) \subset V_{f(p)}$, and $S \cap V_{f(p)} = \lambda^{-1}(0)$. Then there is a neighborhood W_f of f in $C^r(M,N)$ such that $g(U_p) \subset V_{f(p)}$ for all $g \in W_f$. We have that $g \pitchfork_{U_p} S$ if and only if $\lambda \circ g$ is a submersion on U_p. Finally, we note that the evaluation map that is defined near $(f,p) \in \mathcal{X}^r(M,N) \times M$ by

$$(g,q) \to D(\lambda \circ g)(q)$$

is continuous, so that there are open neighborhoods $W_f' \subset W_f, U_p' \subset U_p$ such that $D(\lambda \circ g')$ is surjective as the set of linear surjections is open. This means that for all g in the C^r-neighborhood W_f' of f, we have $g \pitchfork_{U_p} S$. Finally, since $f^{-1}(S)$ is compact, we can cover $f^{-1}(S)$ with a finite number of the U_p', and so the intersection of the corresponding W_f' is a C^r-neighborhood of f of functions transverse to S.

One consequence of the above proof is that we may generalize the above result and claim that the set of maps transverse to a given submanifold is open in any space for which the evaluation map considered above is continuous. For example, the above theorem is not directly applicable in the case of vectorfields, as $\mathcal{X}^r(M)$

is not a space of the form $C^r(M,N)$, but we still have that the set of vectorfields that are transverse to a given submanifold of TM is open in $\mathcal{X}^r(M)$..

There is a corresponding result concerning the density of transversal maps. First, recall *Sard's theorem* for a map $f : M \to N$ of *finite-dimensional* manifolds, which states that the set of regular values of the map is dense in N. The C^∞-version of this theorem is found in most texts on advanced calculus, differential geometry or introductory differential topology. The Sard-Smale theorem is a generalization of this result where the map $f : M \to N$ is now a C^k*-Fredholm* map and M, N are Banach manifolds. In this case, if f is sufficiently smooth, then the set of regular values is again dense. This extension of Sard's theorem is the subject of Appendix E. of Abraham *et. al.* [1983]and also covered in section 16 of Abraham and Robbin [1967].

Given an evaluation map, we may use the following lemma.

Lemma. *Let $F : \Lambda \times M \to N$ be transverse to $S \subset N$. We know that $\tilde{S} = F^{-1}(S)$ is a submanifold of $\Lambda \times M$. Let $\pi_\Lambda : \Lambda \times M \to \Lambda$ be the natural projection map and $\tilde{\pi}_\Lambda = \pi_\Lambda|\tilde{S} : \tilde{S} \to \Lambda$. Then, $F_\lambda = F(\lambda, \cdot)$ is transverse to S if and only if λ is a regular value of $\tilde{\pi}_\Lambda$.*

Proof. If F_λ is not transverse to S, then for some $(\lambda, p) \in \Lambda \times M$ we have

$$DF(\lambda,p) \cdot T_pM + T_{F(\lambda,p)}N \not\supset T_{F(\lambda,p)}N.$$

But, $T_{F(\lambda,p)}S = DF(\lambda,p) \cdot T_{(\lambda,p)}\tilde{S}$, so that

$$DF(\lambda,p) \cdot \left(T_pM + T_{(\lambda,p)}\tilde{S}\right) \not\supset T_{F(\lambda,p)}N.$$

Obviously $T_pM + T_{(\lambda,p)}\tilde{S} \not\supset T_{(\lambda,p)}(\Lambda \times M)$, so that

$$\begin{aligned} D\pi_\Lambda(\lambda,p) \cdot \left(T_pM + T_{(\lambda,p)}\tilde{S}\right) &= D\tilde{\pi}_\Lambda(\lambda,p) \cdot T_{(\lambda,p)}\tilde{S} \\ &\not\supset D\pi_\Lambda(\lambda,p) \cdot T_{(\lambda,p)}(\Lambda \times M), \end{aligned}$$

so that $D\tilde{\pi}_\Lambda(\lambda, p)$ is not surjective, whence λ is not a regular value of $\tilde{\pi}_\Lambda$. Conversely, if λ is not a regular value of $\tilde{\pi}_\Lambda$, then for some $p \in M$, we have

$$D\tilde{\pi}_\Lambda(\lambda, p) \cdot T_{(\lambda,p)}\tilde{S} \not\supset T_\lambda \Lambda,$$

which implies

$$D\pi_\Lambda(\lambda, p) \cdot \left(T_pM + T_{(\lambda,p)}\tilde{S}\right) \not\supset T_\lambda \Lambda,$$

so that

$$T_{(\lambda,p)}\tilde{S} + T_pM \not\supset T_\lambda \Lambda.$$

Letting $(v_\lambda, 0)$ be in $T_\lambda\Lambda \times \{0\}$ but not in $T_{(\lambda,p)}\tilde{S} + T_pM$, we know that

$$\begin{aligned} DF(\lambda, p) \cdot (v_\lambda, 0) &\notin DF(\lambda, p) \cdot \left(T_pM + T_{(\lambda,p)}\tilde{S}\right) \\ &= DF(\lambda, p) \cdot T_pM + T_{F(\lambda,p)}S, \end{aligned}$$

whence $DF(\lambda, p) \cdot T_pM + T_{F(\lambda,p)}S \not\supset T_{F(\lambda,p)}N$, so that F is not transverse to S.

In the case where the above manifolds are finite dimensional, the projection map $\tilde{\pi}_\Lambda$ satisfies Sard's theorem, so that the set of λ which are regular values of $\tilde{\pi}_\Lambda$ are dense in Λ. Therefore the set of λ for which $F_\lambda \pitchfork S$ is dense in Λ. However, the interesting application of this lemma is in the case where F is an *evaluation map* of the form

$$ev_{\mathcal{F}} : \mathcal{F}(M, N) \times M \to N,$$

where M is a finite-dimensional manifold, and $\mathcal{F}(M, N)$ is a Banach space of functions from M to N. One example of this kind of map we have already seen is the evaluation map $ev : \mathcal{X}^r(M) \times M \to TM$ for vectorfields. Another example was the map $(g, x) \to D(\lambda \circ g)$ that was used in the proof of the openness theorem for transversal maps above (the range of this map is actually a linear map bundle, and the proof used the fact that the set of surjections in in this bundle is open– this is done carefully in Abraham and Robbin [1967], section 18.) The set of all one-parameter families of vectorfields can also be considered in this way. A one-parameter vectorfield $X_\lambda, \lambda \in [0,1]$ can be considered as a map $\lambda \to \mathcal{X}^r(M)$ in the Banach space $C^r([0,1], \mathcal{X}^r(M))$. Then there is an evaluation map associated with parameterised vectorfields

$$Ev : C^r([0,1], \mathcal{X}^r(M)) \times [0,1] \longrightarrow \mathcal{X}^r(M)$$

$$(X, \lambda) \longmapsto X(\lambda, \cdot) \in \mathcal{X}^r(M).$$

In order to apply the above lemma to the evaluation map ev or Ev, we must verify that ev and Ev are transverse to any submanifolds their respective ranges, TM and $\mathcal{X}^r(M)$. We have already verified that ev was a submersion in section 3.2. Similarly, it is easy to see that Ev is a submersion, for $DEv(X_\lambda, \lambda_0) \cdot (Y_\lambda, v_\lambda) = Y_{\lambda_0} + \frac{\partial}{\partial \lambda} X_{\lambda_0} \cdot v_\lambda$, and since we can choose anything we want for Y_{λ_0}, the derivative is surjective. Then, given a submanifold S of the range (TM or $\mathcal{X}^r(M)$), we need to verify that the induced projections

$$\tilde{\pi}_0 : (\mathcal{X}^r(M) \times M)|\tilde{S} \to \mathcal{X}^r(M),$$

$$\tilde{\pi}_1 : (C^r([0,1], \mathcal{X}^r(M)) \times [0,1])|\tilde{S} \to C^r([0,1], \mathcal{X}^r(M))$$

are *Fredholm maps*, where $\tilde{S}$ is the pull back of S by the evaluation map ev or Ev. This requires that (i) both the kernel and range of $D\tilde{\pi}_i$ split, (ii) the kernel is finite-dimensional and (iii) the range has finite codimension. If the submanifold S has finite codimension, then $\tilde{S}$ has (the same) finite codimension, and it is not difficult to verify that (i)—(iii) hold. Then the Sard–Smale theorem holds for the $\tilde{\pi}_i$, and we have the same density result as above. Specifically, for any submanifold S (remember that TM is finite dimensional) of TM, the set of vectorfields transverse to S is dense in $\mathcal{X}^r(M)$. For one-parameter vectorfields, we have that for any submanifold $S \subset \mathcal{X}^r(M)$ of *finite-codimension,* the set of one-parameter families X_λ that are transverse to S is dense in $C^r([0,1], \mathcal{X}^r(M))$.

Thus, for a compact submanifold S, we have that the set of maps transverse to S is both open and dense. If S is paracompact, then the resulting set of maps transverse to S will be the countable intersection of open-dense sets, which is called a *residual* or *generic* set. This is where the connection between transversality and generic properties arises. We consider an example.

G_0 is a generic property.

A vectorfield $X \in \mathcal{X}^r(M)$ is said to have property G_0 if all critical points of X are non-degenerate in the sense of section 5.1. Since X has non-degenerate critical points if and only if $X \pitchfork 0_{TM}$, and therefore the set of vectorfields satisfying G_0 is open and dense (since 0_{TM} is compact.)

Generic 1-Parameter Vectorfields

In order to find generic properties of 1-parameter vectorfields, we must look for 1-parameter submanifolds of $\mathcal{X}^r(M)$. In the last section, I hinted that one could project the "fold" that gave rise to the saddle-node bifurcation and get a codimension-1 submanifold $\Sigma_1 \subset \mathcal{X}^r(M)$. If we consider the evaluation map for 2-jets of vectorfields, then we may pull-back the submanifold of $J^2(TM)$ that corresponds to saddle-node critical points. In natural vectorbundle coordinates on $J^2(TM)$ associated with a chart $(U_\alpha, \varphi_\alpha)$, we have that this becomes

$$\{0\} \times \{A \in L(\mathbf{R}^n) : rank(A) = n-1\} \times \{\text{non-degenerate quadratic forms}\}$$

in $\mathbf{R}^n \times L(\mathbf{R}^n) \times L^2(\mathbf{R}^n)$. Since $\{0\} \subset \mathbf{R}^n$ is codimension-n, the set of rank-$n-1$ maps is codimension-1 in $L(\mathbf{R}^n)$ (since it is the pull-back of $0 \in \mathbf{R}$ by the map $A \mapsto det(A)$), and the set of non-degenerate bilinear forms is open in $L^2(\mathbf{R}^n)$, then the submanifold of saddle-node critical points is of codimension $n+1$ in $J^2(TM)$. We may pull this back to a submanifold $\Sigma_1 \subset (\mathcal{X}^r(M) \times M)$ of codimension $n+1$. If this manifold projects to a codimension-1 submanifold $\tilde{\Sigma}_1 \subset \mathcal{X}^r(M)$, then we will have a genericity theorem for one-parameter families of vectorfields.

The openness result for transversal maps tells us that if a 1-parameter family has a saddle-node bifurcation (which crosses $\tilde{\Sigma}_1$ transversely), then nearby (in the sense of $C^r([0,1], \mathcal{X}^r(M))$) 1-parameter families will also have a saddle-node bifurcation. The qualitative change associated with the saddle-node bifurcation occurs when the manifold $\tilde{\sigma}_1$ is crossed transversely—we recall that on one side there are no critical points in a neighborhood of the bifurcating critical point, while on the other side the bifurcating critical point splits into two.

Extension to Other Bifurcations.

The result that G_0 is a generic property for vectorfields is the first part of the Kupka-Smale theorem (in Abraham and Marsden [1978], chapter 7 or Abraham and Robbin [1967]), which gives several generic properties for vectorfields. Each of the generic properties (G_0)—(G_3) is associated with some kind of non-degeneracy of critical points, periodic orbits, or the intersection of the stable and unstable manifolds of a pair of critical elements. In the literature on generic bifurcation theory (Sotomayor [1973a], Sotomayor [1973b], Sotomayor [1974]), codimension-1 bifurcations are examined that correspond to failure of each one of the non-degeneracy conditions/generic properties in the Kupka-Smale theorem, and submanifolds of $\mathcal{X}^r(M)$ are constructed that give rise to a corresponding genericity result for degenerate (bifurcating) equilibria of 1-parameter families of vectorfields.

The content of Abraham and Robbin [1967], is a modernised proof of the Kupka-Smale theorem that relies on evaluation-transversality techniques like that used above to show that G_0 was a generic property in $\mathcal{X}^r(M)$. In order to consider the corresponding codimension-1 bifurcations, it is necessary to extend all of the evaluation transversality results used in Abraham and Robbin [1967]to account for higher orderr terms. For example, the property G_0 involves non-degeneracy or the derivative, whereas it is necessary to have non-degeneracy of the second derivative for the saddle-node bifurcation. If this were done (and I believe it is possible), then the results from the literature on generic bifurcations could be reproduced along the lines of what I have done in this thesis, which would be more geometrically intuitive, and hence of some pedagogical value.

Bibliography

Abraham, R.H., and Marsden, J.E. [1978]. *Foundations of Mechanics.* Benjamin/Cummings: Reading, MA.

Abraham, R.H., Marsden, J.E., and Ratiu, T. [1983]. *Manifolds, Tensor Analysis and Applications.* Addison-Wesley: Reading, MA.

Abraham, R.H., and Robbin, J. [1967]. *Transversal Mappings and Flows.* Benjamin: Reading, MA.

Chillingworth, D.R.J. [1976]. *Differential Topology with a View to Applications.* Pitman: London.

Guckenheimer, J., and Holmes, P.J. [1983]. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields.* Springer-Verlag: New York.

Palis, J. and de Melo, W. [1982]. *Geometric Theory of Dynamical Systems.* Springer-Verlag: New York, Heidelberg, Berlin.

Sotomayor, J. [1974]. Generic one-parameter families of vector fields, *Publ. Math. I.H.E.S. No 43 (1974), 5—46.*

Sotomayor, J. [1973a]. Structural Stability and Bifurcation Theory. In *Dynamical Systems*, M.M. Peixoto (ed.), pp. 549—560. Acedemic Press: New York.

Sotomayor J. [1973b]. Generic Bifurcations of Dynamical Systems. In *Dynamical Systems*, M.M. Peixoto (ed.), pp. 561—582. Acedemic Press: New York.