CHAPTER 9

Simple dynamics as a tool

In this chapter, the progression of the course from simple to complicated dynamics comes full circle. The tools from Chapter 2 are now put to use systematically, culminating in the proof of the stable manifold theorem, which is the fundamental result for the analysis of those highly complex systems that appeared in Chapter 7.

1. Introduction

a. Applications of the Contraction Principle. The collection of simple dynamical systems with complicated orbit structure presented in Chapter 7 and revisited in Section 8.3 is representative of *hyperbolic* dynamical systems. Much of the core theory of hyperbolic dynamics consists of results that are obtained (more or less) directly from the Contraction Principle, which first appeared as an example of a dynamical system with simple dynamics in Chapter 2. Although we already used it in Section 2.5 as a tool that can tell us much about other dynamical systems, its pervasive role in hyperbolic dynamics motivates a more thorough presentation of its uses. Accordingly, the main theme of this chapter is to present case studies of using the Contraction Principle, that is, of putting one important insight about a specific class of simple dynamics to use in an auxiliary space to tell us about analysis as well as (complicated) dynamical systems. Since the results we obtain are rather important we take some time to develop them further, notably when it comes to the basic theory of differential equations. In this chapter we maintain the same standard of proof as in the course.

As in the preceding chapters this intrinsically interesting development has a utilitarian undercurrent. The results obtained here are important for the study of dynamical systems. In the case of existence and uniqueness of solutions of differential equations this is evident. But all other results presented here also figure in our development, and are standard tools in dynamics. This chapter does not present nearly all such applications, but some others are presented elsewhere, such as the Anosov Closing Lemma (Theorem 10.2.2) which follows from the Contraction Principle by way of the Hyperbolic Fixed Point Theorem 9.5.4. The Stable Manifold Theorem 9.5.2 is the foremost example and is featured prominently here.

b. Overview. We begin by deriving two important results in analysis, the Inverse and Implicit Function Theorems. The latter immediately tells us something new about the Contraction Principle itself: The fixed point of a contraction depends smoothly on the contraction. A first and straightforward application of these results is persistence of transverse fixed points in Section 9.3, where we show that a simple condition on the linear part of a map at a fixed point can guarantee that the fixed point persists when the map is perturbed. This is similar to the situation for contractions (Proposition 2.2.20) and very much in the spirit of linearization (which is discussed in Section 2.1, the beginning of Chapter 3, and, for example, Section 6.2b7). In these first applications the space on which the contraction is defined is the same space in which the problem is posed. However, the applications of the Contraction Principle to existence and uniqueness of solutions of differential equations (Section 9.4) and in the theorem on stable manifolds (Section 9.5), like many other important applications in analysis, use the Contraction Principle by reducing the situation at hand to a search for a fixed point in a space of functions, which has infinite dimension, rather than in a euclidean space.

c. Creating a context for the Contraction Principle. While the common feature is the application of the Contraction Principle in some auxiliary space, the degree of cleverness required to set this up varies in these examples. Picard iteration (Section 9.4) is a straightforward application, even though the space in question is not as simple as in the earlier applications. Of course, this is also the oldest example. The initial step in the proof of the Inverse Function Theorem 9.2.2 requires more creativity, but is close to the Newton method. The proof of persistence of transverse fixed points (Proposition 9.3.1) has no equally obvious motivation for the initial step, but it exhibits some features common to other applications of the Contraction Principle in dynamics. The central point is the combination of transversality and closeness (smallness of a perturbation), which is being used to produce an invertible map by transversality whose inverse is composed with a strongly contracting map arising from the perturbation. (The trick is to do this in such a way that the desired object is a fixed point of the resulting contraction.)

Except for Picard iteration all applications of the Contraction Principle in this chapter depend on linearization. This, too, is typical of applications in the theory of smooth dynamical systems.

2. Implicit and Inverse Function Theorems in euclidean space

a. The Inverse Function Theorem. The inverse function theorem says that if a differentiable map has invertible derivative at some point then the map is invertible near that point. This result is related to linearization: If we assume a certain qualitative ("yes-no") fact about the linear part (invertibility) then it holds for the nonlinear map itself—at least in a neighborhood. The version for the real line is familiar from calculus:

THEOREM 9.2.1. *Suppose $I \subset \mathbb{R}$ is an open interval and $f\colon I \to \mathbb{R}$ a differentiable function. If $a \in I$ is such that $f'(a) \neq 0$ and f' is continuous at a then f is invertible on a neighborhood U of a and $(f^{-1})'(y) = 1/f'(x)$, where $y = f(x)$.*

Usually one thinks of invertibility as the easy part and the derivative formula as the hard one, because the basic calculus examples of invertible real-valued functions are given by formulas where invertibility is rather apparent. However, the main content of this result is to conclude invertibility from knowledge only of the linear part of a map *at one point*, without any such extra information. The derivative formula is then an easy postlude. We even get higher derivatives easily. This result is fairly easy in $\mathbb{R}^n$ as well, however, we first give a proof for the simple case of a single variable.

PROOF. Given y, we want to solve the equation $f(x) = y$ for x, which is the same as to find a root of $F_y(x) := y - f(x)$. To this end we first set up a suitable contracting map. *The space*. The space on which the contraction acts is the real line.

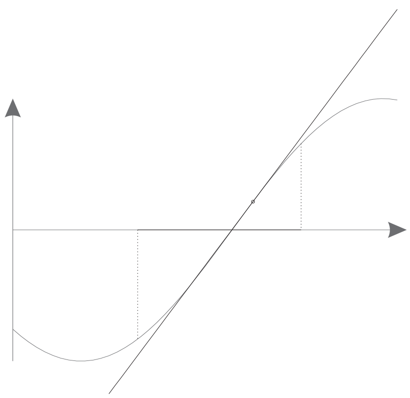

FIGURE 9.2.1. The Inverse Function Theorem

Defining the contraction. The Newton method Section 2.2h suggests to make an initial guess x (where y is fixed for the moment) and improve the guess by repeatedly applying the map

$$F_y(x) = x - \frac{F_y(x)}{F_y'(x)} = x + \frac{y - f(x)}{f'(x)}.$$

To verify that this is a contraction involves taking and estimating the second derivative of f, but we don't assume it exists. It is convenient to instead consider the map

$$\varphi_y(x) := x + \frac{y - f(x)}{f'(a)}$$

on I. Its fixed points are solutions of our problem because $\varphi_y(x) = x$ if and only if

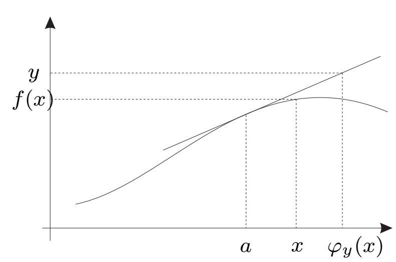

FIGURE 9.2.2. φ_y

$f(x) = y$.

The contraction property. Now we show that φ_y is a contraction of some closed subset of O. Then by the Contraction Principle it has a unique fixed point and hence there exists a unique x such that $f(x) = y$.

To that end let $A := f'(a)$ and $\alpha = |A|/2$. By continuity of f' at a there is an $\epsilon > 0$ such that $W := (a - \epsilon, a + \epsilon) \subset I$ and $|f'(x) - A| < \alpha$ for x in the closure $\overline{W}$ of W.

To see that φ_y is a contraction on $\overline{W}$ note that if $x \in \overline{W}$ then

$$|\varphi_y'(x)| = |1 - \frac{f'(x)}{A}| = |\frac{A - f'(x)}{A}| < \frac{\alpha}{|A|} = 1/2.$$

Using Proposition 2.2.3 we get $|\varphi_y(x) - \varphi_y(x')| \le |x - x'|/2$ for $x, x' \in W$.

We need to show also that $\varphi_y(\overline{W}) \subset \overline{W}$ for y sufficiently close to $b := f(a)$. Let $\delta = A\epsilon/2$ and $V = (b - \delta, b + \delta)$. Then for $y \in V$ we have

$$|\varphi_y(a) - a| = |a + \frac{y - f(a)}{A} - a| = |\frac{y - b}{A}| < |\frac{\delta}{a}| = \frac{\epsilon}{2},$$

so if $x \in \overline{W}$ then

$$|\varphi_y(x) - a| \le |\varphi_y(x) - \varphi_y(a)| + |\varphi_y(a) - a| < \frac{x - a}{2} + \frac{\epsilon}{2} \le \epsilon$$

and hence $\varphi_y(x) \in W$.

Therefore Proposition 2.2.21 applied to $\varphi_y : \overline{W} \to \overline{W}$ for $y \in V$ produces a unique fixed point $g(y) \in W$, which depends continuously on y.

Next we prove that the inverse is differentiable. For $y = f(x) \in V$ we want to show that $g'(y)$ exists and is the reciprocal of $B := f'(g(y))$.

Let $U := g(V) = W \cap f^{-1}(V)$ (the preimage under f), so U is open. Take $y + k = f(x + h) \in V$. Then

$$\frac{|h|}{2} \ge |\varphi_y(x + h) - \varphi_y(x)| = |h + \frac{f(x) - f(x + h)}{A}| = |h - \frac{k}{A}| \ge |h| - |\frac{k}{A}|$$

hence

$$\frac{|h|}{2} \le |\frac{k}{A}| < \frac{|k|}{\alpha} \quad \text{and} \quad \frac{1}{|k|} < \frac{2}{\alpha|h|}.$$

Since $g(y + k) - g(y) - k/B = h - k/B = -(f(x + h) - f(x) - Bh)/B$ we therefore get

$$\frac{|g(y + k) - g(y) - k/B|}{|k|} < \frac{2}{|B|\alpha} \frac{|f(x + h) - f(x) - Bh|}{|h|} \xrightarrow[|h| \le 2|k|/\alpha \to 0]{} 0,$$

which proves $g'(y) = 1/B = 1/f'(g(y))$.

Finally, suppose $f \in C^r$. We show inductively that $g \in C^r$. To that end assume $g \in C^k$ for some $k < r$ (we start the induction with $k = 0$). Then $f'(g(y)) \in C^k$ and so is its reciprocal g'. Thus, $g \in C^{k+1}$. □

Now we adapt this argument to $\mathbb{R}^n$:

THEOREM 9.2.2 (Inverse function theorem). *Suppose $O \subset \mathbb{R}^m$ is open, $f : O \to \mathbb{R}^m$ is differentiable and that Df is invertible at a point $a \in O$ and continuous at a. Then there exist neighborhoods $U \subset O$ of a and V of $b := f(a) \in \mathbb{R}^m$ such that f is a bijection from U to V (that is, f is one-to-one on U and $f(U) = V$). The inverse $g : V \to U$ of f is differentiable with $Dg(y) = (Df(g(y)))^{-1}$. Furthermore, if f is C^r (that is, all partial derivatives of f up to order r exist and are continuous) on U, then so is its inverse.*

PROOF. The proof is actually the same as before. We only need to replace various numbers by linear maps, and some absolute values by norms.
The space. The contraction acts in $\mathbb{R}^m$.

The map. For any given $y \in \mathbb{R}^m$ consider the map

$$\varphi_y(x) := x + Df(a)^{-1}(y - f(x))$$

on O. Notice that $\varphi_y(x) = x$ if and only if $f(x) = y$, so we try to find a unique fixed point for φ_y. We need a set W on which it is a contraction.
The contraction property. Let $A := Df(a)$, $\alpha < \|A^{-1}\|^{-1}/2$, and, using continuity of Df at a, take $\epsilon > 0$ such that $\|Df(x) - A\| < \alpha$ for x in the closure of $W := B(a, \epsilon)$. To see that φ_y is a contraction note that

$$\|D\varphi_y(x)\| = \|\operatorname{Id} - A^{-1}Df(x)\| = \|A^{-1}(A - Df(x))\| < \|A^{-1}\|\alpha =: \lambda < 1/2$$

for $x \in W$ and apply Corollary 2.2.15 to get $\|\varphi_y(x) - \varphi_y(x')\| \leq \lambda\|x - x'\|$ for $x, x' \in W$. Therefore by Proposition 2.2.20 there is a neighborhood V of b such that φ_y is a contraction of $\overline{W}$ for all $y \in V$ and has a unique fixed point $g(y) \in W$ (which depends continuously on y). $U := g(V) = W \cap f^{-1}(V)$ is open.

The determinant of $Df(x)$ depends continuously on Df and hence is continuous at a as a function of x. Thus, by taking V (and hence U) smaller, if necessary, we may assume $\det Df \neq 0$ on U and therefore that $Df(x)$ is invertible on U.

For $y = f(x) \in V$ we want to show that $Dg(y)$ exists and is the inverse of $B := Df(g(y))$. Take $y + k = f(x + h) \in V$. Then

$$\text{(9.2.1)} \quad \frac{\|h\|}{2} \geq \|\varphi_y(x + h) - \varphi_y(x)\| = \|h + A^{-1}(f(x) - f(x + h))\| \\ = \|h - A^{-1}k\| \geq \|h\| - \|A^{-1}\|\|k\|,$$

so

$$\frac{\|k\|}{\alpha} > \|A^{-1}\|\|k\| \geq \frac{\|h\|}{2} \quad \text{and} \quad \frac{1}{\|k\|} < \frac{2}{\alpha\|h\|}.$$

Since $g(y + k) - g(y) - B^{-1}k = h - B^{-1}k = -B^{-1}(f(x + h) - f(x) - Bh)$ we get

$$\frac{\|g(y + k) - g(y) - B^{-1}k\|}{\|k\|} < \frac{\|B^{-1}\|}{\alpha/2} \frac{\|f(x + h) - f(x) - Bh\|}{\|h\|} \xrightarrow[\|h\| \leq 2\|k\|/\alpha \to 0]{} 0,$$

which proves $Dg(y) = B^{-1}$.

Finally, suppose $f \in C^r$ and $g \in C^k$ for some $k < r$. Then $Df(g(y)) \in C^k$ and so is its inverse Dg by using the formula for matrix inverses (the entries of A^{-1} are polynomials in those of A divided by $\det A \neq 0$). Thus, $g \in C^{k+1}$. □

b. The Implicit Function Theorem. A result closely related to the Inverse Function Theorem is the Implicit Function Theorem. It follows easily from the Inverse Function Theorem and is therefore indirectly an application of the Contraction Principle. Furthermore, as we see in the next subsection, it immediately tells us more about the Contraction Principle itself regarding the dependence of the fixed point of a contraction on the contraction (see also Figure 2.2.3).

Like the Inverse Function Theorem the Implicit Function Theorem transfers information about the linear part of a map to the map itself. To see how, consider the question answered by the Implicit Function Theorem in the case of a linear map. Suppose $A\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is a linear map and write it as $A = (A_1, A_2)$, where $A_1\colon \mathbb{R}^n \to \mathbb{R}^n$ and $A_2\colon \mathbb{R}^m \to \mathbb{R}^n$ are linear. Suppose we pick $k \in \mathbb{R}^m$ and want to find $h \in \mathbb{R}^n$ such that $A(h,k)=0$. To see when this can be done rewrite this as $A_1h+A_2k=0$ to conclude that if A_1 is invertible then

$$A(h,k)=0 \Leftrightarrow h=-(A_1)^{-1}A_2k. \tag{9.2.2}$$

One can interpret this as saying that the equation $A(h,k)=0$ implicitly defines a map $h=Lk$ such that $A(Lk,k)=0$. The Implicit Function Theorem says that if this is true for the linear part of a map, then it is true for the map itself: Under some assumptions corresponding to invertibility of A_1, the equation $f(x,y)=0$ implicitly defines a map $x=g(y)$ such that $f(g(y),y)=0$. To state those assumptions for a map $f\colon \mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ we write $Df=(D_1f,D_2f)$ analogously to the above, with $D_1f\colon\mathbb{R}^n\to\mathbb{R}^n$ and $D_2f\colon\mathbb{R}^m\to\mathbb{R}^n$.

The Implicit Function Theorem gives an analogous statement for nonlinear maps. It tells us that if we can solve an equation given a particular value of a parameter then there is a solution for nearby parameter values as well.

THEOREM 9.2.3 (Implicit Function Theorem). *Let $O\subset\mathbb{R}^n\times\mathbb{R}^m$ be open and $f\colon O\to\mathbb{R}^n$ a C^r map. If there is a point $(a,b)\in O$ such that $f(a,b)=0$ and $D_1f(a,b)$ is invertible then there are open neighborhoods $U\subset O$ of (a,b), $V\subset\mathbb{R}^m$ of b such that for every $y\in V$ there exists a unique $x=:g(y)\in\mathbb{R}^n$ with $(x,y)\in U$ and $f(x,y)=0$. Furthermore g is C^r and $Dg(b)=-(D_1f(a,b))^{-1}D_2f(a,b)$.*

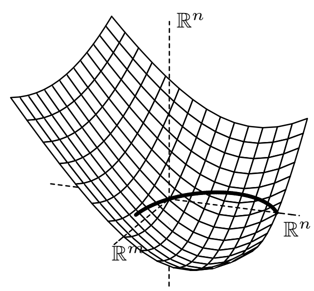

FIGURE 9.2.3. Implicit function theorem

PROOF. $F(x,y):=(f(x,y),y)\colon O\to\mathbb{R}^n\times\mathbb{R}^m$ is C^r and if $A=Df(a,b)$ then $DF(a,b)\,(h,k)=(A(h,k),k)$ by the chain rule. This gives zero only if $k=0$ and $A(h,k)=0$, hence $(h,k)=0$ by (9.2.2). Therefore DF is invertible and by the Inverse Function Theorem 9.2.2 there are open neighborhoods $U\subset O$ of (a,b) and $W\subset\mathbb{R}^n\times\mathbb{R}^m$ of $(0,b)$ such that $F\colon U\to W$ is invertible with C^r inverse $G=F^{-1}\colon W\to U$. Thus, for any $y\in V:=\{y\in\mathbb{R}^m \mid (0,y)\in W\}$ there exists an $x=:g(y)\in\mathbb{R}^n$ such that $(x,y)\in U$ and $F(x,y)=(0,y)$, that is, $f(x,y)=0$.

Now $(g(y),y)=(x,y)=G(0,y)$ and hence g is C^r. To find $Dg(b)$ let $\gamma(y):=(g(y),y)$. Then $f(\gamma(y))\equiv 0$ and hence $Df(\gamma(y))D\gamma(y)=0$ by the chain rule. For $y=b$ this gives $D_1f(a,b)Dg(b)+D_2f(a,b)=Df(a,b)D\gamma(b)=0$, completing the proof. $\square$

c. The smooth Contraction Principle. Returning to dynamics we apply smoothness of the implicit function g to the Contraction Principle To show that the fixed point of a contraction depends smoothly on the contraction itself (see Figure 2.2.3). To express this we write our contractions as maps with a parameter.

THEOREM 9.2.4. *Suppose* $f\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ *is* C^r *and there exists a* $\lambda < 1$ *such that* $d(f(x,y), f(x',y)) \le \lambda d(x,x')$ *for* $x, x' \in X$. *Then for every* $y \in Y$ *there is a unique fixed point* $g(y)$ *of* $x \mapsto f(x,y)$ *and* g *is* C^r.

PROOF. Existence of the fixed point $g(y)$ follows from the Contraction Principle. Now write $F(x,y) := f(x,y) - x$ and notice that this is a C^r function that satisfies the hypotheses of the Implicit Function Theorem 9.2.3: It is zero at $(a,b) = (g(y),y)$ (any choice of y is fine here) and $\|D_1Fv\| = \|D_1fv - v\| \ge \|v\| - \|D_1fv\| \ge (1-\lambda)\|v\| > 0$ for $v \ne 0$, so D_1F is invertible. Thus $g \in C^r$. □

REMARK 9.2.5. Instead of the domain $\mathbb{R}^n \times \mathbb{R}^m$ one can take $A \times O$, where $O \subset \mathbb{R}^m$ is open and A is the closure of an open set, say. (One needs a closed set to apply the Contraction Principle, but a good enough set to be able to differentiate r times.

REMARK 9.2.6. Suppose f_λ depends smoothly on λ and $f := f_0$ is as in Proposition 2.2.20 (p. 39). Show that there is a smooth family $\lambda \mapsto x_\lambda$ with x_0 as in Proposition 2.2.20 and $f_\lambda(x_\lambda) = x_\lambda$.

3. Persistence of transverse fixed points

The fixed point of a contraction simultaneously exhibits two kinds of stability. As an attracting fixed point it is asymptotically stable. Proposition 2.2.20 and Proposition 2.6.14 (as well as Theorem 9.2.4) state that it is also stable under perturbations of the map, that is, perturbations of the map have a unique fixed point nearby. This is an important robustness property of the local dynamics, and we now use the Contraction Principle to describe a general condition under which an analogous conclusion holds. This is a straightforward and simple illustration of the use of the Contraction Principle and Implicit Function Theorem in dynamics where the Contraction Principle is applied to a derived system in the same space.

Recall that two C^1-maps f and g are C^1-close if $|f-g| + \|Df - Dg\|$ is uniformly small.

PROPOSITION 9.3.1. *If* p *is a periodic point of period* m *for a* C^1 *map* f *and the differential* Df_p^m *does not have one as an eigenvalue (in this case* p *is said to be a* transverse periodic point*) then for every map* g *sufficiently* C^1*-close to* f *there is a unique periodic point of period* m *close to* p.

Note that in dimension one the assumption on the derivative simply means that it is not one. Accordingly, in the example of the basic bifurcation of Section 2.3b (see Figure 2.3.2) the single fixed point appears or disappears exactly when there is a tangency with the diagonal, that is, the derivative of the map is one. Figure 9.3.1 illustrates this. The axis of the independent variable points right, the vertical axis is for the "output", and the axis towards the rear gives a parameter with which the map changes. The plane shows the diagonal for various parameters, and the graphs of perturbed maps combine to a surface that intersects the diagonals in the family of fixed points.

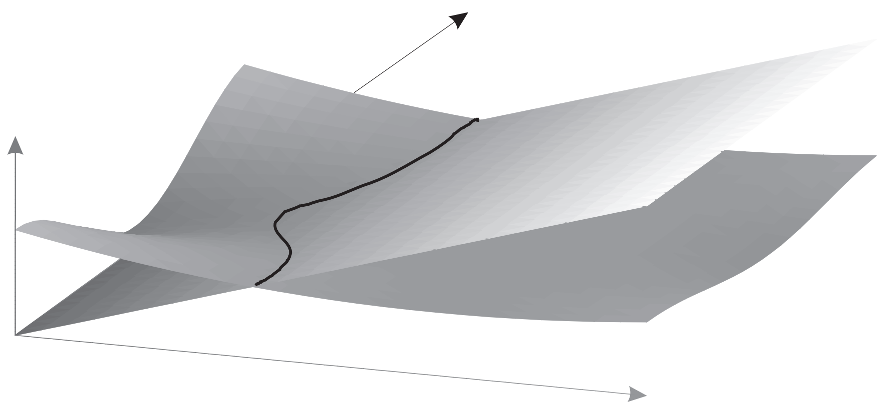

FIGURE 9.3.1. Persistence of a fixed point

PROOF. *The space*. We define a contraction in a neighborhood of p.
The map. Introduce local coordinates near p with p as the origin. In these coordinates Df_0^m becomes a matrix. Since 1 is not among its eigenvalues the map $F = f^m - \mathrm{Id}$ defined locally in these coordinates is locally invertible by the Inverse Function Theorem 9.2.2. Now let g be a map C^1-close to f. Near 0 one can write $g^m = f^m - H$ where H is small together with its first derivatives. A fixed point for g^m can be found from the equation $x = g^m(x) = (f^m - H)(x) = (F + \mathrm{Id} - H)(x)$ or $(F - H)(x) = 0$ or

$$x = F^{-1}H(x).$$

The contraction property. Since F^{-1} has bounded derivatives and H has very small first derivatives one can show that $F^{-1}H$ is a contracting map. More precisely, let $\|\cdot\|_0$ denote the C^0-norm, $\|dF^{-1}\|_0 = L$, and

$$\max(\|H\|_0, \|dH\|_0) \le \epsilon.$$

Then, since $F(0) = 0$, we get $\|F^{-1}H(x) - F^{-1}H(y)\| \le \epsilon L\|x - y\|$ for every x, y close to 0 and $\|F^{-1}H(0)\| \le L\|H(0)\| \le \epsilon L$, and hence $\|F^{-1}H(x)\| \le \|F^{-1}H(x) - F^{-1}H(0)\| + \|F^{-1}H(0)\| \le \epsilon L\|x\| + \epsilon L$. Thus if $\epsilon \le \dfrac{R}{L(1+R)}$ the disc $X = \{x \mid \|x\| \le R\}$ is mapped by $F^{-1}H$ into itself and the map $F^{-1}H \colon X \to X$ is contracting. By the Contraction Principle it has a unique fixed point in X, which is thus a unique fixed point for g^m near 0. $\square$

REMARK 9.3.2. It is easy to show that a transverse fixed point is isolated.

REMARK 9.3.3. If f is a C^1 map with a hyperbolic fixed point p, that is, $Df_{|_p}$ has no eigenvalues on the unit circle, and g is sufficiently C^1-close to f then g has a unique fixed point near p and this fixed point is a hyperbolic fixed point of g.

4. Solutions of differential equations

Differential equations are a natural setting in which dynamical issues arise, and they appear in several important contexts. At the basis of the use of differential equations in