this system. (b) Draw the phase portrait of the system for $\epsilon > 0$. Note: In this case, the term $\epsilon\dot{x}$ models viscous damping. (c) What is the fate of the solution with initial condition $(x(0), \dot{x}(0)) = (4, 0)$ for $\epsilon = 0.1$? Note: To solve this problem you will probably have to resort to numerics. How do we know that the result obtained by a numerical simulation is correct?

Exercise 1.76. [Gradient Systems] If H is a Hamiltonian, then the vector field $\operatorname{grad} H$ is everywhere orthogonal to the corresponding Hamiltonian vector field. What are the properties of the flow of $\operatorname{grad} H$? More generally, for a smooth function $G : \mathbb{R}^n \to \mathbb{R}$ (maybe n is odd), let us define the associated gradient system

$$\dot{x} = \operatorname{grad} G(x).$$

Because a conservative force is the negative gradient of a potential, many authors define the gradient system with potential G to be $\dot{x} = -\operatorname{grad} G(x)$. The choice of sign simply determines the direction of the flow. Prove the following statements: (a) A gradient system has no periodic orbits. (b) If a gradient system has a rest point, then all of the eigenvalues of its linearization at the rest point are real. (c) In the plane, the orbits of the gradient system with potential G are orthogonal trajectories for the orbits of the Hamiltonian system with Hamiltonian G. (d) If $x_0 \in \mathbb{R}^n$ is an isolated maximum of the function $G : \mathbb{R}^n \to \mathbb{R}$, then x_0 is an asymptotically stable rest point of the corresponding gradient system $\dot{x} = \operatorname{grad} G(x)$.

Exercise 1.77. [Rigid Body Motion] A system that is not Hamiltonian, but closely related to this class, is given by Euler's equations for rigid body motion. The angular momentum $M = (M_1, M_2, M_3)$ of a rigid body, relative to a coordinate frame rotating with the body with axes along the principal axes of the body and with origin at its center of mass, is related to the angular velocity vector Ω by $M = A\Omega$, where A is a symmetric matrix called the *inertia matrix*. Euler's equation is $\dot{M} = M \times \Omega$. Equivalently, the equation for the angular velocity is $A\dot{\Omega} = (A\Omega) \times \Omega$. If A is diagonal with diagonal components (moments of inertia) (I_1, I_2, I_3), show that Euler's equations for the components of the angular momentum are given by

$$\begin{aligned}
\dot{M}_1 &= -\left(\frac{1}{I_2} - \frac{1}{I_3}\right) M_2 M_3, \\
\dot{M}_2 &= \left(\frac{1}{I_1} - \frac{1}{I_3}\right) M_1 M_3, \\
\dot{M}_3 &= -\left(\frac{1}{I_1} - \frac{1}{I_2}\right) M_1 M_2.
\end{aligned}$$

Assume that $0 < I_1 \le I_2 \le I_3$. Find some invariant manifolds for this system. Can you use your results to find a qualitative description of the motion? As a physical example, take this book and hold its covers together with a rubber band. Then, toss the book vertically three times, imparting a rotation in turn about each of its axes of symmetry (see Figure 1.13). Are all three rotary motions Lyapunov stable? Do you observe any other interesting phenomena associated with the motion? For example, pay attention to the direction of the front cover of the book after each toss. Hint: Look for invariant quadric surfaces; that is, manifolds defined as level sets of quadratic polynomials (first integrals) in the variables (M_1, M_2, M_3). For example, show that the kinetic energy given by $\frac{1}{2}\langle A\Omega, \Omega\rangle$ is constant along orbits. The total angular momentum (length of the angular momentum) is also conserved. For a complete mathematical description of rigid body motion, see [12]. For a mathematical description of the observed "twist" in the rotation of the tossed book, see [20]. Note that Euler's equations do not describe the motion of the book in space. To do so would require a functional relationship between the coordinate system rotating with the body and the position coordinates relative to a fixed coordinate frame in space.

1.8.2 Smooth Manifolds

Because the modern definition of a smooth manifold can appear quite formidable at first sight, we will formulate a simpler equivalent definition for the class of manifolds called the *submanifolds* of $\mathbb{R}^n$. Fortunately, this class is rich enough to contain the manifolds that are met most often in the study of differential equations. In fact, every manifold can be "embedded" as a submanifold of some Euclidean space. Thus, the class that we will study can be considered to contain all manifolds.

Recall that a manifold is supposed to be a set that is locally the same as $\mathbb{R}^k$. Thus, whatever is meant by "locally the same," every open subset of $\mathbb{R}^k$ must be a manifold.

If $W \subseteq \mathbb{R}^k$ is an open set and $g : W \to \mathbb{R}^{n-k}$ is a smooth function, then the graph of g is the subset of $\mathbb{R}^n$ defined by

$$\operatorname{graph}(g) := \{(w, g(w)) \in \mathbb{R}^n : w \in W\}.$$

The set $\operatorname{graph}(g)$ is the same as $W \subseteq \mathbb{R}^k$ up to a nonlinear change of coordinates. By this we mean that there is a smooth map G with domain W and image $\operatorname{graph}(g)$ such that G has a smooth inverse. In fact, such a map $G : W \to \operatorname{graph}(g)$ is given by $G(w) = (w, g(w))$. Clearly, G is smooth. Its inverse is the linear projection on the first k coordinates of the point $(w, g(w)) \in \operatorname{graph}(g)$; that is, $G^{-1}(w, g(w)) = w$. Thus, G^{-1} is smooth as well.

Open subsets and graphs of smooth functions are the prototypical examples of what we will call submanifolds. But these classes are too restrictive; they include objects that are *globally* the same as some Euclidean space. The unit circle $\mathbb{T}$ in the plane, also called the one-dimensional torus, is an example of a submanifold that is not of this type. Indeed, $\mathbb{T} := \{(x, y) : x^2 + y^2 = 1\}$ is not the graph of a scalar function defined on an open subset of $\mathbb{R}$. On the other hand, every point of $\mathbb{T}$ is contained in a *neighborhood in $\mathbb{T}$ that is the graph of such a function*. In other words, $\mathbb{T}$ is *locally* the same as $\mathbb{R}$. In fact, each point in $\mathbb{T}$ is in one of the four sets

$$\begin{aligned}
S_{\pm} &:= \{(x, y) \in \mathbb{R}^2 : y = \pm\sqrt{1 - x^2}, \quad |x| < 1\}, \\
S^{\pm} &:= \{(x, y) \in \mathbb{R}^2 : x = \pm\sqrt{1 - y^2}, \quad |y| < 1\}.
\end{aligned}$$

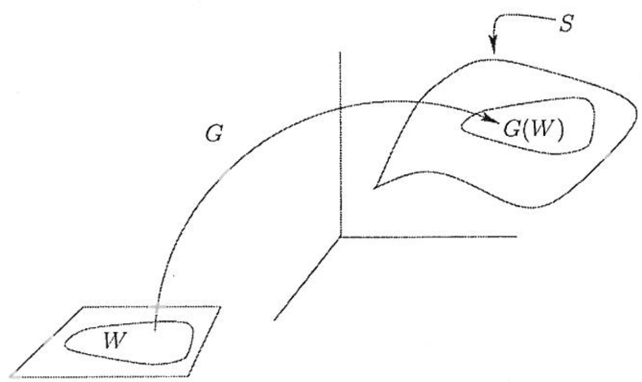

Figure 1.14: A chart for a two-dimensional submanifold in $\mathbb{R}^3$.

Submanifolds of $\mathbb{R}^n$ are subsets with the same basic property: Every point in the subset is in a neighborhood that is the graph of a smooth function.

To formalize the submanifold concept for subsets of $\mathbb{R}^n$, we must deal with the problem that, in the usual coordinates of $\mathbb{R}^n$, not all graphs are given by sets of the form

$$\{(x_1, \dots, x_k, g_{k+1}(x_1, \dots, x_k), \dots, g_n(x_1, \dots, x_k)) : \\ (x_1, \dots, x_k) \in W \subseteq \mathbb{R}^k\}.$$

Rather, we must allow, as in the example provided by $\mathbb{T}$, for graphs of functions that are not functions of the first k coordinates of $\mathbb{R}^n$. To overcome this technical difficulty we will build permutations of the variables into our definition.

Definition 1.78. Suppose that $S \subseteq \mathbb{R}^n$ and $x \in S$. The pair (W, G) where W is an open subset of $\mathbb{R}^k$ for some $k \leq n$ and $G : W \to \mathbb{R}^n$ is a smooth function is called a *k-dimensional submanifold chart for S at x* (see Figure 1.14) if there is an open set $U \subseteq \mathbb{R}^n$ with $x \in U \cap S$ such that $U \cap S = G(W)$ and one of the following two properties is satisfied:
1) The integer k is equal to n and G is the identity map.
2) The integer k is less than n and G has the form

$$G(w) = A\begin{pmatrix} w \\ g(w) \end{pmatrix}$$

where $g : W \to \mathbb{R}^{n-k}$ is a smooth function and A is a nonsingular $n \times n$ matrix.

Definition 1.79. The set $S \subseteq \mathbb{R}^n$ is called a *k-dimensional smooth submanifold* of $\mathbb{R}^n$ if there is a k-dimensional submanifold chart for S at every point x in S.

The map G in a submanifold chart (W, G) is called a *submanifold coordinate map.* If S is a submanifold of $\mathbb{R}^n$, then (even though we have not yet defined the concept), let us also call a submanifold S of $\mathbb{R}^n$ a smooth manifold.

As an example, let us show that $\mathbb{T}$ is a one-dimensional manifold. Consider a point in the subset $S^+ = \{(x, y) : x = \sqrt{1 - y^2}, |y| < 1\}$ of $\mathbb{T}$. Define the set $W := \{t \in \mathbb{R} : |t| < 1\}$, the function $g : W \to \mathbb{R}$ by $g(t) = \sqrt{1 - t^2}$, the set $U := \{(x, y) \in \mathbb{R}^2 : (x - 1)^2 + y^2 < 2\}$, and the matrix

$$A := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then we have

$$\mathbb{T} \cap U = \left\{\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} t \\ g(t) \end{pmatrix}, \quad t \in W\right\}.$$

Similarly, $\mathbb{T}$ is locally the graph of a smooth function at points in the subsets S^- and $S_\pm$, as required.

A simple but important result about submanifold charts is the following proposition.

Proposition 1.80. *If (W, G) is a submanifold chart for a k-dimensional submanifold of $\mathbb{R}^n$, then the function $G : W \to G(W) \subseteq S$ is invertible. Moreover, the inverse of G is the restriction of a smooth function that is defined on all of $\mathbb{R}^n$.*

Proof. The result is obvious if $k = n$. If $k < n$, then define $\Pi : \mathbb{R}^n \to \mathbb{R}^k$ to be the linear projection on the first k-coordinates; that is, $\Pi(x_1, \dots, x_n) = (x_1, \dots, x_k)$, and define

$$F : G(W) \to W$$

by

$$F(s) = \Pi A^{-1} s.$$

Clearly, F is smooth as a function defined on all of $\mathbb{R}^n$. Also, if $w \in W$, then

$$F \circ G(w) = F\Big(A\begin{pmatrix} w \\ g(w) \end{pmatrix}\Big) = \Pi A^{-1} A\begin{pmatrix} w \\ g(w) \end{pmatrix} = w.$$

If $s \in G(W)$, then $s = A\begin{pmatrix} w \\ g(w) \end{pmatrix}$ for some $w \in W$. Hence, we also have

$$G(F(s)) = G(w) = s.$$

This proves that F is the inverse of G. □

If S is a submanifold, then we can use the submanifold charts to define the open subsets of S.

Definition 1.81. Suppose that S is a submanifold. The *open* subsets of S are all possible unions of all sets of the form $G(W)$ where (W, G) is a submanifold chart for S.

The next proposition is an immediate consequence of the definitions.

Proposition 1.82. *If S is a submanifold of $\mathbb{R}^n$ and if V is an open subset of S, then there is an open set U of $\mathbb{R}^n$ such that $V = S \cap U$; that is, the topology defined on S using the submanifold charts agrees with the subspace topology on S.*

As mentioned above, one of the main reasons for defining the manifold concept is to distinguish those subsets of $\mathbb{R}^n$ on which we can use the calculus. To do so, let us first make precise the notion of a smooth function.

Definition 1.83. Suppose that S_1 is a submanifold of $\mathbb{R}^m$, S_2 is a submanifold of $\mathbb{R}^n$, and F is a function $F : S_1 \to S_2$. We say that F is *differentiable* at $x_1 \in S_1$ if there are submanifold charts (W_1, G_1) at x_1 and (W_2, G_2) at $F(x_1)$ such that the map $G_2^{-1} \circ F \circ G_1 : W_1 \to W_2$ is differentiable at $G_1^{-1}(x_1) \in W_1$. If F is differentiable at each point of an open subset V of S_1, then we say that F is differentiable on V.

Definition 1.84. Suppose that S_1 and S_2 are manifolds. A smooth function $F : S_1 \to S_2$ is called a *diffeomorphism* if there is a smooth function $H : S_2 \to S_1$ such that $H(F(s)) = s$ for every $s \in S_1$ and $F(H(s)) = s$ for every $s \in S_2$. The function H is called the inverse of F and is denoted by F^{-1}.

With respect to the notation in Definition 1.83, we have defined the concept of differentiability for the function $F : S_1 \to S_2$, but we have not yet defined what we mean by its derivative. We have, however, determined the derivative relative to the submanifold charts used in the definition. Indeed, the *local representative* of the function F is given by $G_2^{-1} \circ F \circ G_1$, a function defined on an open subset of a Euclidean space with range in another Euclidean space. By definition; the *local representative of the derivative* of F relative to the given submanifold charts is the usual derivative in Euclidean space of this local representative of F. In the next subsection, we will interpret the derivative of F without regard to the choice of a submanifold chart; that is, we will give a coordinate-free definition of the derivative of F (see also Exercise 1.85).

Exercise 1.85. Prove: The differentiability of a function defined on a manifold does not depend on the choice of submanifold chart.

Exercise 1.86. (a) Show that $\dot\theta = f(\theta)$ can be viewed as a (smooth) differential equation on the unit circle if and only if f is periodic. To be compatible with the usual (angular) coordinate on the circle it is convenient to consider only 2π-periodic functions (see Section 1.8.5). (b) Describe the bifurcations that occur for the family $\dot\theta = 1 - \lambda \sin\theta$ with $\lambda \geq 0$. (c) For each $\lambda < 1$, the corresponding differential equation has a periodic orbit. Determine the period of this periodic orbit and describe the behavior of the period as $\lambda \to 1$ (see [218, p. 98]).

We have used the phrase "smooth function" to refer to a function that is continuously differentiable. In view of Definition 1.83, the smoothness of a function defined on a manifold is determined by the smoothness of its local representatives—functions that are defined on open subsets of Euclidean spaces. It is clear that smoothness of all desired orders can be defined in the same manner by imposing the requirement on local representatives. More precisely, if F is a function defined on a manifold S, then we will say that F is an element of $C^r(S)$, for r a nonnegative integer, $r = \infty$, or $r = \omega$, provided that at each point of S there is a local representative of F all of whose partial derivatives up to and including those of order r are continuous. If $r = \infty$, then all partial derivatives are required to be continuous. If $r = \omega$, then all local representatives are all required to have convergent power series representations valid in a neighborhood of each point of their domains. A function in C^ω is called *real analytic.*

In the subject of differential equations, specifying the minimum number of derivatives of a function required to obtain a result often obscures the main ideas that are being illustrated. Thus, as a convenient informality, we will often use the phrase "smooth function" to mean that the function in question has as many continuous derivatives as needed. In cases where the exact requirement for the number of derivatives is essential, we will refer to the appropriate class of C^r functions.

The next definition formalizes the concept of a coordinate system.

Definition 1.87. Suppose that S is a k-dimensional submanifold. The pair (V, Ψ) is called a *coordinate system* or *coordinate chart* on S if V is an open subset of S, W is an open subset of $\mathbb{R}^k$, and $\Psi : V \to W$ is a diffeomorphism.

Exercise 1.88. Prove: If (W, G) is a submanifold chart for a manifold S, then $(G(W), G^{-1})$ is a coordinate chart on S.

The abstract definition of a manifold is based on the concept of coordinate charts. Informally, a set S together with a collection of subsets S is defined to be a k-dimensional manifold if every point of S is contained in at least one set in S and if, for each member V of S, there is a corresponding open subset W of $\mathbb{R}^k$ and a function $\Psi : V \to W$ that is bijective. If two such subsets V_1 and V_2 overlap, then the domain of the map

$$\Psi_1 \circ \Psi_2^{-1} : \Psi_2(V_1 \cap V_2) \to W_1$$

is an open subset of $\mathbb{R}^k$ whose range is contained in an open subset of $\mathbb{R}^k$. The set S is called a manifold provided that all such "overlap maps" are smooth (see [120] for the formal definition). This abstract notion of a manifold has the advantage that it does not require a manifold to be a subset of a Euclidean space.

Exercise 1.89. Prove: If $F : \mathbb{R}^m \to \mathbb{R}^n$ is smooth and $F(S_1) \subseteq S_2$ for submanifolds S_1 and S_2, then the restriction of F to S_1 is differentiable.

Exercise 1.90. Prove: If $\alpha \in \mathbb{R}$, then the map $\mathbb{T} \to \mathbb{T}$ given by

$$(x, y) \mapsto (x \cos\alpha - y \sin\alpha, x \sin\alpha + y \cos\alpha)$$

is a diffeomorphism.

Now that we know the definition of a manifold, we are ready to prove that linear subspaces of $\mathbb{R}^n$ and regular level sets of smooth functions are manifolds.

Proposition 1.91. *A linear subspace of $\mathbb{R}^n$ is a submanifold.*

Proof. Let us suppose that S is the span of the k linearly independent vectors $v_1, \dots, v_k$ in $\mathbb{R}^n$. We will show that S is a k-dimensional submanifold of $\mathbb{R}^n$.

Let $\mathbf{e}_1, \dots, \mathbf{e}_n$ denote the standard basis of $\mathbb{R}^n$. By a basic result from linear algebra, there is a set consisting of $n-k$ standard basis vectors $f_{k+1}, \dots, f_n$ such that the vectors

$$v_1, \dots, v_k, f_{k+1}, \dots, f_n$$

are a basis for $\mathbb{R}^n$. (Why?) Let us denote the remaining set of standard basis vectors by $f_1, \dots, f_k$. For each $j = 1, \dots, k$, there are scalars λ_i^j and μ_i^j such that

$$f_j = \sum_{i=1}^{k} \lambda_i^j v_i + \sum_{i=k+1}^{n} \mu_i^j f_i.$$

Hence, if $(t_1, \dots, t_k) \in \mathbb{R}^k$, then the vector

$$\sum_{j=1}^{k} t_j f_j - \sum_{j=1}^{k} t_j \Big(\sum_{i=k+1}^{n} \mu_i^j f_i \Big) = \sum_{j=1}^{k} t_j \Big(\sum_{i=1}^{k} \lambda_i^j v_i \Big)$$

is in S; and, relative to the basis $f_1, \cdots, f_n$, the vector

$$(t_1, \dots, t_k, -\sum_{j=1}^{k} t_j \mu_{k+1}^j, \dots, -\sum_{j=1}^{k} t_j \mu_n^j)$$

is in S.

Define $g : \mathbb{R}^k \to \mathbb{R}^{n-k}$ by

$$g(t_1, \dots, t_k) := \Big(-\sum_{j=1}^{k} t_j \mu_{k+1}^j, \dots, -\sum_{j=1}^{k} t_j \mu_n^j \Big)$$

and let A denote the permutation matrix given by $A\mathbf{e}_j = f_j$. It follows that the pair $(\mathbb{R}^k, G)$, where $G : \mathbb{R}^k \to \mathbb{R}^n$ is defined by

$$G(w) = A \begin{pmatrix} w \\ g(w) \end{pmatrix},$$

is a k-dimensional submanifold chart such that $G(\mathbb{R}^k) = \mathbb{R}^n \cap S$. In fact, by the construction, it is clear that the image of G is a linear subspace of S. Moreover, because the image of G has dimension k as a vector space, the subspace $G(\mathbb{R}^k)$ is equal to S. □

As mentioned previously, linear subspaces often arise as invariant manifolds of differential equations. For example, consider the differential equation given by $\dot{x} = Ax$ where $x \in \mathbb{R}^n$ and A is an $n \times n$ matrix. If S is an invariant subspace for the matrix A, for example, one of its generalized eigenspaces, then, by Proposition 1.91, S is a submanifold of $\mathbb{R}^n$. Also, S is an invariant set for the corresponding linear system of differential equations. Although a complete proof of this proposition requires some results from linear systems theory that will be presented in Chapter 2, the essential features of the proof are simply illustrated in the special case where the linear transformation A restricted to S has a complete set of eigenvectors. In other words, S is a k-dimensional subspace of $\mathbb{R}^n$ spanned by k linearly independent eigenvectors $v_1, \dots, v_k$ of A. Under this assumption, if $Av_i = \lambda_i v_i$, then $t \to e^{\lambda_i t} v_i$ is a solution of $\dot{x} = Ax$. Also, note that $e^{\lambda_i t} v_i$ is an eigenvector of A for each $t \in \mathbb{R}$. Therefore, if $x_0 \in S$, then there are scalars $(a_1, \dots, a_k)$ such that $x_0 = \sum_{i=1}^k a_i v_i$ and

$$t \mapsto \sum_{i=1}^{k} e^{\lambda_i t} a_i v_i$$

is the solution of the ordinary differential equation with initial condition $x(0) = x_0$. Clearly, the corresponding orbit stays in S for all $t \in \mathbb{R}$.

Linear subspaces can be invariant sets for nonlinear differential equations. For example, consider the Volterra–Lotka system

$$\dot{x} = x(a - by), \qquad \dot{y} = y(cx - d).$$

In case a, b, c, and d are all positive, this system models the interaction of the population y of a predator and the population x of its prey. For this system, the x-axis and the y-axis are each invariant sets. Indeed, suppose that $(0, y_0)$ is a point on the y-axis corresponding to a population of predators with no prey, then $t \mapsto (0, e^{-dt}y_0)$ is the solution of the system starting at this point that models this population for all future time. This solution stays on the y-axis for all time, and, as there are is no spontaneous generation of prey, the predator population dies out in positive time.

Let us now discuss level sets of functions. Recall that the *level set with energy* c of a smooth function $H : \mathbb{R}^n \to \mathbb{R}$ is the set

$$S_c := \{x \in \mathbb{R}^n : H(x) = c\}.$$

Moreover, the level set S_c is called a regular level set if $\operatorname{grad} H(x) \neq 0$ for each $x \in S_c$.

Proposition 1.92. *If $H : \mathbb{R}^n \to \mathbb{R}$ is a smooth function, then each of its regular level sets is an $(n-1)$-dimensional submanifold of $\mathbb{R}^n$.*

It is instructive to outline a proof of this result because it provides our first application of a nontrivial and very important theorem from advanced calculus, namely, the implicit function theorem.

Suppose that S_c is a regular level set of H, choose $a \in S_c$, and define $F : \mathbb{R}^n \to \mathbb{R}$ by

$$F(x) = H(x) - c.$$

Let us note that $F(a) = 0$. Also, because $\operatorname{grad} H(a) \neq 0$, there is at least one integer $1 \leq i \leq n$ such that the corresponding partial derivative $\partial F/\partial x_i$ does not vanish when evaluated at a. For notational convenience let us suppose that $i = 1$. All other cases can be proved in a similar manner.

We are in a typical situation: We have a function $F : \mathbb{R} \times \mathbb{R}^{n-1} \to \mathbb{R}$ given by $(x_1, x_2, \ldots, x_n) \mapsto F(x_1, \ldots, x_n)$ such that

$$F(a_1, \ldots, a_n) = 0, \qquad \frac{\partial F}{\partial x_1}(a_1, a_2, \ldots, a_n) \neq 0.$$

This calls for an application of the implicit function theorem. A preliminary version of the theorem is stated here; a more general version will be proved later (see Theorem 1.259).

If $f : \mathbb{R}^\ell \times \mathbb{R}^m \to \mathbb{R}^n$ is given by $(p, q) \mapsto f(p, q)$, then, for fixed $b \in \mathbb{R}^m$, consider the function $\mathbb{R}^\ell \to \mathbb{R}^n$ defined by $p \mapsto f(p, b)$. Its derivative at $a \in \mathbb{R}^\ell$ will be denoted by $f_p(a, b)$. Of course, with respect to the usual bases of $\mathbb{R}^\ell$ and $\mathbb{R}^n$, this derivative is represented by an $n \times \ell$ matrix of partial derivatives.

Theorem 1.93 (Implicit Function Theorem). *Suppose that $F : \mathbb{R}^m \times \mathbb{R}^k \to \mathbb{R}^m$ is a smooth function given by $(p, q) \mapsto F(p, q)$. If (a, b) is in $\mathbb{R}^m \times \mathbb{R}^k$ such that $F(a, b) = 0$ and the linear transformation $F_p(a, b) : \mathbb{R}^m \to \mathbb{R}^m$ is invertible, then there exist two open metric balls $U \subseteq \mathbb{R}^m$ and $V \subseteq \mathbb{R}^k$ with $(a, b) \in U \times V$ together with a smooth function $g : V \to U$ such that $g(b) = a$ and $F(g(v), v) = 0$ for each $v \in V$. Moreover, if $(u, v) \in U \times V$ and $F(u, v) = 0$, then $u = g(v)$.*

Continuing with our outline of the proof of Proposition 1.92, let us observe that, by an application of the implicit function theorem to F, there is an open set $Z \subseteq \mathbb{R}$ with $a_1 \in Z$, an open set $W \subseteq \mathbb{R}^{n-1}$ containing the point $(a_2, \ldots, a_n)$, and a smooth function $g : W \to Z$ such that $g(a_2, \ldots, a_n) = a_1$ and

$$H(g(x_2, \ldots, x_n), x_2, \ldots, x_n) - c \equiv 0.$$

The set

$$U := \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_1 \in Z \text{ and } (x_2, \ldots, x_n) \in W\} = Z \times W$$

is open. Moreover, if $x = (x_1, \ldots, x_n) \in S_c \cap U$, then $x_1 = g(x_2, \ldots, x_n)$. Thus, we have that

$$\begin{aligned} S_c \cap U &= \{(g(x_2, \ldots, x_n), x_2, \ldots, x_n) : (x_2, \ldots, x_n) \in W\} \\ &= \{u \in \mathbb{R}^n : u = A\begin{pmatrix} w \\ g(w) \end{pmatrix} \text{ for some } w \in W\} \end{aligned}$$

where A is the permutation of $\mathbb{R}^n$ given by

$$(y_1, \ldots, y_n) \mapsto (y_n, y_1, \ldots, y_{n-1}).$$

In particular, it follows that S_c is an $(n-1)$-dimensional manifold.

Exercise 1.94. Show that $\mathbb{S}^{n-1} := \{(x_1, \ldots, x_n) \in \mathbb{R}^n : {x_1}^2 + \cdots + {x_n}^2 = 1\}$ is an $(n-1)$-dimensional manifold.

Exercise 1.95. For $p \in \mathbb{S}^2$ and $p \neq \pm e_3$ (the north and south poles) define $f(p) = v$ where $\langle v, p\rangle = 0$, $\langle v, e_3\rangle = 1 - z^2$, and $\langle p \times e_3, v\rangle = 0$. Define $f(\pm e_3) = 0$ Prove that f is a smooth function $f : \mathbb{S}^2 \to \mathbb{R}^3$.

Exercise 1.96. Show that the surface of revolution S obtained by rotating the circle given by $(x-2)^2 + y^2 = 1$ around the y-axis is a two-dimensional manifold. This manifold is diffeomorphic to a (two-dimensional) torus $\mathbb{T}^2 := \mathbb{T} \times \mathbb{T}$. Construct a diffeomorphism.

Exercise 1.97. Suppose that J is an interval in $\mathbb{R}$ and $\gamma : J \to \mathbb{R}^n$ is a smooth function. The image C of γ is, by definition, a curve in $\mathbb{R}^n$. Is C a one-dimensional submanifold of $\mathbb{R}^n$? Formulate and prove a theorem that gives sufficient conditions for C to be a submanifold. Hint: Consider the function $t \mapsto (t^2, t^3)$ for $t \in \mathbb{R}$ and the function $t \mapsto (1 - t^2, t - t^3)$ for two different domains: $t \in \mathbb{R}$ and $t \in (-\infty, 1)$. Can you imagine a situation where the image of a smooth curve is a dense subset of a manifold with dimension $n > 1$? Hint: Consider curves mapping into the two-dimensional torus.

Exercise 1.98. Show that the closed unit disk in $\mathbb{R}^2$ is not a manifold. Actually, it is a *manifold with boundary.* How should this concept be formalized?

Exercise 1.99. Prove that for $\epsilon > 0$ there is a $\delta > 0$ and a root r of the polynomial $x^3 - ax + b$ such that $|r| < \epsilon$ whenever $|a - 1| + |b| < \delta$.

Exercise 1.100. Show that if $f : \mathbb{R}^n \to \mathbb{R}^n$, A is a nonsingular $(n \times n)$-matrix and $|\epsilon|$ is sufficiently small, then the differential equation $\dot{x} = Ax + \epsilon f(x)$ has a rest point.

1.8.3 Tangent Spaces

We have used, informally, the following proposition: *If S is a manifold in $\mathbb{R}^n$, and $(x, f(x))$ is tangent to S for each $x \in S$, then S is an invariant manifold for the differential equation $\dot{x} = f(x)$.* To make this proposition precise, we will give a definition of the concept of a tangent vector on a manifold. This definition is the main topic of this section.

Let us begin by considering some examples where the proposition on tangents and invariant manifolds can be applied.

The vector field on $\mathbb{R}^3$ associated with the system of differential equations given by

$$\begin{aligned}\dot{x} &= x(y+z),\\ \dot{y} &= -y^2 + x\cos z,\\ \dot{z} &= 2x + z - \sin y\end{aligned} \tag{1.18}$$

is "tangent" to the linear two-dimensional submanifold $S := \{(x, y, z) : x = 0\}$ in the following sense: If $(a, b, c) \in S$, then the value of the vector function

$$(x, y, z) \mapsto (x(y+z), y^2 + x\cos z, 2x + z - \sin y)$$

at (a, b, c) is a vector in the linear space S. Note that the vector assigned by the vector field depends on the point in S. For this reason, we will view the vector field as the function

$$(x, y, z) \mapsto (x, y, z, x(y+z), -y^2 + x\cos z, 2x + z - \sin y)$$

where the first three component functions specify the *base point*, and the last three components, called the *principal part*, specify the vector that is assigned at the base point.

To see that S is an invariant set, choose $(0, b, c) \in S$ and consider the initial value problem

$$\dot{y} = -y^2, \quad \dot{z} = z - \sin y, \qquad y(0) = b, \quad z(0) = c.$$

Note that if its solution is given by $t \mapsto (y(t), z(t))$, then the function $t \mapsto (0, y(t), z(t))$ is the solution of system (1.18) starting at the point $(0, b, c)$. In particular, the orbit corresponding to this solution is contained in S. Hence, S is an invariant set. In this example, the solution is not defined for all $t \in (-\infty, \infty)$. (Why?) But, every solution that starts in S stays in S, as required by Definition 1.66.

The following system of differential equations,

$$\begin{aligned}\dot{x} &= x^2 - (x^3 + y^3 + z^3)x,\\ \dot{y} &= y^2 - (x^3 + y^3 + z^3)y,\\ \dot{z} &= z^2 - (x^3 + y^3 + z^3)z\end{aligned} \tag{1.19}$$

has a nonlinear invariant submanifold; namely, the unit sphere

$$\mathbb{S}^2 := \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}.$$

This fact follows from our proposition, provided that the vector field associated with the differential equation is everywhere tangent to the sphere. To prove this requirement, recall from Euclidean geometry that a vector in space is defined to be tangent to the sphere if it is orthogonal to the normal line passing through the base point of the vector. Moreover, the normal lines to the sphere are generated by the outer unit normal field given by the restriction of the vector field

$$\eta(x, y, z) := (x, y, z, x, y, z)$$

to $\mathbb{S}^2$. By a simple computation, it is easy to check that the vector field associated with the differential equation is everywhere orthogonal to η on $\mathbb{S}^2$; that is, at each base point on $\mathbb{S}^2$ the corresponding principal parts of the two vector fields are orthogonal, as required.

We will give a definition for tangent vectors on a manifold that generalizes the definition given in Euclidean geometry for linear subspaces and spheres. Let us suppose that S is a k-dimensional submanifold of $\mathbb{R}^n$ and (G, W) is a submanifold coordinate chart at $p \in S$. Our objective is to define the tangent space to S at p.

Definition 1.101. The *tangent space* to $\mathbb{R}^k$ with base point at $w \in \mathbb{R}^k$ is the set

$$T_w\mathbb{R}^k := \{w\} \times \mathbb{R}^k.$$

We have the following obvious proposition: If $w \in \mathbb{R}^k$, then the tangent space $T_w\mathbb{R}^k$, with addition defined by

$$(w, \xi) + (w, \zeta) := (w, \xi + \zeta)$$

and scalar multiplication defined by

$$a(w, \xi) := (w, a\xi),$$

is a vector space that is isomorphic to the vector space $\mathbb{R}^k$.

To define the *tangent space of the submanifold S at $p \in S$*, denoted T_pS, we simply move the space $T_w\mathbb{R}^k$, for an appropriate choice of w, to S

with a submanifold coordinate map. More precisely, suppose that (W, G) is a submanifold chart at p. By Proposition 1.80, the coordinate map G is invertible. If $q = G^{-1}(p)$, then define

$$T_pS := \{p\} \times \{v \in \mathbb{R}^n : v = DG(q)\xi,\ \xi \in \mathbb{R}^k\}. \tag{1.20}$$

Note that the set

$$\mathcal{V} := \{v \in \mathbb{R}^n : v = DG(q)\xi,\ \xi \in \mathbb{R}^k\}$$

is a k-dimensional subspace of $\mathbb{R}^n$. If $k = n$, then $DG(q)$ is the identity map. If $k < n$, then $DG(q) = AB$ where A is a nonsingular matrix and the $n \times k$ block matrix

$$B := \begin{pmatrix} I_k \\ Dg(q) \end{pmatrix}$$

is partitioned by rows with I_k the $k \times k$ identity matrix and g a map from W to $\mathbb{R}^{n-k}$. Thus, we see that $\mathcal{V}$ is just the image of a linear map from $\mathbb{R}^k$ to $\mathbb{R}^n$ whose rank is k.

Proposition 1.102. *If S is a manifold and $p \in S$, then the vector space T_pS is well-defined.*

Proof. If K is a second submanifold coordinate map at p, say $K : Z \to S$ with $K(r) = p$, then we must show that the tangent space defined using K agrees with the tangent space defined using G. To prove this fact, let us suppose that $(p, v) \in T_pS$ is given by

$$v = DG(q)\xi.$$

Using the chain rule, it follows that

$$v = \frac{d}{dt}G(q + t\xi)\Big|_{t=0}.$$

In other words, v is the directional derivative of G at q in the direction ξ. To compute this derivative, we simply choose a curve, here $t \mapsto q + t\xi$, that passes through q with tangent vector ξ at time $t = 0$, move this curve to the manifold by composing it with the function G, and then compute the tangent to the image curve at time $t = 0$.

The curve $t \mapsto K^{-1}(G(q + t\xi))$ is in Z (at least this is true for $|t|$ sufficiently small). Thus, we have a vector $\alpha \in \mathbb{R}^k$ given by

$$\alpha := \frac{d}{dt}K^{-1}(G(q + t\xi))\Big|_{t=0}.$$

We claim that $DK(r)\alpha = v$. In fact, we have

$$K^{-1}(G(q)) = K^{-1}(p) = r,$$

and

$$\begin{aligned} DK(r)\alpha &= \frac{d}{dt}K(K^{-1}(G(q + t\xi)))\Big|_{t=0} \\ &= \frac{d}{dt}G(q + t\xi)\Big|_{t=0} \\ &= v. \end{aligned}$$

In particular, T_pS, as originally defined, is a subset of the "tangent space at p defined by K." But this means that this subset, which is itself a k-dimensional affine subspace (the translate of a subspace) of $\mathbb{R}^n$, must be equal to T_pS, as required. □

Exercise 1.103. Prove: If $p \in \mathbb{S}^2$, then the tangent space $T_p\mathbb{S}^2$, as in Definition 1.20, is equal to

$$\{p\} \times \{v \in \mathbb{R}^3 : \langle p, v \rangle = 0\}.$$

Definition 1.104. The *tangent bundle* TS of a manifold S is the union of its tangent spaces; that is, $TS := \bigcup_{p \in S} T_pS$. Also, for each $p \in S$, the vector space T_pS is called the *fiber* of the tangent bundle over the base point p.

Definition 1.105. Suppose that S_1 and S_2 are manifolds, and $F : S_1 \to S_2$ is a smooth function. The *derivative*, also called the *tangent map*, of F is the function $F_* : TS_1 \to TS_2$ defined as follows: For each $(p, v) \in T_pS_1$, let (W_1, G_1) be a submanifold chart at p in S_1, (W_2, G_2) a submanifold chart at $F(p)$ in S_2, $(G_1^{-1}(p), \xi)$ the vector in $T_{G_1^{-1}(p)}W_1$ such that $DG_1(G_1^{-1}(p))\xi = v$, and $(G_2^{-1}(F(p)), \zeta)$ the vector in $T_{G_2^{-1}(F(p))}W_2$ such that

$$\zeta = D(G_2^{-1} \circ F \circ G_1)(G_1^{-1}(p))\xi.$$

The tangent vector $F_*(p, v)$ in $T_{F(p)}S_2$ is defined by

$$F_*(p, v) = \big(F(p), DG_2(G_2^{-1}(F(p)))\zeta\big).$$

Although definition 1.105 seems to be rather complex, the idea is natural: we simply use the local representatives of the function F and the definition of the tangent bundle to define the derivative F_* as a map with two component functions. The first component is F (to ensure that base points map to base points) and the second component is defined by the derivative of a local representative of F at each base point.

The following proposition is obvious from the definitions.

Proposition 1.106. *The tangent map is well-defined and it is linear on each fiber of the tangent bundle.*

The derivative, or tangent map, of a function defined on a manifold has a geometric interpretation that is the key to understanding its applications in the study of differential equations. We have already discussed this interpretation several times for various special cases. But, because it is so important, let us consider the geometric interpretation of the derivative in the context of the notation introduced in Definition 1.105. If $t \mapsto \gamma(t)$ is a curve—a smooth function defined on an open set of $\mathbb{R}$—with image in the submanifold $S_1 \subseteq \mathbb{R}^m$ such that $\gamma(0) = p$, and if

$$v = \dot{\gamma}(0) = \frac{d}{dt}\gamma(t)\Big|_{t=0},$$

then $t \mapsto F(\gamma(t))$ is a curve in the submanifold $S_2 \subseteq \mathbb{R}^n$ such that $F(\gamma(0)) = F(p)$ and

$$F_*(p,v) = \Big(F(p), \frac{d}{dt}F(\gamma(t))\Big|_{t=0}\Big).$$

We simply find a curve that is tangent to the vector v at p and move the curve to the image of the function F to obtain a curve in the range. The tangent vector to the new curve at $F(p)$ is the image of the tangent map.

Proposition 1.107. *A submanifold S of $\mathbb{R}^n$ is an invariant manifold for the ordinary differential equation $\dot{x} = f(x)$, $x \in \mathbb{R}^n$ if and only if*

$$(x, f(x)) \in T_xS$$

for each $x \in S$. If, in addition, S is compact, then each orbit on S is defined for all $t \in \mathbb{R}$.

Proof. Suppose that S is k-dimensional, $p \in S$, and (W, G) is a submanifold chart for S at p. The idea of the proof is to change coordinates to obtain an ordinary differential equation on W.

Recall that the submanifold coordinate map G is invertible and G^{-1} is the restriction of a linear map defined on $\mathbb{R}^n$. In particular, we have that $w \equiv G^{-1}(G(w))$ for $w \in W$. If we differentiate both sides of this equation and use the chain rule, then we obtain the relation

$$I = DG^{-1}(G(w))DG(w) \tag{1.21}$$

where I denotes the identity transformation of $\mathbb{R}^n$. In particular, for each $w \in W$, we have that $DG^{-1}(G(w))$ is the inverse of the linear transformation $DG(w)$.

Under the hypothesis, we have that $(x, f(x)) \in T_xS$ for each $x \in S$. Hence, the vector $f(G(w))$ is in the image of $DG(w)$ for each $w \in W$. Thus, it follows that

$$(w,\ DG^{-1}(G(w))f(G(w))) \in T_w\mathbb{R}^k,$$

and, as a result, the map

$$w \mapsto (w,\ DG^{-1}(G(w))f(G(w)))$$

defines a vector field on $W \subseteq \mathbb{R}^n$. The associated differential equation on W is given by

$$\dot{w} = DG^{-1}(G(w))f(G(w)). \tag{1.22}$$

Suppose that $G(q) = p$, and consider the initial value problem on W given by the differential equation (1.22) together with the initial condition $w(0) = q$. By the existence theorem, this initial value problem has a unique solution $t \mapsto \omega(t)$ that is defined on an open interval containing $t = 0$.

Define $\phi(t) = G(\omega(t))$. We have that $\phi(0) = p$ and, using equation (1.21), that

$$\begin{aligned}\frac{d\phi}{dt}(t) &= DG(\omega(t))\dot{\omega}(t)\\ &= DG(\omega(t)) \cdot DG^{-1}(G(\omega(t)))f(G(\omega(t)))\\ &= f(\phi(t)).\end{aligned}$$

Thus, $t \mapsto \phi(t)$ is the solution of $\dot{x} = f(x)$ starting at p. Moreover, this solution is in S because $\phi(t) = G(\omega(t))$. The solution remains in S as long as it is defined within the submanifold chart. The same result is true for every submanifold chart. Thus, the solution remains in S as long as it is defined.

Suppose that S is compact and note that the solution just defined is a solution of the differential equation $\dot{x} = f(x)$ defined on $\mathbb{R}^n$. By the extension theorem, if a solution of $\dot{x} = f(x)$ does not exist for all time, for example, if it exists only for $0 \le t < \beta < \infty$, then it approaches the boundary of the domain of definition of f or it blows up to infinity as t approaches β. As long as the solution stays in S, both possibilities are excluded if S is compact. Since the manifold S is covered by coordinate charts, the solution stays in S and it is defined for all time.

If S is invariant, $p \in S$ and $t \mapsto \gamma(t)$ is the solution of $\dot{x} = f(x)$ with $\gamma(0) = p$, then the curve $t \to G^{-1}(\gamma(t))$ in $\mathbb{R}^k$ has a tangent vector ξ at $t = 0$ given by

$$\xi := \frac{d}{dt}G^{-1}(\gamma(t))\Big|_{t=0}.$$

As before, it is easy to see that $DG(q)\xi = f(p)$. Thus, $(p, f(p)) \in T_pS$, as required. □

Exercise 1.108. Show that the function $f(\theta) = 1 - \lambda \sin\theta$ defines a (smooth) vector field on $\mathbb{T}^1$, but $f(\theta) = \theta - \lambda \sin\theta$ does not.

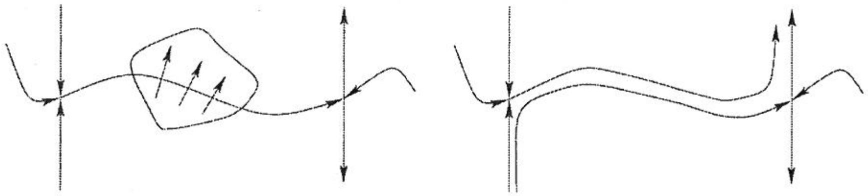

Figure 1.15: The left panel depicts a heteroclinic saddle connection and a locally supported perturbation. The right panel depicts the phase portrait of the perturbed vector field.

Exercise 1.109. State and prove a proposition that is analogous to Proposition 1.107 for the case where the submanifold S is not compact.

Exercise 1.110. We have mentioned several times the interpretation of the derivative of a function whereby a curve tangent to a given vector at a point is moved by the function to obtain a new curve whose tangent vector is the directional derivative of the function applied to the original vector. This interpretation can also be used to define the tangent space at a point on a manifold. In fact, let us say that two curves $t \mapsto \gamma(t)$ and $t \mapsto \nu(t)$, with image in the same manifold S, are equivalent if $\gamma(0) = \nu(0)$ and $\dot\gamma(0) = \dot\nu(0)$. Prove that this is an equivalence relation. A tangent vector at $p \in S$ is defined to an equivalence class of curves all with value p at $t = 0$. As a convenient notation, let us write $[\gamma]$ for the equivalence class containing the curve γ. The tangent space at p in S is defined to be the set of all equivalence classes of curves that have value p at $t = 0$. Prove that the tangent space at p defined in this manner can be given the structure of a vector space and this vector space has the same dimension as the manifold S. Also prove that this definition gives the same tangent space as defined in equation 1.20. Finally, for manifolds S_1 and S_2 and a function $F : S_1 \to S_2$, prove that the tangent map F_* is given by $F_*[\gamma] = [F \circ \gamma]$.

Exercise 1.111. Let A be an invertible symmetric $(n \times n)$-matrix. (a) Prove that the set $M := \{x \in \mathbb{R}^2 : \langle Ax, x\rangle = 1\}$ is a submanifold of $\mathbb{R}^n$. (b) Suppose that $x_0 \in M$. Describe the tangent space to M at x_0. Hint: Apply Exercise 1.110.

Exercise 1.112. [General Linear Group] The general linear group $GL(\mathbb{R}^n)$ is the set of all invertible real $n \times n$-matrices where the group structure is given by matrix multiplication (see also Exercise 2.55). (a) Prove that $GL(\mathbb{R}^n)$ is a submanifold of $\mathbb{R}^{n^2}$. Hint: Consider the determinant function. (b) Determine the tangent space of $GL(\mathbb{R}^n)$ at its identity. Hint: Apply Exercise 1.110. (c) Prove that the map $GL(\mathbb{R}^n) \times GL(\mathbb{R}^n) \to GL(\mathbb{R}^n)$ given by $(A, B) \mapsto AB$ is smooth. (d) Prove that the map $GL(\mathbb{R}^n) \to GL(\mathbb{R}^n)$ given by $A \mapsto A^{-1}$ is smooth. Note: A Lie group is a group that is also a smooth manifold such that the group operations are smooth. The vector space $T_I GL(\mathbb{R}^n)$ is called the Lie algebra of the Lie group when endowed with the multiplication $[A, B] = AB - BA$.

Exercise 1.113. (a) Prove that the tangent bundle of the torus $\mathbb{T}^2$ is trivial; that is, it can be viewed as $T\mathbb{T}^2 = \mathbb{T}^2 \times \mathbb{R}^2$. (b) (This exercise requires some knowledge of topology) Prove that the tangent bundle of $\mathbb{S}^2$ is not trivial.

Exercise 1.114. Suppose that $f : \mathbb{R}^n \to \mathbb{R}^n$ is smooth and the differential equation $\dot x = f(x)$ has a first integral all of whose level sets are compact. Prove that the corresponding flow is complete.

Exercise 1.115. Prove: The diagonal

$$\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : x = y\}$$

in $\mathbb{R}^n \times \mathbb{R}^n$ is an invariant set for the system

$$\dot x = f(x) + h(y - x), \qquad \dot y = f(y) + g(x - y)$$

where $f, g, h : \mathbb{R}^n \to \mathbb{R}^n$ and $g(0) = h(0)$.

Exercise 1.116. [An Open Problem in Structural Stability] Let $H(x, y, z)$ be a homogeneous polynomial of degree n and η the outer unit normal on the unit sphere $\mathbb{S}^2 \subset \mathbb{R}^3$. Show that the vector field $X_H = \operatorname{grad} H - nH\eta$ is tangent to $\mathbb{S}^2$.

Call a rest point *isolated* if it is the unique rest point in some open set. Prove that if n is fixed, then the number of isolated rest points of X_H is uniformly bounded over all homogeneous polynomials H of degree n. Suppose that $n = 3$, the uniform bound for this case is B, and m is an integer such that $0 \le m \le B$. What is B? Is there some H such that X_H has exactly m rest points? If not, then for which m is there such an H? What if $n > 3$?

Note that the homogeneous polynomials of degree n form a finite dimensional vector space $\mathcal{H}_n$. What is its dimension? Is it true that for an open and dense subset of $\mathcal{H}_n$ the corresponding vector fields on $\mathbb{S}^2$ have only hyperbolic rest points?

In general, if X is a vector field in some class of vector fields $\mathcal{H}$, then X is called *structurally stable* with respect to $\mathcal{H}$ if X is contained in some open subset $U \subset \mathcal{H}$ such that the phase portrait of every vector field in U is the same; that is, if Y is a vector field in U, then there is a homeomorphism of the phase space that maps orbits of X to orbits of Y. Let us define $\mathcal{X}_n$ to be the set of all vector fields on $\mathbb{S}^2$ of the form X_H for some $H \in \mathcal{H}_n$. It is an interesting unsolved problem to determine the structurally stable vector fields in $\mathcal{X}_n$ with respect to $\mathcal{X}_n$.

One of the key issues that must be resolved to determine the structural stability of a vector field on a two-dimensional manifold is the existence of *heteroclinic orbits*. A heteroclinic orbit is an orbit that is contained in the stable manifold of a saddle point q and in the unstable manifold of a different saddle point p. If $p = q$, such an orbit is called *homoclinic*. A basic fact from the theory of structural stability is that if two saddle points are connected by a heteroclinic orbit, then the local phase portrait near this orbit can be changed by an arbitrarily small smooth perturbation. In effect, a perturbation can be chosen such that, in the phase portrait of the perturbed vector field, the saddle connection is broken (see Figure 1.15). Thus, in particular, a vector field with two saddle points connected by a heteroclinic orbit is not structurally stable with respect to the class of all smooth vector fields. Prove that a vector field X_H in $\mathcal{X}_n$ cannot have a homoclinic orbit. Also, prove that X_H cannot have a periodic orbit. Construct a homogeneous

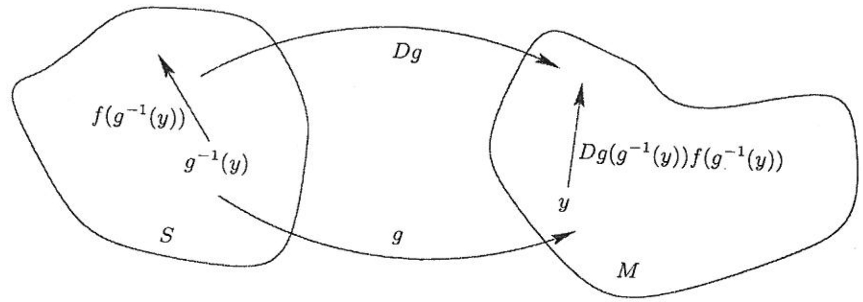

Figure 1.16: The "push forward" of a vector field f by a diffeomorphism $g : S \to M$.

polynomial $H \in \mathcal{H}_3$ such that X_H has hyperbolic saddle points p and q connected by a heteroclinic orbit.

Is every heteroclinic orbit of a vector field $X_H \in \mathcal{X}_3$ an arc of a great circle? The answer to this question is not known. But if it is true that all heteroclinic orbits are arcs of great circles, then the structurally stable vector fields, with respect to the class $\mathcal{X}_3$, are exactly those vector fields with all their rest points hyperbolic and with no heteroclinic orbits. Moreover, this set is open and dense in $\mathcal{X}_n$. A proof of these facts requires some work. But the main idea is clear: if X_H has a heteroclinic orbit that is an arc of a great circle, then there is a homogeneous polynomial K of degree $n = 3$ such that the perturbed vector field $X_{H+\epsilon K}$ has no heteroclinic orbits for $|\epsilon|$ sufficiently small. In fact, K can be chosen to be of the form

$$K(x, y, z) = (ax + by + cz)(x^2 + y^2 + z^2)$$

for suitable constants a, b, and c. (Why?) Of course, the conjecture that heteroclinic orbits of vector fields in $\mathcal{H}_3$ lie on great circles is just one approach to the structural stability question for $\mathcal{X}_3$. Can you find another approach?

There is an extensive and far-reaching literature on the subject of structural stability (see, for example, [192] and [204]).

1.8.4 Change of Coordinates

The proof of Proposition 1.107 contains an important computation that is useful in many other contexts; namely, the formula for changing coordinates in an autonomous differential equation. To reiterate this result, suppose that we have a differential equation $\dot{x} = f(x)$ where $x \in \mathbb{R}^n$, and $S \subseteq \mathbb{R}^n$ is an invariant k-dimensional submanifold. If g is a diffeomorphism from S to some k-dimensional submanifold $M \subseteq \mathbb{R}^m$, then the ordinary differential equation (or, more precisely, the vector field associated with the differential equation) can be "pushed forward" to M. In fact, if $g : S \to M$ is the diffeomorphism, then

$$\dot{y} = Dg(g^{-1}(y))f(g^{-1}(y)) \tag{1.23}$$

is a differential equation on M. Since g is a diffeomorphism, the new differential equation is the same as the original one up to a change of coordinates as schematically depicted in Figure 1.16.

Example 1.117. Consider $\dot{x} = x - x^2$, $x \in \mathbb{R}$. Let $S = \{x \in \mathbb{R} : x > 0\}$, $M = S$, and let $g : S \to M$ denote the diffeomorphism defined by $g(x) = 1/x$. Here, $g^{-1}(y) = 1/y$ and

$$\begin{aligned}\dot{y} &= Dg(g^{-1}(y))f(g^{-1}(y))\\ &= -\Big(\frac{1}{y}\Big)^{-2}\Big(\frac{1}{y} - \frac{1}{y^2}\Big)\\ &= -y + 1.\end{aligned}$$

The diffeomorphism g defines the change of coordinates $y = 1/x$ used to solve this special form of Bernoulli's equation; it is encountered in elementary courses on differential equations.

Exercise 1.118. According to the Hartman-Grobman theorem 1.47, there is a homeomorphism (defined on some open neighborhood of the origin) that maps orbits of $\dot{y} = y$ to orbits of $\dot{x} = x - x^2$. In this case, the result is trivial; the homeomorphism h given by $h(y) = y$ satisfies the requirement. For one and two-dimensional systems (which are at least twice continuously differentiable) a stronger result is true: There is a diffeomorphism h defined on a neighborhood of the origin with $h(0) = 0$ such that h transforms the linear system into the nonlinear system. Find an explicit formula for h and describe its domain.

Exercise 1.119. [Bernoulli's Equation] Show that the differential equation

$$\dot{x} = g(t)x - h(t)x^n$$

is transformed to a linear differential equation by the change of coordinates $y = 1/x^{n-1}$.

Coordinate transformations are very useful in the study of differential equations. New coordinates can reveal unexpected features. As a dramatic example of this phenomenon, we will show that all autonomous differential equations are the same, up to a smooth change of coordinates, near each of their regular points. Here, a *regular point* of $\dot{x} = f(x)$ is a point $p \in \mathbb{R}^n$, such that $f(p) \neq 0$. The following precise statement of this fact, which is depicted in Figure 1.17, is called the *rectification lemma*, the *straightening out theorem*, or the *flow box theorem*.

Lemma 1.120 (Rectification Lemma). *Suppose that $\dot{x} = f(x)$, $x \in \mathbb{R}^n$. If $p \in \mathbb{R}^n$ and $f(p) \neq 0$, then there are open sets U, V in $\mathbb{R}^n$ with $p \in U$, and a diffeomorphism $g : U \to V$ such that the differential equation in the new coordinates, that is, the differential equation*

$$\dot{y} = Dg(g^{-1}(y))f(g^{-1}(y)),$$