GENERALIZED IMPLICIT FUNCTION THEOREMS AND APPLICATIONS TO ORDINARY DIFFERENTIAL EQUATIONS

MVUYISI MDEKAZI

June 1993

A thesis submitted in fulfilment of the requirements for the degree of

Master of Science

Supervisor: Professor R.I. Becker

Department of Mathematics, University of Cape Town

Typeset by $\mathcal{AMS}$-TEX

Acknowledgements

The warm encouragement and the valuable suggestions from Dr. E. Brüning meant much to me. I would like to thank him for the limitless patience with which he kept explaining some concepts to me.

My heartfelt thanks go to Professor R.I. Becker. He suggested the topic and cheerfully supervised the transformation of the manuscript into a thesis. This thesis would not have come to fruition without his help.

I have been fortunate in having the help of an efficient, patient and excellent typist. I am grateful to Wanda Fouquet for spending gruelling hours in typing the various versions of this thesis.

Contents

Chapter 0. NOTATION, DEFINITIONS AND SOME PRELIMINARY RESULTS

0.1 Some Facts from Functional Analysis

In the discussion below we use $\mathbb{K}$ to denote $\mathbb{R}$ or $\mathbb{C}$.

Here we list definitions and theorems used in Chapter 1 to Chapter 4. The letters X, Y and Z will stand for Banach spaces over $\mathbb{K}$ and F, G, H will stand for mappings between these spaces.

Definition 0.1.1. Let X, Y be two Banach spaces. Then a mapping $F : X \to Y$ is called a linear operator of X into Y if

(i) $F(x + y) = Fx + Fy$, for all $x, y \in X$

(ii) $F(\alpha x) = \alpha Fx$, for all $\alpha \in K$ and $x \in X$.

Definition 0.1.2. Let $F : X \to Y$ be a linear operator. F is said to be bounded if there is a real number $k > 0$ such that $\|Fx\| \leq k \|x\|$, for all $x \in X$.

Definition 0.1.3. The norm of a bounded linear operator F is defined to be $\|F\| = \sup_{\|x\|_X \leq 1} \|Fx\|_Y$.

Theorem 0.1.4. (cf.[19],p.144) A linear operator $F : X \to Y$ is bounded if and only if it is continuous. By continuity we mean that $\lim_{n\to\infty} x_n = x$ in X implies $\lim_{n\to\infty} Fx_n = Fx$ in Y.

Definition 0.1.5. Let U and V be subspaces of X. X is said to be the direct sum of U and V, written $X = U \oplus V$, if and only if

(i) $X = U + V$

(ii) $U \cap V = \{0\}$.

In this case V is called a direct complement of U in X and written U'. If $X = U \oplus V$ then any element x in X can be written in the form $x = u + v$, $u \in U$, $v \in V$, where $u = u_x \in U$ and $v = v_x \in V$ are uniquely determined in X.

Definition 0.1.6. Let X have the direct sum decomposition $X = U \oplus V$ and let $P : X \to X$ be a linear operator satisfying

(i) $Pu = u$, for all $u \in U$

(ii) $Pv = 0$, for all $v \in V$.

Then P is called the projection of X onto U along V. Since every $x \in X$ can be written uniquely as $x = u + v$ we see that $x = Px + v$ so that $v = (I - P)x$ implying that $I - P$ is the projection of X onto V along U. Clearly $(I - P)P = P(I - P) = 0$. Hence $P = P^2$.

Theorem 0.1.7. (cf.[19],p.43) If a bounded linear operator $P : X \to X$ is idempotent ($P^2 = P$), then there exists subspaces U and V such that $X = U \oplus V$ and P is the projection of X onto U along V.

Definition 0.1.8. A linear operator $F : X \to \mathbb{K}$ is said to be a linear functional on X.

Let $\mathcal{L}(X, Y)$ be a set of bounded linear operators from X to Y.

Theorem 0.1.9. (cf.[24],p.43) If Y is a Banach space, so is $\mathcal{L}(X, Y)$.

Adjoint Operator:

Let X, Y be Banach spaces and $F \in \mathcal{L}(X, Y)$. Let X^* denote the space of bounded linear functionals on X.

Theorem 0.1.10. (cf.[24],p.39) X^* is a Banach space.

Now for each $y^* \in Y^*$, the expression $y^*(Fx)$ assigns to each $x \in X$ a scalar. Thus by definition 0.1.8 it is a functional $G(x)$. Clearly G is linear. Moreover it is bounded, since $|G(x)| = |y^*(Fx)| \leq \|y^*\| \, \|Fx\| \leq \|y^*\| \, \|F\| \, \|x\|$. There is a $x^* \in X^*$ such that $y^*(Fx) = x^*(x)$, $x \in X$. x^*, a functional, is unique. Thus to each $y^* \in Y^*$ there is assigned a unique $x^* \in X^*$. The map making this assignment is denoted by F^* and is a linear operator from Y^* to X^*. Hence $y^*(Fx) = F^*y^*(x)$. F^* is called the adjoint operator of F.

Theorem 0.1.11. (cf.[24],p.59) $F^* \in \mathcal{L}(Y^*, X^*)$ and $\|F^*\| = \|F\|$.

Denote by $N(F)$ the nullspace of F and by $R(F)$ the range of F. Note that if F is a linear operator from X to Y with finite dimensional range and nullspace, then the rank of F is the dimension of the range of F and the nullity of F is the dimension of the nullspace of F. Clearly $N(F)$ and $R(F)$ are subspaces of X and Y respectively since F is linear. Suppose $Fx = y$ where $x \in X$, $y \in Y$. For any $y^* \in Y^*$, $y^*(Fx) = y^*(y)$. Taking adjoints we get $F^*y^*(x) = y^*(y)$. If $y^* \in N(F^*)$ then $y^*(y) = 0$. We have: A necessary condition that $y \in R(F)$ is that $y^*(y) = 0$ for all $y^* \in N(F^*)$.

Annihilators:

Let S be a subset of X. A functional $x^* \in X^*$ is called an annihilator of S if $x^*(x) = 0$, $x \in S$. We denote by $S^\perp$ the set of annihilators of S. Also if T is a subset of X^*, we call an $x \in X$ an annihilator of T if $x^*(x) = 0$, $x^* \in T$. We denote by $T^\perp$ the set of annihilators of T. It follows from the previous paragraph that in terms of annihilators a necessary condition that $y \in R(F)$ is that $R(F) \subseteq N(F^*)^\perp$.

Theorem 0.1.12. (cf.[24],p.61) $S^\perp$ and $T^\perp$ are closed subspaces.

Theorem 0.1.13. cf.[19],p.54) If V and W are subspaces of X and $V^\perp$ and $W^\perp$ are their annihilators, then

(i) $(V + W)^\perp = V^\perp \cap W^\perp$;
(ii) $(V \cap W)^\perp = V^\perp + W^\perp$;
(iii) if $W^\perp$ is of finite dimension k then W is of codimension k.

Theorem 0.1.14. (cf.[24],p.62,p.74) $R(F) = N(F^*)^\perp$ if and only if $R(F)$ is closed in Y. If $R(F)$ is closed in Y then $R(F^*) = N(F)^\perp$ and hence closed in X^*. Furthermore $R(F)^\perp = N(F^*)$ and $R(F^*)^\perp = N(F)$.

Inverse Operator:

Suppose $Fx = y$ where $F \in \mathcal{L}(X, Y)$.

Theorem 0.1.15. (cf.[24],p.63) If X, Y are Banach spaces, and $F \in \mathcal{L}(X, Y)$ with $R(F) = Y$, $N(F) = \{0\}$, then $F^{-1} \in \mathcal{L}(Y, X)$.

Fredholm Operator:

Let X, Y be Banach spaces. An operator $F \in \mathcal{L}(X,Y)$ is said to be a Fredholm operator from X to Y if

(i) $\dim N(F)$ is finite

(ii) $\operatorname{codim} R(F)$ is finite

(iii) $R(F)$ is closed.

The index of a Fredholm operator is defined by $i(F) = \dim N(F) - \operatorname{codim} R(F)$.

Theorem 0.1.16. (cf.[6],p.322) Let N be a finite dimensional linear subspace of a normed space X. Then there is a closed linear subspace M of X with $X = N \oplus M$.

Theorem 0.1.17. (cf.[24],p.108) Let X be a normed space and let N be a closed subspace such that $N^{\perp}$ is of finite dimension n. Then, there is an n–dimensional subspace M of X such that $X = N \oplus M$.

Theorem 0.1.18. (cf.Theorem 0.1.20) Let X, Y be Banach spaces and $L \in \mathcal{L}(X,Y)$ be a Fredholm operator. Then there exists closed complementary subspaces M and N of X and Y respectively such that

$$X = N(L) \oplus M$$

$$Y = N \oplus R(L).$$

Theorem 0.1.19. (cf.[6],p331) Suppose that L is a Fredholm operator. Then there exists $\delta > 0$ such that, if $K \in \mathcal{L}(X,Y)$ and $\|K\| < \delta$, then $L + K$ is a Fredholm operator, $\dim N(L+K) \leq \dim N(L)$, $\operatorname{codim}(R(L+K)) \leq \operatorname{codim}(R(L))$ and index $(L+K) =$ index (L).

Note: Let L be a Fredholm operator. Look back at Theorem 0.1.18 and note that L is one to one on M. Now let $\hat{L}$ be the restriction of L to M. Then $\hat{L} \in \mathcal{L}(M, R(L))$. Moreover $R(\hat{L}) = R(L)$: $y \in R(L)$ implies there is $x \in X$ such that $Lx = y$. But $x = x_0 + x_1$, where $x_0 \in M$, $x_1 \in N(L)$. $Lx_0 = Lx - Lx_1 = y$ showing that $y \in R(\hat{L})$. Thus $\hat{L}$ has an inverse $\hat{L}^{-1} \in \mathcal{L}(R(L), M)$.

Theorem 0.1.20. (cf.[24],p.108) If L is a Fredholm operator from X to Y (Banach spaces), there is a closed subspace X_0 of X such that $X = X_0 \oplus N(L)$ and a subspace Y_0 of Y of dimension $\beta(L) = \dim N(L^*)$ such that $Y = R(L) \oplus Y_0$. Moreover there is an operator $L_0 \in \mathcal{L}(Y, X)$ such that

$$N(L_0) = Y_0$$
$$R(L_0) = X_0$$
$$L_0 L = I \text{ on } X_0$$
$$L L_0 = I \text{ on } R(L).$$

In addition

$$L_0 L = I - F_1 \text{ on } X$$
$$L L_0 = I - F_2 \text{ on } Y$$

where $F_1 \in \mathcal{L}(X, X)$ with $R(F_1) = N(L)$ and $F_2 \in \mathcal{L}(Y, Y)$ with $R(F_2) = Y_0$. Consequently the operators F_1 and F_2 are of finite rank.

Theorem 0.1.21. Let X be a Banach space and let S, T, N be subspaces such that

(i) $X = T \oplus N$

(ii) $\dim S = \dim T = m < \infty$

(iii) $S \cap N = \{0\}$.

Then $X = S \oplus N$.

Proof. Let $\{t_1, \ldots, t_m\}$ be a basis of T and let $\{s_1, \ldots, s_m\}$ be a basis of S. Then there exist $n_1, \ldots, n_m \in N$ and $(a_{ij})_{1 \leq i,j \leq m}$ such that

$$s_1 = (a_{11}t_1 + \cdots + a_{1m}t_m) + n_1$$

$$\vdots$$

$$s_m = (a_{mm}t_1 + \cdots + a_{mm}t_m) + n_m$$

or in matrix form

$$\begin{pmatrix} s_1 \\ \vdots \\ s_m \end{pmatrix} = A \begin{pmatrix} t_1 \\ \vdots \\ t_m \end{pmatrix} + \begin{pmatrix} n_1 \\ \vdots \\ n_m \end{pmatrix}. \tag{0.1.1}$$

We show that A is non–singular:

Suppose not. Then A^* is also singular, and there exists $c = \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix} \neq 0$ such that

$$A^* c = 0$$

or

$$c^* A = 0.$$

Multiplying (0.1.1) by c^* gives

$$c^* \begin{pmatrix} s_1 \\ \vdots \\ s_m \end{pmatrix} = c^* A \begin{pmatrix} t_1 \\ \vdots \\ t_m \end{pmatrix} + c^* \begin{pmatrix} n_1 \\ \vdots \\ n_m \end{pmatrix}$$

$$= c^* \begin{pmatrix} n_1 \\ \vdots \\ n_m \end{pmatrix}.$$

That is

$$c_1 s_1 + \cdots + c_m s_m = c_1 n_1 + \cdots + c_m n_m \in N.$$

But $c_1s_1 + \cdots + c_ms_m \in S$, so by (iii), $c_1n_1 + \cdots + c_mn_m = 0$ (since both sides belong to $S \cap N$). Thus

$$c_1s_1 + \cdots + c_ms_m = 0$$

and $c \neq 0$, contradicting linear independence of $\{s_1, \ldots, s_m\}$. Hence A is nonsingular.

From (0.1.1)

$$A^{-1}\begin{pmatrix} s_1 \\ \vdots \\ s_m \end{pmatrix} = \begin{pmatrix} t_1 \\ \vdots \\ t_m \end{pmatrix} + A^{-1}\begin{pmatrix} n_1 \\ \vdots \\ n_m \end{pmatrix}$$

or

$$\begin{pmatrix} t_1 \\ \vdots \\ t_m \end{pmatrix} = A^{-1}\begin{pmatrix} s_1 \\ \vdots \\ s_m \end{pmatrix} - A^{-1}\begin{pmatrix} n_1 \\ \vdots \\ n_m \end{pmatrix}.$$

Thus

$$t_i \in \text{span}\{s_1, \ldots, s_m\} + \text{span}\{n_1, \ldots, n_m\} = S + N \qquad (1 \leq i \leq m)$$

and therefore $T \subset S + N$. Since $N \subset S + N$ we have $X = T + N \subset S + N$. Also $S + N \subset X$. By (iii) we have $X = S \oplus N$.

Theorem 0.1.22. (cf.[6],p.323) If L is a Fredholm operator from X to Y, then L^* is a Fredholm operator from Y^* to X^* with index $i(L^*) = -i(L)$.

Note:

$$X^* = R(L^*) \oplus W$$

$$Y^* = N(L^*) \oplus Z$$

W, Z closed in X^*, Y^* respectively.

0.2 Differentiability in Banach Spaces

Let X, Y be Banach spaces and $M \subset X$ be an open nonempty subset. A function $F : M \to Y$ is said to be Fréchet differentiable at a point $x_0 \in M$ if and only if a continuous linear mapping $\ell : X \to Y$ exists, such that

$$F(x_0 + h) = F(x_0) + \ell h + o(x_0, h) \tag{0.2.1}$$

holds for all $h \in X$, $x_0 + h \in M$, where $o(x_0, h)$ satisfies

(i) $o(x_0, 0) = 0$

(ii) $\lim_{h \to 0} \frac{\|o(x_0, h)\|_Y}{\|h\|_X} = 0.$

The following can be shown: There is at most one such mapping ℓ and if it exists it is called the Fréchet derivative of F at a point x_0 denoted by $\ell = DF(x_0)$. If F is differentiable at all points in M, then F is said to be differentiable in M, in this case, the mapping $DF : M \to \mathcal{L}(X, Y)$ which associates with every point $x \in M$ the derivative of F at x, is called the derivative of F on M. If $DF : M \to \mathcal{L}(X, Y)$ is a continuous mapping, then F is said to be continuously differentiable in M or F is of class C^1 (or that $F \in C^1(M, Y)$.

Suppose F is Fréchet differentiable at $x \in M$. If the Fréchet derivative of $DF : M \to \mathcal{L}(X, Y)$ at x exists it is called the second Fréchet derivative of F at x and is written $D^2F(x) \triangleq D(DF(x))$. Note $\mathcal{L}(X, Y)$ is a Banach space therefore $D^2F(x) = D(DF(x)) \in \mathcal{L}(X, \mathcal{L}(X, Y))$ since now $\mathcal{L}(X, Y)$ plays the role of Y in the derivative.

$\mathcal{L}(X, \mathcal{L}(X, Y))$ is isomorphic to $\mathcal{L}(X^2, Y)$, the Banach space of the continuous bilinear mappings of the Banach space X^2 into the Banach space Y. Here we define $X^1 \triangleq X$ and $X^n \triangleq X \times X^{n-1}$. Likewise the n^{th} derivative is defined as the derivative of the $(n-1)^{th}$ derivative:

$D^nF \triangleq D(D^{n-1}F) \in \mathcal{L}(X, \mathcal{L}(X^{n-1}, Y))$ where $D^{n-1}F : M \to \mathcal{L}(X^{n-1}, Y)$ and $\mathcal{L}(X, \mathcal{L}(X^{n-1}, Y))$ is isomorphic to $\mathcal{L}(X^n, Y)$.

Partial derivatives are defined as derivatives of the corresponding partial maps. For example, suppose that $F : M \subset X \times Z \to Y$ is given. Then F has a partial derivative $D_1F(x_0, z_0) \in \mathcal{L}(X, Y)$ if the partial map $F(\cdot, z_0)$ is differentiable at x_0 with derivative $D_1F(x_0, z_0)$. Likewise $D_2F(x_0, z_0) \in \mathcal{L}(Z, Y)$ if the partial map $F(x_0, \cdot)$ is differentiable at z_0 with derivative $D_2F(x_0, z_0)$.

The Fréchet derivative has the usual properties of derivatives:

(i) The Fréchet derivative is a linear function: That is, if $F_i : M \to Y$, $i = 1, 2$ where $M \subset X$ is an open subset, are two continuously differentiable functions on M and if Y is $\mathbb{K}$ then for any $\lambda_i \in \mathbb{K}$, $\lambda_1 F_1 + \lambda_2 F_2$ is a continuously differentiable function on M and $D(\lambda_1 F_1 + \lambda_2 F_2) = \lambda_1 DF_1 + \lambda_2 DF_2$.

(ii) The Fréchet derivative satisfies the chain rule: That is, if X, Y, Z are Banach spaces and $M \subset X$, $N \subset Y$, $O \subset Z$ are open nonempty sets such that $F : M \to N$ any $G : N \to O$, then $G \circ F : M \to O$ is also continuously differentiable for all $x \in M$ and $D(G \circ F)(x) = DG(F(x)) \circ DF(x)$.

Let X, Y be Banach spaces and let $U \subset X$. Then by a C^r–map in U, $f : U \to Y$, $r \geq 1$, we mean that f has r derivatives which are continuous at each point of U; C^0 means that F is continuous at each point of U. Given a C^r–map $f : X_1 \times X_2 \times \cdots \times X_s \to Y$, $D_i^k f(x_1, x_2, \ldots, x_s)$, $0 \leq k \leq r$, $1 \leq i \leq s$, will denote the (partial) k^{th} derivative of f with respect to the i^{th} argument. Where there is ambiguity we will use the notation $D_{(x_i)} f(x_1, x_2, \ldots, x_s)$ to denote differentiation with respect to x_i.

We will need the following.

Definition 0.2.1. Let $M \subset X$ be a finite dimensional manifold of X. We define $T_bM = \{y | y = \frac{d}{dt}x(t)|_{t=0}$ for some $x : [-1,1] \to M$ such that $x \in C^1$ and $x(0) = b\}$ so that T_bM is the tangent space of M at b.

Theorem 0.2.2. (cf.[12],p.148) Let X, Y, Z be Banach spaces, $U \subset X$ and $V \subset Y$ neighbourhoods of x_0 and y_0 respectively, $F : U \times V \to Z$, a C^2–function. Suppose also that $F(x_0, y_0) = 0$ and $(D_2F(x_0, y_0))^{-1} \in \mathcal{L}(Z, Y)$. Then there exists balls $\overline{B}_r(x_0) \subset U$, $\overline{B}_\delta(y_0) \subset V$, $r, \delta > 0$ and exactly one map $T : B_r(x_0) \to B_\delta(y_0)$ such that $T(x_0) = y_0$ and $F(x, T(x)) = 0$ on $B_r(x_0)$. The map T is a C^2–function.

Theorem 0.2.3. (cf.[18],p.113) Let F be a Banach space and let $E = E_1 \times \cdots \times E_n$, be a product of Banach spaces. Let U_i be open subsets such that $U = U_1 \times \cdots \times U_n \subset E_1 \times \cdots \times E_n$. Let $f : U \to F$ be a map. This map is C^r if and only if each partial derivative $D_if : U_1 \times \cdots \times U_n \to \mathcal{L}(E_i, F)$ is of class C^{r-1}. If this is the case, and $v = (v_1, \ldots, v_n) \in E_1 \times \cdots \times E_n$, then $D_f(x)v = \sum_{i=1}^{n} D_if(x)v_i$.

Remark.

With $r = 2$, assuming that f is C^2 in U implies that $D_1f : U \to \mathcal{L}(E, F)$ is C^1 in U and hence is continuous on U. So if $x(\mu) \in E$ where $\mu \in (-\mu_0, \mu_0) \subset \mathbb{R}$ and if $x(\mu) \to x(0) = 0$ as $\mu \to 0$ then $D_1f(x(\mu), \mu) \to D_1f(0, 0)$ in operator norm as $\mu \to 0$.

Theorem 0.2.4. (cf.[18],p.103) Let E, F be Banach spaces. Let U be an open subset of E. Let $y \in E$. Let $f : U \to F$ be a C^1–map. Assume that the line segment $x + ty$ with $0 \le t \le 1$ is contained in U. Then

$$f(x + y) - f(x) = \int_0^1 Df(x + ty)y\,dt = \int_0^1 Df(x + ty)dt \cdot y \ .$$

Theorem 0.2.5. (cf.[18],p.110) Let E, F be Banach spaces and let U be an open subset of E. Let $f : U \to F$ be of class C^r. Let $x \in U$ and let $y \in E$ be such that the segment $x + ty$, $0 \leq t \leq 1$, is contained in U. Denote by $y^{(k)}$ the k–tuple $(y, y, \ldots, y)$. Then

$$f(x+y) = f(x) + \frac{Df(x)y}{1!} + \cdots + \frac{D^{p-1}f(x)y^{(p-1)}}{(p-1)!} + \tilde{R}_p(y)y^{(p)}$$

where $\tilde{R}_p(y) = \int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} D^p f(x+ty)dt,$ and

$$\lim_{y \to 0} \tilde{R}_p(y) = \frac{1}{p!} D^p f(x).$$

Remark. (i) Theorem 0.2.5 has been adapted from Lang [18] p.110 but we have modified it slightly in the last statement so that it can be used in this form in the following chapters.

(ii) It is not difficult to show that $\lim_{y \to 0} \tilde{R}_p(y) = \frac{1}{p!} D^p f(x)$ since f is continuous.

0.3 Some Definitions and Results Relating to Ordinary Differential Equations (ODE)

Here we list a few definitions and results from the theory of ordinary differential equations and dynamical systems as they pertain to the discussion in the chapters that follow.

We regard an ODE as a system of equations having the following form

$$\dot{x} = f(t, x), \qquad (t, x) \in \mathbb{R} \times \mathbb{R}^n \tag{0.3.1}$$

where $f : U \to \mathbb{R}^n$ with U an open set in $\mathbb{R} \times \mathbb{R}^n$ and $\dot{x} \triangleq \frac{dx}{dt}$. The space of dependent variables is often referred to as the phase space of the system (0.3.1). By a solution of (0.3.1) we will mean a map

$$\phi : I \to \mathbb{R}^n \tag{0.3.2}$$

where I is some open interval in $\mathbb{R}$ such that

$$\phi^{\cdot}(t) = f(t, \phi(t)) \qquad (t \in I). \tag{0.3.3}$$

Existence and Uniqueness of Solutions:

Suppose that f is C^r in U, and for $\xi_1, \xi_2 > 0$ let $I_1 = \{t \in \mathbb{R} \mid t_0 - \xi_1 < t < t_0 + \xi_1\}$ and $I_2 = \{t \in \mathbb{R} \mid t_0 - \xi_2 < t < t_0 + \xi_2\}$; then we have the following theorem.

Theorem 0.3.1. (cf.[1],p.57) Let (t_0, x_0) be a point in U. Then for ξ_1 sufficiently small there exists a solution of (0.3.1), $\phi_1 : I_1 \to \mathbb{R}^n$, satisfying $\phi_1(t_0) = x_0$. Moreover, if f is C^1 in U, and $\phi_2 : I_2 \to \mathbb{R}^n$ is also a solution of (0.3.1) satisfying $\phi_2(t_0) = x_0$, then $\phi_1(t) = \phi_2(t)$ for all $y \in I_3 = \{t \in \mathbb{R} \mid t_0 - \xi_3 < t < t_0 + \xi_3\}$ where $\xi_3 = \min\{\xi_1, \xi_2\}$.

Remarks.

(1) For a solution of (0.3.1) to exist, only continuity of f is required; however, in this case the solution passing through a given point in U may not be unique. If f is at least C^1 in U, then there is a unique solution passing through any given point of U.

(2) In denoting the solutions of (0.3.1) it may be useful to note the dependence on initial conditions explicitly. For ϕ, a solution of (0.3.1) passing through the point $x = x_0$ at $t = t_0$, the notation would be

$$\phi(t, t_0, x_0) \text{ with } \phi(t_0, t_0, x_0) = x_0$$

or

$$\phi(t, x_0) \text{ with } \phi(t_0, x_0) = x_0.$$

Theorem 0.3.2. (cf.Theorem 0.3.3) If $f(t, x)$ is C^r in U, then the solution of (0.3.1) $\phi(t, t_0, x_0)$, $(t_0, x_0) \in U$, is a C^r function of t, t_0 and x_0.

Now suppose (0.3.1) depends on parameters

$$\dot{x} = f(t, x, \xi), \quad (t, x, \xi) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^N \tag{0.3.4}$$

where $f : U \to \mathbb{R}^n$ with U an open set in $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^N$. We denote by $\phi(t, t_0, x_0, \xi)$ the solution $x(t)$ satisfying $x(t_0) = x_0$. We have the following theorem.

Theorem 0.3.3. (cf.[1],p.58) Suppose $f(t, x, \xi)$, is C^r in U. Then the solution of (0.3.4), $\phi(t, t_0, x_0, \xi)$, $(t_0, x_0, \xi) \in U$, is a C^r–function of ξ.

Some Terminology:

(1) A solution $\phi(t, t_0, x_0)$ of (0.3.1) is also called a trajectory through the point x_0.

(2) Suppose we have a solution $\phi(t, t_0, x_0)$; we define an orbit through x_0 to be $\Gamma_{x_0} = \{x \in \mathbb{R}^n \mid x = \phi(t, t_0, x_0), t \in I\}$ so that Γ_{x_0} is the set of points in $\mathbb{R}^n$ through which this solution passes as t varies through I.

Autonomous Systems:

An autonomous system of ODE has the following form

$$\dot{x} = f(x), \ \ x \in \mathbb{R}^n \tag{0.3.5}$$

where $f : U \to \mathbb{R}^n$ with U an open set in $\mathbb{R}^n$. We assume f is C^r, $r \geq 1$ and let $\phi(t)$ be a solution of (0.3.5).

Theorem 0.3.4. (cf.[11],p.116) If $\phi(t)$ is a solution of (0.3.5), then so is $\phi(t + t_0)$ for any $t \in \mathbb{R}$.

Special Solutions:

(1) **Fixed Points:** A point p in the phase space of an ODE $\dot{x} = f(x)$ which satisfies $f(p) = 0$ is called a fixed point. An equilibrium solution of (0.3.5) is a function $x(t) = p \in \mathbb{R}^n$ such that $f(p) = 0$, i.e. a solution which does not change in time. An equilibrium solution is also called a fixed point. Fixed points of vector fields which have the property that the eigenvalues of the matrix associated with the linearization of the vector field about the fixed point have nonzero real parts are called hyperbolic fixed points.

(2) **Homoclinic and Heteroclinic Solutions:** Before we give a definition to this kind we digress slightly to discuss invariant manifolds.

Invariant Manifolds:

A manifold M is invariant if for all $t \in \mathbb{R}$, $\phi_t(M) \subset M$. ϕ_t will be defined below. The most important invariant manifolds for autonomous ODE are the stable and unstable manifolds which we shall define below.

To motivate the concept of invariant manifolds consider a nonlinear autonomous ODE

$$\dot{x} = f(x), \quad x(0) = x_0, \quad x \in \mathbb{R}^n \tag{0.3.6}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is at least C^1.

Definition 0.3.1. Let $\phi(t, x_0)$ be a solution of (0.3.6). The mapping $\phi_t : \mathbb{R}^n \to \mathbb{R}^n$ defined by $\phi_t(x_0) = \phi(t, x_0)$ is called the flow of (0.3.5); ϕ_t is also referred to as the flow of the vector field $f(x)$.

Properties

For all $x \in \mathbb{R}^n$ ϕ_t satisfies

(i) $\phi_o(x) = x$

(ii) $\phi_s(\phi_t(x)) = \phi_{s+t}(x)$ for all $s, t \in \mathbb{R}$.

(iii) $\phi_{-t}(\phi_t(x)) = \phi_t(\phi_{-t}(x)) = x$ for all $s, t \in \mathbb{R}$.

Assume that

(1) $x = 0$ is a fixed point.

(2) $Df(0)$ has $n - k$ eigenvalues having positive real parts and k eigenvalues having negative real parts.

Denote the linearization of (0.3.6) by

$$\dot{z} = Df(0)z, \quad z \in \mathbb{R}^n \tag{0.3.7}$$

and note that $z = 0$ is a fixed point. Let $v^1, \ldots, v^{n-k}$ denote generalized eigenvectors corresponding to the eigenvalues having positive real parts, and $v^{n-k+1}, \ldots, v^n$

denote generalized eigenvectors corresponding to eigenvalues having negative real parts. Then the linear subspaces of $\mathbb{R}^n$ defined as

$$\begin{aligned} E^u &= \operatorname{span}\{v^1, \ldots, v^{n-k}\} \\ E^s &= \operatorname{span}\{v^{n-k+1}, \ldots, v^n\} \end{aligned} \tag{0.3.8}$$

are invariant manifolds for (0.3.7) which are known as the unstable and stable subspaces of (0.3.7) respectively.

The stable manifold theorem for fixed points tells us that in a neighbourhood N of the fixed point $x = 0$ for (0.3.6), there exists a differentiable (as differentiable as $f(x)$) $n-k$ dimensional surface $W_{\text{loc}}(0)$, tangent to E^u at $x = 0$ and a differentiable k dimensional surface $V_{\text{loc}}(0)$, tangent to E^s at $x = 0$ with the properties that orbits of points on $W_{\text{loc}}(0)$ approach $x = 0$ as $t \to -\infty$, and orbits of points on $V_{\text{loc}}(0)$ approach $x = 0$ as $t \to +\infty$. $W_{\text{loc}}(0)$ and $V_{\text{loc}}(0)$ are known as the local unstable and stable manifolds, respectively, of $x = 0$. Denote the flow generated by (0.3.6) by $\phi_t(\cdot)$, then we define the global stable and unstable manifolds of $x = 0$ by

$$\begin{aligned} W &= \bigcup_{t \geq 0} \phi_t(W_{\text{loc}}(0)) \\ V &= \bigcup_{t \leq 0} \phi_t(V_{\text{loc}}(0)), \end{aligned} \tag{0.3.9}$$

respectively. We call W and V the unstable and stable manifolds, respectively.

Suppose we add a small autonomous perturbation $\varepsilon g(x)$ to (0.3.6) where $g(x)$ is as differentiable as $f(x)$ and $\varepsilon \in I \subset \mathbb{R}$ where $I = \{\varepsilon \in \mathbb{R} \mid 0 < \varepsilon < \varepsilon_0\}$, we denote the perturbed system by

$$\dot{x} = f(x) + \varepsilon g(x), \quad x(0) = x_0, \quad x \in \mathbb{R}^n. \tag{0.3.10}$$

Since a fixed point of the unperturbed system is known we can under certain conditions find a fixed point of (0.3.10) in a neighbourhood of the fixed point of the unperturbed system by using the Implicit Function Theorem. Let us set up the problem for application of this theorem by considering the functions

$$\begin{aligned} G : \mathbb{R}^n \times I &\rightarrow \mathbb{R}^n \\ (x, \varepsilon) &\mapsto f(x) + \varepsilon g(x). \end{aligned} \tag{0.3.11}$$

Then $G(0,0) = 0$, and we wish to determine if there exists a solution of $G(x,\varepsilon) = 0$ for (x,ε) close to $(0,0)$. Now $D_1G(0,0) = Df(0)$ and by assumption $\det[D_1G(0,0)] = \det[Df(0)] \neq 0$, so that by the Implicit Function Theorem there exists a function of ε, $\overline{x}(\varepsilon)$, such that $G(\overline{x}(\varepsilon),\varepsilon) = 0$ for ε sufficiently small in I. The persistance theory [cf. [26], Theorem 1.3.6] tells us that in some neighbourhood $\tilde{N}$ containing $x = 0$ and $x = \overline{x}(\varepsilon)$ there exists differentiable manifolds $\tilde{W}_{\text{loc}}(\overline{x}(\varepsilon))$ and $\tilde{V}_{\text{loc}}(\overline{x}(\varepsilon))$ passing through $\overline{x}(\varepsilon)$ with the properties that orbits of points in $\tilde{W}_{\text{loc}}(\overline{x}(\varepsilon))$ under the perturbed flow approach $x = \overline{x}(\varepsilon)$ as $t \rightarrow -\infty$ and orbits of points in $\tilde{V}_{\text{loc}}(\overline{x}(\varepsilon))$ under the perturbed flow approach $x = \overline{x}(\varepsilon)$ as $t \rightarrow +\infty$. $\tilde{W}_{\text{loc}}(\overline{x}(\varepsilon))$ and $\tilde{V}_{\text{loc}}(\overline{x}(\varepsilon))$ have the same dimensions and differentiability as $W_{\text{loc}}(0)$ and $V_{\text{loc}}(0)$, respectively. The global stable and unstable manifolds of $x = \overline{x}(\varepsilon)$ are defined similarly to the way they are defined for the unperturbed system.

We are now ready to define homoclinic and heteroclinic motions.

Definition 0.3.2. Let p be a hyperbolic fixed point of the nonlinear system $\dot{x} = f(x)$ and let ϕ_t be the flow of this system. The stable and unstable manifolds of the system at p are defined by $V = \bigcup_{t\leq 0} \phi_t(V_{loc}(p))$ and $W = \bigcup_{t\geq 0}(W_{loc}(p))$ respectively.

Definition 0.3.3. Let x_0, x_1 be hyperbolic fixed points of $\dot{x} = f(x)$. If p lies in the phase space of $\dot{x} = f(x)$ and if p lies in both the stable and unstable manifolds of x_0 and Γ_p, the orbit of p, approaches x_0 as $|t| \rightarrow \infty$ we say that Γ_p is a homoclinic orbit. If p lies in the unstable manifold of x_0 and the stable manifold of x_1 and Γ_p

approaches x_0 as $t \to -\infty$ and x_1 as $t \to \infty$, then Γ_p is a heteroclinic orbit. The point p is referred to as a homoclinic (resp. heteroclinic) point and, if the stable and unstable manifolds intersect transversally at x_0 it is called a transverse homoclinic (resp. heteroclinic) point.

Chapter 1. INTRODUCTION

1.1 Introduction

This thesis is concerned with the study of small perturbation problems of the form

$$\dot{z} = g(z) + \mu h(t, z, \mu\xi^0) \tag{1.1.1}$$

where $z \in \mathbb{R}^n$, $\mu \in J \subset \mathbb{R}$, $\xi^0 \in \mathbb{R}^N$ and g and h are C^r–functions and bounded on bounded sets. We wish to develop a perturbation technique which will allow us to find bounded solutions for (1.1.1). As with most perturbation theories, we will begin with some assumptions on the unperturbed system

$$\dot{z} = g(z). \tag{1.1.2}$$

Let x_0 and x_1 be hyperbolic fixed points of (1.1.2). That is, the real parts of eigenvalues of $Dg(x_0)$ and $Dg(x_1)$ are nonzero. Then we need the following assumptions on (1.1.2).

A1. The equation (1.1.2) has a bounded solution $\gamma(t)$ such that $\lim_{t\to-\infty} \gamma(t) = x_0$ and $\lim_{t\to\infty} \gamma(t) = x_1$.

A2. The corresponding variational system along $\gamma(t)$ has an exponential dichotomy on both $\mathbb{R}_+$ and $\mathbb{R}_-$ [cf. Section 2.2 of Chapter 2].

Melnikov was the first person to give a criterion to determine the existance of bounded solutions for system (1.1.1). Since his paper in 1963 several authors have considered systems of the form (1.1.1), originally in the case where $n = 2$, $N = 1$ and system (1.1.2) has a homoclinic orbit. We will extend the results of these authors to the case where $n \geq 2$, $N \geq 1$ and (1.1.2) has either a homoclinic or heteroclinic orbit.

Below we will give a short overview of Melnikov's original perturbative method. The Melnikov technique is geometric. In this thesis the Melnikov–type technique is developed for n–dimensional systems possessing hyperbolic fixed points connected to one another by orbits. Our method is not geometric but functional analytic in nature and it makes use of the concept of exponential dichotomies.

The thesis is based on several papers by different authors, namely, Chow, Hale and Mallet–Paret [7], Palmer [20], Hale [16] and Battelli and Lazzari [4]. We do not present any new results but rather give a connected account of the papers mentioned above. A direct proof of Theorem 2 in [4] is also given here [cf. Theorem 32 of this thisis]. At this point let us also mention that the bounded solution $\gamma(t)$ of (1.1.2) and the solution $\gamma(t,\mu)$ we wish to find for (1.1.1) are in fact homoclinic or heteroclinic solutions, or $\gamma(t,\mu)$ approaches periodic solutions $\gamma_0(t,\mu)$ and $\gamma_1(t,\mu)$ in the neighbourhoods of hyperbolic fixed points x_0 and x_1 as $t \to -\infty$ and $t \to \infty$, respectively. That is, there exists $\alpha \in \mathbb{R}$ such that $\lim_{t\to-\infty} |\gamma(t,\mu) - \gamma_0(t+\alpha,\mu)| = 0$ and $\lim_{t\to\infty} |\gamma(t,\mu) - \gamma_1(t+\alpha,\mu)| = 0$. If the latter case occurs we will say $\gamma(t,\mu)$ is a general homoclinic or heteroclinic solution, respectively.

1.2 Motivation

Heteroclinic and homoclinic orbits were defined in Chapter 0 [cf. Definition 0.3.3]. As in Definition 0.3.3 let x_0, x_1 be hyperbolic fixed points of (1.1.2). Since x_0, x_1 are hyperbolic, there are solutions $x_0(t,\mu)$, $x_1(t,\mu)$ of (1.1.1), bounded for $t \in \mathbb{R}$ and existing for μ small such that $x_0(t,0) = x_0$, $x_1(t,0) = x_1$ [cf. [15], p.164, Theorem 4.1]. Let $V_{0\mu}(V_{1\mu})$ be the stable manifold of $x_0(\cdot,\mu)(x_1(\cdot,\mu))$ and $W_{0\mu}(W_{1\mu})$ be the unstable manifold of $x_0(\cdot,\mu)(x_1(\cdot,\mu))$. By hypothesis A1 there is an orbit Γ connecting x_0 and x_1 (this follows from $\lim_{t\to\infty} \gamma(t) = x_1$, $\lim_{t\to-\infty} \gamma(t) = x_0$). In turn the latter statement implies that for $\mu = 0$, $V_{00} \cap W_{10} \neq \emptyset$. From this it follows therefore that finding conditions for the existance of bounded solutions $z(t,\mu)$ of (1.1.1) is equivalent to determining conditions such that $V_{0\mu} \cap W_{1\mu} \neq \emptyset$.

There are many problems that motivate this kind of investigation. If the perturbative term h is independent of t then $x_0(\cdot,\mu) = x_0(\mu), x_1(\cdot,\mu) = x_1(\mu)$ and $x_0(\mu)$ and $x_1(\mu)$ are constant. One then will want conditions which ensure that there is an orbit connecting $x_0(\mu)$ and $x_1(\mu)$, that is, a solution $z(t,\mu)$ of (1.1.1) such that $z(t,\mu) \to x_0(\mu)(x_1(\mu))$ as $t \to -\infty, (\infty)$. An example of this problem is that of travelling waves in parabolic equations. Another example is that of a buckled beam which can be described mathematically by a normalized Duffing's equation $\ddot{x} - x + x^3 = 0$, or as a system

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= x - x^3, \qquad (x,y) \in \mathbb{R} \times \mathbb{R}. \end{aligned}$$

This system has an unstable fixed point at $(0,0)$. This unstable fixed point is connected to itself by two homoclinic orbits which are characterized by the (non-transverse) intersection of the stable and unstable manifolds of $(0,0)$. If we force the system by a periodic force $\mu\cos\theta$ the system becomes $\ddot{x} - x + x^3 = \mu\cos\theta$, or equivalently

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= x - x^3 + \mu\cos\theta \\ \dot{\theta} &= 1, \qquad (x,y,\theta) \in \mathbb{R} \times \mathbb{R} \times S^1. \end{aligned}$$

The unstable fixed point of the unforced system now becomes an unstable periodic orbit of period 2π. Again the stable and unstable manifolds of this periodic orbit intersect nontransversally. For the forced system the stable and unstable manifolds of the periodic orbit may intersect transversally in the phase space $\mathbb{R} \times \mathbb{R} \times S^1$ and as is well known transversal intersection results in chaotic behaviour[cf.[14],section

2.2]. For this reason it is important to study systems of the form (1.1.1). The two examples above provide motivation for the study of weakly perturbed systems like (1.1.1).

1.3 Melnikov Method

In this section we give a short overview of Melnikov's original technique for perturbed systems of the form (1.1.1). As mentioned above the study of homoclinic and heteroclinic orbits is extremely important if one wants to determine, among other things, chaotic behaviour for perturbed systems (1.1.1). The method goes as follows:

Method

Melnikov considered system (1.1.1) for the case where $n = 2$, $N = 1$ with g and h C^r–functions ($r \geq 2$) which are bounded on bounded sets and with h periodic in t with period 2π, i.e., $h(t, z, \mu\xi^0) = h(t + 2\pi, z, \mu\xi^0)$. Following Arrowsmith and Place [cf. [2], pp 170–180], Perko [[22], pp 378–392], Wiggins [[25], pp 484–509], we assume the following conditions.

(C1) For $\mu = 0$ the system (1.1.2) has a homoclinic orbit $q_0(t)$, to a hyperbolic fixed point p_0.

(C2) Let

$$\begin{aligned}\gamma_{p_0} &= \{q_0(t) \mid t \in \mathbb{R}\} \cup \{p_0\}\\ &= V(p_0) \cap W(p_0),\end{aligned}$$

where $W(p_0)$ and $V(p_0)$ are the unstable and stable manifolds of p_0, respectively. The interior of Γ_{p_0} is filled with a continuous family of periodic orbits $q_\alpha(t)$ with period T_α, $\alpha \in (-1,0)$. Assume that $\lim_{\alpha\to 0} q_\alpha(t) = q_0(t)$, $\lim_{\alpha\to 0} T_\alpha = \infty$.

Consider the suspended system equivalent to (1.1.1), namely

$$\begin{aligned} \dot{z} &= g(z) + \mu h(t, z, \mu\xi^0, \theta) \\ \dot{\theta} &= 1 \end{aligned} \tag{1.3.1}$$

where $(z,\theta) \in \mathbb{R}^2 \times S^1$. Then the hyperbolic fixed point p_0 of the z–component of the unperturbed system

$$\begin{aligned} \dot{z} &= g(z) \\ \dot{\theta} &= 1 \end{aligned} \tag{1.3.2}$$

becomes a periodic orbit $\gamma_0(t)$. Denote by $W(\gamma_0(t))$ and $V(\gamma_0(t))$ the unstable and stable manifolds of $\gamma_0(t)$, respectively. In Wiggins [[25], Proposition 4.5.1] it is stated that for ξ sufficiently small, the periodic orbit $\gamma_0(t)$ of the unperturbed system (1.3.2) persists as a periodic orbit $\gamma_\epsilon(t) = \gamma_0(t) + O(\epsilon)$, of the perturbed system (1.3.1) having the same stability type as $\gamma_0(t)$ with $\gamma_\epsilon(t)$ depending on ϵ in a C^r manner. Moreover, $W_{loc}(\gamma_\epsilon(t))$ and $V_{loc}(\gamma_\epsilon(t))$ are C^r ϵ–close to $W_{loc}(\gamma_0(t))$ and $V_{loc}(\gamma_0(t))$ respectively.

The global stabe and unstable manifolds of $\gamma_\epsilon(t)$ are obtained from the local stable and unstable manifolds of $\gamma_\epsilon(t)$. If $\phi_t(\cdot)$ is a flow generated by (1.3.1) then the global stable and unstable manifolds are defined as

$$V(\gamma_\epsilon(t)) = \bigcup_{t\le 0} \phi_t(V_{loc}(\gamma_\epsilon(t)) \quad \text{and}$$

$$W(\gamma_\epsilon(t)) = \bigcup_{t\ge 0} \phi_t(W_{loc}(\gamma_\epsilon(t)), \quad \text{respectively [cf. Definition 0.3.2].}$$

By (C1) $W(\gamma_0(t)) \cap V(\gamma_0(t)) \neq \emptyset$. We wish to find conditions on the perturbative term h to guarantee the non-empty intersection of $V(\gamma_\epsilon(t))$ and $W(\gamma_\epsilon(t))$. The question is dependent on the Melnikov function which is related to the "distance" between these two manifolds. Melnikov had the following geometric approach: Let Σ_{θ_0} be a plane such that $x_{\epsilon,\theta} = \gamma_\epsilon(t) \cap \Sigma_{\theta_0}$ is near $x_{0,\theta} = \gamma_0(t) \cap \Sigma_{\theta_0}$. Note that this is guaranteed by [25] Proposition 4.5.1 mentioned above. Let L be a perpendicular section such that L intersects $\Gamma_{p_0} \times \{\theta_0\}$ at x_0 and $W(\gamma_\epsilon(t)) \cap \Sigma_{\theta_0} = W(x_{\epsilon,\theta})$ at p^u, and $W(\gamma_\epsilon(t)) \cap \Sigma_{\theta_0} = V(x_{\epsilon,\theta})$ at p^s. Note that $\Sigma_{\theta_0} = \mathbb{R}^2 \times \{\theta_0\}$.

With the description above the Melnikov function is defined to be

$$M(\theta_0) \triangleq \int_{-\infty}^{\infty} \exp\,[-\int_{\theta_0}^{t} Dg(\gamma_0(s))ds]g(\gamma_0(t)) \wedge h(t+\theta_0, \gamma_0(t), \mu\xi^0)$$

where the wedge product is defined by $a \wedge b = a_1 b_2 - a_2 b_1$ with $a, b \in \mathbb{R}^2$ such that $a = (a_1, a_2)$, $b = (b_1, b_2)$. The right hand side of the above equation is the measure of the distance between $V(x_{\epsilon,0})$ and $W(x_{\epsilon,\theta})$ along L.

An important result in the Melnikov theory is the following.

Theorem 1.3.1. If $M(\theta_0)$ has a simple zero, then, for sufficiently small $\epsilon > 0$, $W(x_{\epsilon,\theta})$ and $V(x_{\epsilon,\theta})$ intersect transversally for some $\theta_0 \in [0, 2\pi]$. On the other hand if $M(\theta_0)$ is bounded away from zero, then $W(x_{\epsilon,\theta})$ and $V(x_{\epsilon,\theta})$ do not intersect for all θ_0.

Since then a great deal of work has been done on the subject. Without giving any full details we will comment on some other perturbative techniques that have been employed.

1.4 Generalizations of the Melnikov Method and Equivalent Techniques.

As mentioned above many authors have considered systems of the form (1.1.1). Alternative methods have been used to obtain a similar result to that of Melnikov. Among other mthods is that of Chow et al. [7]. Chow et al consider system (1.1.1) for the case where $n = 2$, $N = 1$ and (1.1.2) has a homoclinic orbit Γ. First they obtain conditions for the existance of solutions bounded on $\mathbb{R}$ for the nonhomogeneous linear variational equation around Γ. They prove a version of the Fredholm alternative for solutions bounded on $\mathbb{R}$ [cf. [7], Lemma 2.1]. They then use the Liapunov–Schmidt Reduction Method to obtain two equations that are equivalent to (1.1.1) and these equations are analysed using the application of the Implicit Function Theorem. In short they look at $z(t) = \gamma(t) + x(t)$, $x(t)$ some variation, as a perturbation of a solution $\gamma(t)$ of (1.1.2). They substitute $\gamma(t) + x(t)$ for $z(t)$ in (1.1.1) to get

$$\dot{x} + Dg(\gamma(t)) = F(t, x, \dot{x}, \mu\xi^0), \tag{1.4.1}$$

where $F(t, x, \dot{x}, \mu\xi^0) = \mu h(t, x(t)+\gamma(t), \dot{x}(t)+\dot{\gamma}(t), \mu\xi^0)+Dg(\gamma(t))-g(x(t)+\gamma(t))+g(\gamma(t))$. In Lemma 2 they show that (1.4.1) has bounded solutions on $\mathbb{R}$ if and only if $\int_{-\infty}^{\infty} \dot{\gamma}(t)F(t, x, \dot{x}, \mu\xi^0)dt = 0$, i.e. if and only if $PF = 0$ where P is a projection defined by $PF = \frac{1}{\eta}\dot{\gamma}\int_{-\infty}^{\infty} \dot{\gamma}F$, $\eta = \int_{-\infty}^{\infty} \dot{\gamma}^2$.

From the above result they apply the Liapunov–Schmidt Reduction Method to reduce the problem to solving the equations

$$\begin{aligned} &(a) \qquad PF = 0 \\ &(b) \qquad x = K(I - P)F. \end{aligned} \tag{1.4.2}$$

They use the Implicit Function Theorem to find $x^*(\mu)$ as a solution of 1.4.2 (b). They then conclude that (1.1.1) has a solution $z(t,\mu)$ if and only if

$$\frac{1}{\eta}\int_{-\infty}^{\infty}\dot{\gamma}(t)F(t,x^*(\mu)(t),\dot{x}^*(\mu)(t),\mu\xi^0)dt = 0.$$

The above result has been extended by Palmer [20] to the case where $h(t,z,\mu\xi^0)$ is bounded and the system (1.1.2) has a bounded solution $\gamma(t)$ whose corresponding variational system has an exponential dichotomy on both $\mathbb{R}_+$ and $\mathbb{R}_-$. This concept says that bounded solutions of the variational system tend to zero as $|t| \to \infty$ and unbounded solutions tend to $+\infty$ in absolute value as $|t| \to \infty$. Note that this is not the requirement in [7], i.e. bounded solutions do not necessarily have to tend to 0 as $|t| \to \infty$. Furthermore he considers the case where $n \geq 2$, $N = 1$. A further extension of Palmer's method is that of Hale [16]. He considers a more general case $n \geq 2$, $N \geq 1$ and the variational system along $\gamma(t)$ may have more than one bounded solution [cf. Example 2 in Chapter 5]. Like Palmer he showed that the existance of bounded solutions of (1.1.1) depends on the solvability of an algebraic system whose coefficients depend also on the solution to a suitable system of ODE [cf. [16],p.129].

Battelli and Lazzari [4] have considered this general case and gave conditions on the perturbative term h to guarantee the existance of bounded solutions of (1.1.1). Even though Battelli and Lazzari consider Hale's general case, the method of approach is slightly different. Hale's method is an extension of Chow, Hale and Mallet–Paret's method. Hale proves Lemma 5.4 [cf. [16], p.128] and uses this lemma to apply the Liapunov–Schmidt Reduction Method to obtain bifurcation equations for bounded solutions of (1.1.1) [cf. [16], p.129]. He defines the Melnikov–type function to be $G(a,\mu) = \int_{-\infty}^{\infty}\psi^*(t)F(t,x(a,\mu)(t),\mu\xi^0)dt$, where $\mu \in (-\mu_0,\mu_0) \subset \mathbb{R}$, $a \in \mathbb{R}^{q-1}$ such that $b = (0,a)$ and $x(a,0)(0) = \phi(0)b$ where $\phi = (\dot{\gamma},\phi_1,\ldots,\phi_2)$ is the basis of $N(L)$ [cf. section 2.3 for the definition of $N(L)$], and $\psi(t)$ is a bounded solution of the adjoint equation $\dot{x}(t) - A^*(t)x(t)$, $A(t) \triangleq Dg(\gamma(t))$ [cf. section 2.3 for details]. Note that $G(a,\mu)$ above reduces to $\frac{1}{\eta}\int_{-\infty}^{\infty} \dot{\gamma}(t)F(t, x^*(\mu)(t), \dot{x}^*(\mu)(t), \mu\xi^0)dt$ defined in [7] if equations of the form $\ddot{x} + g(x) = \mu f(t, x, \dot{x}, \mu\xi^0)$ are considered as in [7]. In the Melnikov–type functions in [7] and [16] we have suppressed the dependence on α. $G(a,\mu)$ defined above differs from the Melnikov–type function defined in [4] where it is defined to be $\int_{-\infty}^{\infty} \psi^*(t)h(t,\gamma(t),0)$, the difference being that in [16] the function involves F where F takes the form in equation (1.4.1) whereas in [4] the Melnikov–type function involves h evaluated at $(t,\gamma(t),0)$. The condition on h in order to have bounded solutions for (1,1,1) are $G(a,\mu) = 0$ in [16] and in [4] they are $\int_{-\infty}^{\infty} \psi^*(t)h(t,\gamma(t),0)dt = 0$ and $\int_{-\infty}^{\infty} \psi^*(t)D_1h(t,\gamma(t),0) \neq 0$.

Battelli and Lazzari's method is an extension of Palmer's method. They prove results analogous to Theorem 4.1 in [20] but the index of the Fredholm operator L is greater than or equal to zero [cf. Theorems 2 and 3 in [4]]. These results say that there is an implicit solution to (1.1.1). The results are proved using the Liapunov–Schmidt Reduction Method and the usual Implicit Function Theorem [cf. Chapter 3 of this thesis]. They then use these results and Lemma 4.2 of [20] to show that under certain conditions on the perturbative term h (1.1.1) has a bounded solution [cf. Chapter 4 of this thesis]. Another extension to Palmer's paper is that their method includes the case where $\gamma(t)$ is a heteroclinic orbit. Battelli and Lazzari's paper also includes Gruendler's results [cf. [4], p.363].

Gruendler [13] has a different approach to the methods mentioned above. He constructs the Melnikov–type function from special solutions for the corresponding variational equation $\dot{x} = Dg(\gamma(t))x$. The properties that these solutions must have are summarized in [13] p.912 or in [4] p.363. In Theorem 2.1. [cf. [13], p.913] he states that there are such solutions for $\dot{x} = Dg(\gamma(t))x$. We do not go into details of this author's method but we do mention that the Melnikov–type functions he gets are the same as the ones in Battelli and Lazzari's paper where it is also shown that the formulas Gruendler obtains hold even if $\gamma(t)$ is a heteroclinic orbit provided that $I = n_1 + m_1 - n \geq 0$, where n_1 and m_1 are the dimensions of the stable and unstable subspaces for $\dot{x} = Dg(\gamma(t))x$, respectively. Furthermore the same formulas obtained by Gruendler hold even in a case where $\gamma(t)$ is a bounded solution satisfying the hypothesis of Proposition 1 in [4] or Theorem 4.2 of this thesis.

Lastly as we see above recent authors have generalized the Melnikov function using different methods. They have come up with what some call Generalized Melnikov Functions.

Note also that our theory does not include Cherry's Example [cf. Palmer [21]] where the problem is not put in a Banach space setting but rather a modification of standard integral manifold theory and a finite dimensional version of Theorem 4.1 in [20] (Theorem 3.1 in this thesis) are used to prove the Melnikov–type result (Corollary 4.3) in [20] (Theorem 4.1 in this thesis).

1.5 Contents

In Chapter 2 we state the problem precisely. We then discuss the concept of exponential dichotomies [cf. [9] or section 2.2 of this thesis]. We reformulate equation (1.1.1) to the form

$$f(x, \mu\xi^0) = 0. \tag{1.4.1}$$

We wish to solve this equation using the standard Implicit Function Theorem on Banach spaces. However, it turns out that the linear operator $D_1 f(0,0) \triangleq L$ is not invertible. We end Chapter 2 by characterizing the nullspace and range of L.

As mentioned above L is not invertible and therefore the Implicit Function Theorem cannot be applied in its standard form. In order to use the theorem we derive bifurcation equations from (1.1.1) by using the Liapunov–Schmidt Reduction Method suggested by Chow, Hale and Mallet–Paret [7]. In Chapter 3 we state and prove the Generalized Implicit Function Theorems. In order to give a better understanding and easy reading of Chapter 3 we first give a summary of the proofs of these theorems. Our contribution has been to give a direct proof of Theorem 3.2 [cf. [4], Theorem 2].

In Chapter 4 the results of the previous chapters are used to show the existance of bounded solutions of (1.1.1) [cf. [20], Corollary 4.3 and [4], Proposition 1]. The theorems give conditions on the perturbative term h in order to have bounded solutions. Also note that Theorem 4.2 includes a case where $\gamma(t)$ (solution of the unperturbed system) is a heteroclinic orbit joining hyperbolic z_1, z_2 fixed points of the unperturbed system, such that the number of nonzero eigenvalues of $Dg(z_1)$ and $Dg(z_2)$ is not necessarily the same.

In Chapter 5 we give two examples [cf. [4], p.360 and [20], p.253] to show the applicability of the theory developed in the previous chapters. We give an example for the case where the index I of the Fredholm operator is zero and an example where $I > 0$.

Chapter 2. FUNCTIONAL FORMULATION OF THE SOLUTION OF NON–LINEAR ODE

In this chapter we will state the problem to be solved precisely. The perturbation method to be used makes use of the concept of exponential dichotomies and so we will have a short overview of this concept and state some known results. We then reformulate the perturbed system

$$\dot{z} = g(z) + \mu h(t, z, \mu\xi^0)$$

in the form $f(x, \mu\xi) = 0$ which, under certain circumstances, can be solved using the standard Implicit Function Theorem. However, as we will see, it turns out that the linear operator $L \triangleq D_1 f(0,0)$ is not invertible. In the following chapter we will prove theorems whose proofs rely on the use of the Implicit Function Theorem and then use these theorems to solve $f(x, \mu\xi^0) = 0$. We end this chapter by characterizing the nullspace and range of L.

2.1 Description of the problem

Problem. Given the perturbed system $\dot{z} = g(z) + \mu h(t, z, \mu\xi^0)$ and the fact that the unperturbed system $\dot{z} = g(z)$ has a bounded solution $\gamma(t)$, the problem we wish to solve is that of finding conditions on the perturbative term h to guarantee the existence of bounded solutions for the perturbed system. Let us be more precise.

We consider an n–dimensional system of ODE of the form

$$\dot{z} = g(z) + \mu h(t, z, \mu\xi^\circ) \tag{2.1.1}$$

where $t \in \mathbb{R}$, $z \in \mathbb{R}^n$, $\mu \in J \subset \mathbb{R}$, J an open interval such that $0 \in J$, $\xi^\circ \in U \subset \mathbb{R}^N$, U an open subset such that $0 \in U$. Let $S \subset \mathbb{R}^m$ $(m > 0)$ be an open set. Denote by

$C^k(S, \mathbb{R}^n)$ the space of continuous R^n-valued functions whose Fréchet derivatives up to order k exist and are continuous on S, and by $C_b^k(\mathbb{R}, \mathbb{R}^n)$ a subspace of $C^k(\mathbb{R}, \mathbb{R}^n)$ with bounded continuous $\mathbb{R}^n$-valued functions whose Fréchet derivatives up to order k exist and are bounded and continuous on $\mathbb{R}$. In this chapter $X \triangleq C_b^1(\mathbb{R}, \mathbb{R}^n)$ and $Y \triangleq C_b^0(\mathbb{R}, \mathbb{R}^n)$.

Note that we may write

$$C_b^k(\mathbb{R}, \mathbb{R}^n) = \{\phi \in C^k(\mathbb{R}, \mathbb{R}^n) : \sup_{\substack{t\in\mathbb{R} \\ 0\le j\le k}} |D^j\phi(t)| < \infty\}.$$

On the space $C_b^k(\mathbb{R}, \mathbb{R}^n)$ we use the norm

$$\|\phi\| = \sup_{\substack{t\in\mathbb{R} \\ 0\le j\le k}} |D^j\phi(t)|.$$

With this norm $C_b^k(\mathbb{R}, \mathbb{R}^n)$ is a Banach space. Let $\Omega \subset \mathbb{R}^n$ be an open subset. We assume :

(H1) $$g \in C_b^2(\Omega, \mathbb{R}^n), \quad h \in C_b^2(\mathbb{R} \times \Omega \times U, \mathbb{R}^n).$$

2.2 Exponential Dichotomies

Let $X(t)$ be the fundamental matrix for the linear differential equation

$$\dot{x} = A(t)x, \tag{2.2.1}$$

which is normalized by $X(0) = I$, where I is the identity matrix. Here $A(t)$ is an $n \times n$ matrix that is continuous on an interval J. The equation (2.2.1) is said to possess an exponential dichotomy on J if there exists a projection P and positive constants k, α such that

$$\begin{aligned} &(i) \quad \|X(t)PX^{-1}(s)\| \le ke^{-\alpha(t-s)}, \quad \text{for any } s,t \in J,\ s \le t \\ &(ii) \quad \|X(t)(I-P)X^{-1}(s)\| \le ke^{-\alpha(s-t)}, \quad \text{for any } s,t \in J,\ s \ge t. \end{aligned} \tag{2.2.2}$$

The most interesting cases for the interval J in an exponential dichotomy are $\mathbb{R}_+ \triangleq [0,\infty)$, $\mathbb{R}_- \triangleq (-\infty, 0]$ and $\mathbb{R}$.

The inequalities in (2.2.2) are equivalent to the existence of a projection P and constants $k, \alpha > 0$ such that for all $\xi \in \mathbb{R}^n$ we have

$$\begin{aligned}
&(i) \quad |X(t)P\xi| \leq ke^{-\alpha(t-s)}\,|X(s)P\xi|\,, \text{ for any } s,t \in J,\ s \leq t \\
&(ii) \quad |X(t)(I-P)\xi| \leq ke^{-\alpha(s-t)}\,|X(s)(I-P)\xi|\,, \text{ for any } s,t \in J,\ s \geq t \\
&(iii) \quad \left\|X(t)PX^{-1}(t)\right\| \leq k \text{ for any } t \in J.
\end{aligned} \tag{2.2.3}$$

Interpretation: Let $J = \mathbb{R}_+$. Suppose that P has rank m, then (2.2.3)(i), says there is an m–dimensional subspace of solutions of (2.2.1) given by $R(P)$, called the stable subspace, such that solutions with initial values in $R(P)$ tend to zero as $t \to +\infty$.

The second condition, (2.2.3)(ii), says there is an $(n-m)$–dimensional subspace of solutions of (2.2.1) given by $R(I-P)$, such that solutions with initial values in $R(I-P)$ tend to infinity as $s \to +\infty$.

The projection P appearing in the definition of an exponential dichotomy is not unique when $J = \mathbb{R}_+$ or $\mathbb{R}_-$ [cf. [9], pp. 16-19]. If one has such a P, then any other projection $\tilde{P}$ which has $R(P) = R(\tilde{P})$, but which may have a different nullspace from that of P will satisfy the definition. Therefore when $J = \mathbb{R}_+$, $R(P)$ is the stable subspace of $\mathbb{R}^n$ given by $R(P) = \{\xi \in \mathbb{R}^n : \sup_{t\geq 0}|X(t)\xi| < \infty\}$ but the nullspace may be any complementary subspace. When $J = \mathbb{R}_-$, $N(P)$ is the unstable subspace of $\mathbb{R}^n$ given by $N(P) = \{\xi \in \mathbb{R}^n : \sup_{t\leq 0}|X(t)\xi| < \infty\}$ but the range may be any complementary subspace. Coppel [[10], Theorem 3, p.134] shows that the definition of $R(P)$ and $N(P)$ above is equivalent to the inequalities (2.2.2) and (2.2.3).

Remarks.

(i) In the case of an exponetial dichotomy on the whole line $\mathbb{R}$ the projection P is uniquely determined. If $X(t)$ is the fundamental matrix of (2.2.1) with $X(0) = I$, the range of P is the subspace of initial values of solutions bounded on the positive half–line $\mathbb{R}_+$ and the nullspace of P is the subspace of initial values of solutions bounded on the negative half–line $\mathbb{R}_-$.

(ii) If the equation (2.2.1) has an exponential dichotomy (2.2.2) on each of the half–lines $\mathbb{R}_+$, $\mathbb{R}_-$ with the same projection P then it has an exponential dichotomy on the whole line $\mathbb{R}$ with projection P.

The following are some known results about exponential dichotomies:

Proposition 2.2.1. (cf. [9], p.19) Let $A(t)$ be an $n \times n$ matrix function defined and continuous on $\mathbb{R}$. Then the equation (2.2.1) has an exponential dichotomy on $\mathbb{R}$ if and only if it has an exponential dichotomy on both $\mathbb{R}_+$ and $\mathbb{R}_-$, and $\mathbb{R}^n$ is the direct sum of the stable and unstable subspaces.

Remark. When $A(t) = A$ is constant, system (2.2.1) has an exponential dichotomy on an infinite interval if and only if all the eigenvalues of A have nonzero real parts.

For example if

$$A(t) = A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

then the system $\dot{x} = Ax$ does not have an exponential dichotomy. The important fact about exponential dichotomies is that, if equation (2.2.1) has bounded solutions, then all bounded solutions must tend to zero exponentially as $|t| \to \infty$, and all non–bounded solutions tend to $+\infty$ in absolute value as $|t| \to \infty$.

Proposition 2.2.2. (cf. [16], p.126) Let J be either $\mathbb{R}_+$, $\mathbb{R}_-$ or $\mathbb{R}$. Let equation (2.2.1) have an exponential dichotomy on J and let $B(t)$ be a continuous $n \times n$ matrix function on J. Then for all δ sufficiently small, $\|B(t)\| \le \delta$ implies that

$$\dot{x} = (A(t) + B(t))x \tag{2.2.4}$$

has an exponential dichotomy on J .

Proposition 2.2.3. (cf. [9], p.13) If the equation (2.2.1) has an exponential dichotomy (2.2.2) on a subinterval $[t_0, \infty)$ then it has an exponential dichotomy on the half-line $\mathbb{R}_+$ with the same projection, say P, and the same exponent α.

Proposition 2.2.4. (cf. [9], p.22) Let J be either $\mathbb{R}_+, \mathbb{R}_-$ or $\mathbb{R}$ and let $A(t)$ be continuous. Then the equation

$$\dot{x} = A(t)x + \theta(t) \tag{2.2.5}$$

has at least one bounded solution on J for every $\theta \in C_b^0(\mathbb{R}, \mathbb{R}^n)$ if and only if equation (2.2.1) has an exponential dichotomy on J.

2.3 Functional Formulation

Consider equation (2.1.1) and look for a solution of the form $z(t) = x(t) + \gamma(t)$, where $\gamma(t)$ is a solution of the unperturbed system

$$\dot{y} = g(y) \tag{2.3.1}$$

and $x(t)$ is some variation of this solution. The strategy is to treat equation (2.1.1) as a pertubation of equation (2.3.1), so we regard solutions of equation (2.1.1) as perturbations of solutions of equation (2.3.1). Now $\dot{z}(t) = \dot{x}(t) + \dot{\gamma}(t)$ and since $\dot{\gamma}(t) = g(\gamma(t))$ and $\dot{z}(t) = g(z) + \mu h(t, z, \mu\xi^0)$ we get $\dot{x}(t) = g(x(t) + \gamma(t)) - g(\gamma(t)) +$

$\mu h(t, x(t) + \gamma(t), \mu\xi^0)$. Since g has continuous derivatives with respect to z we can write

$$\dot{x}(t) = Dg(\gamma(t))x(t) + g(x(t) + \gamma(t)) - g(\gamma(t)) - Dg(\gamma(t))x(t) + \mu h(t, x(t) + \gamma(t), \mu\xi^0).$$

This gives the following equation for the variation x.

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + H(t, x(t), \mu\xi^0) \\ A(t) &\triangleq Dg(\gamma(t)) \\ H(t, x(t), \mu\xi^0) &\triangleq g(x(t) + \gamma(t)) - g(\gamma(t)) - Dg(\gamma(t))x(t) \\ &\quad + \mu h(t, x(t) + \gamma(t), \mu\xi^0). \end{aligned} \tag{2.3.2}$$

We assume:

(H2) the unperturbed autonomous system (2.3.1) has a bounded solution γ whose values $\gamma(t)$ are confined to some compact subset $\overline{V}$ of Ω for any $t \in \mathbb{R}$ and the corresponding variational system

$$\dot{x} = Dg(\gamma(t))x \tag{2.3.3}$$

has an exponential dichotomy on both $\mathbb{R}_+$ and $\mathbb{R}_-$ with constants $k, \alpha > 0$ and projections P and Q, respectively.

Remarks.

(i) $R(P)$ is the stable subspace of (2.3.3) and $N(Q)$ is the unstable subspace of (2.3.3).

(ii) Note that in equation (2.2.2) when $J = \mathbb{R}_+$ the equation stays the same and when $J = \mathbb{R}_-$, P is replaced by Q.

(iii) $\dot{\gamma}(t)$ is a non–trivial bounded solution of (2.3.3) and is in the intersection of the stable and unstable subspaces so that $\mathbb{R}^n$ ($n \geq 2$) is not the direct sum of these spaces and hence we cannot have an exponential dichotomy on $\mathbb{R}$ by Proposition 2.2.1.

Let us write equation (2.3.2) in the form

$$\dot{x}(t) - A(t)x(t) - H(t, x(t), \mu\xi^0) = 0, t \in \mathbb{R}. \tag{2.3.4}$$

We are looking for a global bounded solution $x(t)$ of equation (2.3.4). We decide to use the following strategy to solve this equation. The left hand side of (2.3.4) is just the value $f(x, \mu\xi^0)$ of a map f from the Banach space $X \times \mathbb{R}^N$, into the Banach space Y defined for $x \in X$ and $\mu \in \mathbb{R}$ as the function $f(x, \mu\xi^0) : \mathbb{R} \to \mathbb{R}^n$.

$$f(x, \mu\xi^0)(t) = \dot{x}(t) - A(t)x(t) - H(t, x(t), \mu\xi^0) \text{ for all } t \in \mathbb{R}.$$

Since by assumption $H : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^N \to \mathbb{R}^n$ and $Dg : \mathbb{R}^n \to \mathbb{R}^n$ are bounded on bounded sets and continuous, $f(x, \mu\xi^0)$ indeed is an element of Y, thus $f : X \times \mathbb{R}^N \to Y$ accordingly. To solve equation (2.3.4) one has to solve the equation

$$f(x, \mu\xi^0) = 0 \in Y \tag{2.3.5}$$

for $x \in X$, $\mu \in \mathbb{R}$, $\xi^0 \in \mathbb{R}^N$. We now show that f is a C^2–function.

Let $\xi = \mu\xi^0$, then $f(x, \mu\xi^0) = f(x, \xi)$. Using the definition of the Fréchet derivative we have

$$\begin{aligned} f(x+y, \xi) - f(x, \xi) &= \dot{x} + \dot{y} - [g(\gamma + x + y) - g(\gamma)] - \mu h(\cdot, \gamma + x + y, \xi) \\ &\quad - \dot{x} + [g(\gamma + x) - g(\gamma)] + \mu h(\cdot, \gamma + x, \xi) \\ &= \dot{y} - [g(\gamma + x + y) - g(\gamma + x)] - \mu[h(\cdot, \gamma + x + y, \xi) - h(\cdot, \gamma + x, \xi)] \\ &= \dot{y} - Dg(\gamma + x)y + o(\gamma + x, y) - \mu D_2 h(\cdot, \gamma + x, \xi)y + o(\gamma + x, y) \\ &= \dot{y} - [Dg(\gamma + x)y + \mu D_2 h(\cdot, \gamma + x, \xi)]y + o(\gamma + x, y) \\ &= \ell_1 y + o(\gamma + x, y) \end{aligned}$$

$$\ell_1 \triangleq D_1 f(x, \xi) = \frac{d}{dt} - [Dg(\gamma + x) + \mu D_2 h(\cdot, \gamma + x, \xi)].$$

This shows that f is C^1 since by definition of Banach spaces X and Y, $\frac{d}{dt}$ is a bounded linear operator from $X = C_b^1(\mathbb{R}, \mathbb{R}^n)$ into $Y = C_b^0(\mathbb{R}, \mathbb{R}^n)$ and by (H1) g and h are C_b^2 functions.

Now

$$\begin{aligned}
D_1 f(x+w,\xi) - D_1 f(x,\xi) &= \frac{d}{dt} - Dg(\gamma + x + w) + \mu D_2 h(\cdot, \gamma + x + w, \xi) \\
&\quad - [\frac{d}{dt} - Dg(\gamma + x) + \mu D_2 h(\cdot, \gamma + x, \xi)] \\
&= -[Dg(\gamma + x + w) - Dg(x + \gamma)] \\
&\quad + \mu[D_2 h(\cdot, \gamma + x + w) - D_2 h(\cdot, \gamma + x, \xi)]
\end{aligned}$$

$$\begin{aligned}
&= -D^2 g(\gamma + x)w + o(\gamma + x, w) + \mu D_2^2 h(\cdot, \gamma + x, \xi)w \\
&\quad + o(\gamma + x, w) \\
&= -[D^2 g(\gamma + x) - \mu D_2^2 h(\cdot, \gamma + x, \mu)]w + o(\gamma + x, w) \\
&= \ell_2 w + o(\gamma + x, w)
\end{aligned}$$

$$\ell_2 \triangleq D_1^2 f(x,\xi) = -[D^2 g(\gamma + x) - \mu D_2^2 h(\cdot, \gamma + x, \xi)]$$

and this shows that f is C^2 by (H1).

From the equation above we have for any $y \in X$

$$\begin{aligned}
[D_1 f(x,\xi)y](t) &= \dot{y}(t) - [Dg(\gamma(t) + x(t)) \\
&\quad + \mu D_2 h(t, \gamma(t) + x(t), \xi)]y(t). \qquad (2.3.6)
\end{aligned}$$

Evaluate (2.3.6) at $(x,\xi) = (0,0) \in X \times \mathbb{R}^N$ to get

$$(Ly)(t) \triangleq [D_1 f(0,0)y](t) = \dot{y}(t) - A(t)y(t), \text{ where } A(t) \triangleq Dg(\gamma(t)).$$

The equation of the form $f(x,\xi) = 0$ can under certain circumstances be solved locally using the Implicit Function Theorem. The theorem can be applied in its standard form if there is a point (x_0, ξ_0) such that $f(x_0, \xi_0) = 0$ and $D_1 f(x_0, \xi_0)$, the Fréchet derivative with respect to x, is an invertible operator from X to Y. In our situation we know that $f(0,0) = 0$ so we need to study $L = D_1 f(0,0)$.

If L were an invertible operator from X to Y, the standard version of the Implicit Function Theorem [cf. Theorem 0.2.2] would allow us to solve $f(x, \xi) = 0$ locally. However, typically the operator L is not invertible since (2.3.3) has a nontrivial solution $\gamma(t)$, that is, $N(L)$ is non–trivial. In order to be able to reduce the above problem to a standard application of the Implicit Function Theorem we need some information about $N(L)$ and $R(L)$. Let $V = R(P)$, the stable subspace, and $W = N(Q)$ the unstable subspace, for the fundamental solution $X(t)$ of $\dot{y}(t) = A(t)y(t)$, $A(t) \triangleq Dg(\gamma(t))$.

We now characterize $N(L)$ and $R(L)$, the nullspace and range of L, respectively. We have the following :

Lemma 2.3.1. Let $A(t)$ be an $n \times n$ matrix function bounded and continuous on $\mathbb{R}$ such that (2.3.3) has an exponential dichotomy on $\mathbb{R}_+$ and $\mathbb{R}_-$. Then the linear operator $L : X \to Y$ defined by $(Lx)(t) = \dot{x}(t) - A(t)x(t)$ is Fredholm and a function $\rho \in Y$ is in the range $R(L)$ of L if and only if

$$\int_{-\infty}^{\infty} \psi^*(t)\rho(t)dt = 0 \tag{2.3.7}$$

for all bounded solutions $\psi(t)$ of the adjoint system

$$\dot{x} = -A^*(t)x. \tag{2.3.8}$$

L is a Fredholm operator from X to Y and the index of L is $\dim V + \dim W - n$, where V and W are the stable and unstable subspaces for the fundamental solution $X(t)$ of (2.3.3).

Proof. Let (2.3.3) have an exponential dichotomy on $\mathbb{R}_+$ and on $\mathbb{R}_-$ with projections P and Q, respectively, and suppose that the associated fundamental matrix $X(t)$ satisfies $X(0) = I$. Then the adjoint system has the fundamental matrix $X^{*-1}(t)$. Note that for any matrix B we have $\|B\| = \|B^*\|$ [cf. Theorem 0.1.11]. By exchanging t and s in inequalities (2.2.2) and using the observation above one gets

$$\begin{aligned} &(i) \quad \left\|X^{*-1}(t)P^*X^*(s)\right\| \leq ke^{-\alpha(s-t)}, t \leq s, \text{ for all } s,t \in \mathbb{R}_+ \\ &(ii) \quad \left\|X^{*-1}(t)(I-P^*)X^*(s)\right\| \leq ke^{-\alpha(t-s)}, s \leq t, \text{ for all } s,t \in \mathbb{R}_+ \end{aligned} \tag{2.3.9}$$

and the same with Q^* and $\mathbb{R}_-$ replacing P^* and $\mathbb{R}_+$ respectively. Then (2.3.8) has an exponential dichotomy on $\mathbb{R}_+$ with projection $I - P^*$ and on $\mathbb{R}_-$ with projection $I - Q^*$. Let $V = R(P)$ (stable subspace) and let $W = N(Q)$ (unstable subspace), then the subspace of initial values (at $t = 0$) of bounded global solutions of (2.3.3) is $V \cap W$ since solutions with initial values in V tend to zero as $t \to \infty$, and solutions with initial values in W tend to zero as $t \to -\infty$. For (2.3.8) the subspace of initial values (at $t = 0$) of bounded global solutions is $V^\perp \cap W^\perp$, where $V^\perp$ and $W^\perp$ are the orthogonal complements of V and W respectively. [$V^\perp = R(P)^\perp = N(P^*) = R(I - P^*)$ and $W^\perp = N(Q)^\perp = R(Q^*) = N(I - Q^*)$ cf. Theorem 0.1.14]. $V^\perp$ is the stable subspace and $W^\perp$ is the unstable subspace of the adjoint equation.

By definition $N(L) = \{x \in X : \dot{x}(t) - A(t)x(t) = 0\}$. Note that if K is a subspace of $N(L)$ then $\dim K = \dim\{x(0) \mid x \in K\}$. Let $S = \{\xi \in \mathbb{R}^n : \xi = x(0)$ for some $x \in N(L)\}$. Then since $x(0) \in V \cap W$ for $x \in N(L)$, we have $S = V \cap W$ and hence $\dim N(L) = \dim S = \dim(V \cap W)$. Since $\dot{\gamma}(t)$ is a nontrivial bounded solution of (2.3.3) it follows that $\dim N(L) \geq 1$. By the same argument above one can show that $N(L^*) = \{x \in Y^* : L^*x = 0,\ x(0) \in V^\perp \cap W^\perp\}$ so that $\dim N(L^*) = \dim(V^\perp \cap W^\perp)$.

Let $\rho \in R(L)$. We will show that equation (2.3.7) holds. Since $\rho \in R(L)$ there is an x in X satisfying $\rho(t) = \dot{x}(t) - A(t)x(t)$, for all $t \in \mathbb{R}$. If $\psi(t)$ is a bounded solution of (2.3.8) we have

$$\begin{aligned}
\int_{-\infty}^{\infty} \psi^*(t)\rho(t)dt &= \int_{-\infty}^{\infty} \psi^*(t)[\dot{x}(t) - A(t)x(t)]dt \\
&= \int_{-\infty}^{\infty} [\psi^*(t)\dot{x}(t) - \psi^*(t)A(t)x(t)]dt \\
&= \int_{-\infty}^{\infty} [\psi^*(t)\dot{x}(t) + \dot{\psi}^*(t)x(t)]dt \\
&\qquad [\text{since } \dot{\psi}(t) = -A^*(t)\psi(t) \implies \dot{\psi}^*(t) = -\psi^*(t)A(t)] \\
&= \int_{-\infty}^{\infty} \frac{d}{dt}(\psi^*(t)x(t))dt \\
&= \psi^*(t)x(t) \mid_{-\infty}^{\infty} \\
&= 0
\end{aligned}$$

[since $\psi(t) \to 0$ exponentially as $|t| \to \infty$]. So if $\rho \in R(L)$, (2.3.7) holds for all bounded solutions $\psi(t)$ of (2.3.8).

This completes one direction.

Remark. The existence of integrals above is guaranteed by the fact that the adjoint system has an exponential dichotomy on both $\mathbb{R}_+$ and $\mathbb{R}_-$. ρ is bounded ($\rho \in Y \triangleq C_b^0(\mathbb{R}, \mathbb{R}^n)$) and $\psi(t) \to 0$ exponentially as $|t| \to \infty$ since the adjoint system has an exponential dichotomy on both $\mathbb{R}_+$ and $\mathbb{R}_-$.

Conversely, suppose that $\rho \in Y$ and that (2.3.7) holds for all bounded solutions $\psi(t)$ of the adjoint system (2.3.8). We will show that $\rho \in R(L)$. From linear algebra the equation $B\xi = b$ has a solution ξ if and only if $\eta^* b = 0$ for all vectors η such that $\eta^* B = 0$. Now choose a vector η satisfying

$$\eta^*[P - (I - Q)] = 0 \tag{2.3.10}$$

and let

$$\psi(t) = X^{*-1}(t)(I - P^*)\eta = X^{*-1}(t)Q^*\eta, \qquad t \in \mathbb{R}.$$

We will show that $\psi(t)$ defined above is a bounded solution of (2.3.8). We have for $t \in \mathbb{R}$

$$\begin{aligned}\dot{\psi}(t) &= \frac{d}{dt}(X^{*-1}(t))(I - P^*)\eta \\ &= -A^*(t)X^{*-1}(t)(I - P^*)\eta \quad \text{(since } X^{*-1}(t) \text{ is a matrix solution)} \\ &= -A^*(t)\psi(t).\end{aligned}$$

Since $A(t)$ and $\psi(t)$ are continuous $\dot{\psi}(t)$ is continuous so that indeed ψ is a solution of (2.3.8) in C^1. Also with $\eta \in \mathbb{R}^n$, $(I - P^*)\eta$ is in $V^\perp = R(I - P^*)$ and $\psi(t) = X^{*-1}(t)(I - P^*)\eta$ is a solution of (2.3.8) with initial values in $V^\perp$ and hence by estimate (2.3.9) tends to zero as $t \to +\infty$. Hence the function ψ is bounded on $[0, \infty)$. Also with $\eta \in \mathbb{R}^n$, $Q^*\eta$ is in $W^\perp = R(Q^*)$ and $\psi(t) = X^{*-1}(t)Q^*\eta$ is a solution of (2.3.8) with initial values in $W^\perp$ and hence by estimate (2.3.9) tends to zero as $t \to -\infty$. Hence the function ψ is bounded on $(-\infty, 0]$. Therefore $\psi(t)$ is a bounded solution of (2.3.8).

Since we have assumed that ρ satisfies (2.3.7) we get

$$\begin{aligned}0 &= \int_{-\infty}^{\infty} \psi^*(t)\rho(t)dt = \int_{-\infty}^{0} \psi^*(t)\rho(t)dt + \int_{0}^{\infty} \psi^*(t)\rho(t)dt \\ &= \int_{-\infty}^{0} \eta^* QX^{-1}(t)\rho(t)dt + \int_{0}^{\infty} \eta^*(I - P)X^{-1}(t)\rho(t)dt \\ &= \eta^*[\int_{-\infty}^{0} QX^{-1}(t)\rho(t)dt + \int_{0}^{\infty} (I - P)X^{-1}(t)\rho(t)dt].\end{aligned}$$

We thus have

$$\eta^*[P - (I - Q)] = \eta^*[\int_{-\infty}^{0} QX^{-1}(t)\rho(t)dt + \int_{0}^{\infty} (I - P)X^{-1}(t)\rho(t)dt]. \quad (2.3.11)$$

If we take $B = P - (I - Q)$ and $b = \int_{-\infty}^{0} QX^{-1}(t)\rho(t)dt + \int_{0}^{\infty}(I - P)X^{-1}(t)\rho(t)dt$ then from (2.3.10) and the argument above it, it follows that (2.3.11) has a solution ξ satisfying

$$[P - (I - Q)]\xi = \int_{-\infty}^{0} QX^{-1}(t)\rho(t)dt + \int_{0}^{\infty}(I - P)X^{-1}(t)\rho(t)dt \qquad (2.3.12)$$

Now define a function $x : \mathbb{R} \to \mathbb{R}^n$ by

$$x(t) = \begin{cases} X(t)P\xi + \int_0^t X(t)PX^{-1}(s)\rho(s)ds - \int_t^\infty X(t)(I - P)X^{-1}(s)\rho(s)ds \text{ if } t \geq 0 \\ X(t)(I - Q)\xi + \int_{-\infty}^t X(t)QX^{-1}(s)\rho(s)ds - \int_t^0 X(t)(I - Q)X^{-1}(s)\rho(s)ds \text{ if } t \leq 0. \end{cases} \qquad (2.3.13)$$

We will show that (i) $x \in X$ (ii) $Lx = \rho$.

Note that $x(t)$ in (2.3.13) is well defined since ρ is bounded on $\mathbb{R}$ and $(I - P)X^{-1}(s)$ and $QX^{-1}(s)$ are bounded by a decaying exponential and hence $(I - P)X^{-1}(s)\rho(s)$ and $QX^{-1}(s)\rho(s)$ are bounded by a constant times a decaying exponential so that the integrals in (2.3.13) exist.

(i) We now show that $x(t)$ is a solution of $\dot{x}(t) = A(t)x(t) + \rho(t)$

In (2.3.13) for $t \geq 0$ we have

$$\begin{aligned} x(t) &= X(t)P\xi + \int_0^t X(t)PX^{-1}(s)\rho(s)ds - \int_t^\infty X(t)(I - P)X^{-1}(s)\rho(s)ds \\ &= X(t)[P\xi + \int_0^t PX^{-1}(s)\rho(s)ds - \int_t^\infty (I - P)X^{-1}(s)\rho(s)ds] \end{aligned}$$

so that

$$
\begin{aligned}
\dot{x}(t) &= \dot{X}(t)[P\xi + \int_0^t X^{-1}(s)\rho(s)ds - \int_t^\infty X(t)(I-P)X^{-1}(s)\rho(s)ds] \\
&+ X(t)[PX^{-1}(t)\rho(t) + (I-P)X^{-1}(t)\rho(t)] \\
&= A(t)X(t)[P\xi + \int_0^t PX^{-1}(s)\rho(s)ds - \int_t^\infty (I-P)X^{1}(s)\rho(s)ds] \\
&+ X(t)[P + I - P]X^{-1}(t)\rho(t) \\
&= A(t)x(t) + \rho(t)
\end{aligned}
$$

and for $t \leq 0$ we have

$$
x(t) = X(t)[(I-Q)\xi + \int_{-\infty}^t QX^{-1}(s)\rho(s)ds - \int_t^0 (I-Q)X^{-1}(s)\rho(s)ds]
$$

so that

$$
\begin{aligned}
\dot{x}(t) &= \dot{X}(t)[(I-Q)\xi + \int_{-\infty}^t QX^{-1}(s)\rho(s)ds - \int_t^0 (I-Q)X^{-1}(s)\rho(s)ds] \\
&+ X(t)[QX^{-1}(t)\rho(t) + (I-Q)X^{-1}(t)\rho(t)] \\
&= A(t)X(t)[(I-Q)\xi + \int_{-\infty}^t QX^{-1}(s)\rho(s)ds - \int_t^0 (I-Q)X^{-1}(s)\rho(s)ds] \\
&+ X(t)[Q + I - Q]X^{-1}(t)\rho(t) \\
&= A(t)x(t) + \rho(t).
\end{aligned}
$$

Also from (2.3.13) we get

$$
\begin{aligned}
x(0^+) &= P\xi - \int_0^\infty (I-P)X^{-1}(t)\rho(t)dt \\
x(0^-) &= (I-Q)\xi + \int_{-\infty}^0 X^{-1}(t)\rho(t)dt
\end{aligned}
$$

and by (2.3.12)

$$
\begin{aligned}
& [P - (I-Q)]\xi = \int_{-\infty}^0 X^{-1}(t)dt + \int_0^\infty (I-P)X^{-1}(t)\rho(t)dt \\
\Longrightarrow\ & P\xi - \int_0^\infty (I-P)X^{-1}(t)\rho(t)dt = (I-Q)\xi + \int_{-\infty}^0 X^{-1}(t)\rho(t)dt.
\end{aligned}
$$

Thus (2.3.12) implies $x(0^+) = x(0^-)$. Since $A(t)$, $x(t)$ and $\rho(t)$ are continuous and $\dot{x}(t) = A(t)x(t) + \rho(t)$, it follows that $\dot{x}(t)$ is continuous so that indeed $x \in C^1$.

We now show that the solution defined in (2.3.13) is bounded and tends to zero as $|t| \to \infty$. That is, $x(t)$ is bounded and hence $x \in X$.

For $t \geq 0$ we have

$$x(t) = X(t)P\xi + \int_0^t X(t)PX^{-1}(s)\rho(s)ds - \int_t^\infty X(t)(I-P)X^{-1}(s)\rho(s)ds.$$

By definition of $R(P)$ (see Section 2.2) $|X(t)P\xi| = |X(t)\xi| \leq N$, N some positive constant. Also $\rho \in Y$ so that ρ is bounded, say by M, that is $|\rho(t)| \leq M$ for all $t \in \mathbb{R}$.

Then we have

$$\begin{aligned}
\beta(t) &= \left| \int_0^t X(t)PX^{-1}(s)\rho(s)ds - \int_t^\infty X(t)(I-P)X^{-1}(s)\rho(s)ds \right| \\
&\leq \left| \int_0^t X(t)PX^{-1}(s)\rho(s)ds \right| + \left| \int_t^\infty X(t)(I-P)X^{-1}(s)\rho(s)ds \right| \\
&\leq \int_0^t \left|X(t)PX^{-1}(s)\rho(s)\right| ds + \int_t^\infty \left|X(t)(I-P)X^{-1}(s)\rho(s)\right| ds \\
&\leq \int_0^t \left\|X(t)PX^{-1}(s)\right\| |\rho(s)|\, ds + \int_t^\infty \left\|X(t)(I-P)X^{-1}(s)\right\| |\rho(s)|\, ds \\
&\leq M \int_0^t \left\|X(t)PX^{-1}(s)\right\| ds + M \int_t^\infty \left\|X(t)(I-P)X^{-1}(s)\right\| ds \\
&\leq kM \left[\int_0^t e^{-\alpha(t-s)} ds + \int_t^\infty e^{\alpha(t-s)} ds \right]
\end{aligned}$$

[since $|\rho| \leq M$, $\left\|X(t)PX^{-1}(s)\right\| \leq ke^{-\alpha(t-s)}$, $t \geq s$ and $\left\|X(t)(I-P)X^{-1}(s)\right\| \leq ke^{-\alpha(s-t)}$, $s \geq t$]

$$\leq kM[e^{-\alpha t}\int_0^t e^{\alpha s}ds + e^{\alpha t}\int_t^\infty e^{-\alpha s}ds]$$
$$\leq kM[e^{-\alpha t}\frac{1}{\alpha}e^{\alpha s}\mid_0^t + e^{\alpha t}\lim_{B\to\infty}\{-\frac{1}{\alpha}e^{-\alpha s}\mid_t^B\}]$$
$$\leq \frac{kM}{\alpha}[1-e^{-\alpha t} - \lim_{B\to\infty}(e^{-\alpha B}-e^{-\alpha t})e^{\alpha t}]$$
$$\leq \frac{kM}{\alpha}[1-e^{-\alpha t}-(0-1)]$$
$$\leq \frac{kM}{\alpha}[2-e^{-\alpha t}]$$
$$\leq \frac{2kM}{\alpha}.$$

Hence $|x(t)| \leq |X(t)P\xi| + \beta(t) \leq N + \dfrac{2kM}{\alpha}$ on $[t_0,\infty)$ for some $t_0 \in \mathbb{R}_+$.

For $t \leq 0$ we have

$$x(t) = X(t)(I-Q)\xi + \int_{-\infty}^t X(t)QX^{-1}(s)\rho(s)ds - \int_t^0 X(t)(I-Q)X^{-1}(s)\rho(s)ds$$

By definition of $N(Q)$ in Section 2 $|X(t)(I-Q)\xi| \leq |X(t)\xi| \leq N$, $N > 0$. Also $\rho \in Y$ so that ρ is bounded, say by M, that is, $|\rho(t)| \leq M$, for all $t \in \mathbb{R}$.

$$\tilde{\beta}(t) = \left|\int_{-\infty}^t X(t)QX^{-1}(s)\rho(s)ds - \int_t^0 X(t)(I-Q)X^{-1}(s)\rho(s)ds\right|$$
$$\leq \left|\int_{-\infty}^t X(t)QX^{-1}(s)\rho(s)ds\right| + \left|\int_t^0 X(t)(I-Q)X^{-1}(s)\rho(s)ds\right|$$
$$\leq \int_{-\infty}^t \left|X(t)QX^{-1}(s)\rho(s)\right| ds + \int_t^0 \left|X(t)(I-Q)X^{-1}(s)\rho(s)\right| ds$$
$$\leq \int_{-\infty}^t \|X(t)QX^{-1}(s)\|\,|\rho(s)|\,ds + \int_t^0 \|X(t)(I-Q)X^{-1}(s)\|\,|\rho(s)|\,ds$$

$$\begin{aligned}
&\leq \int_{-\infty}^{t} ke^{-\alpha(t-s)}Mds + \int_{t}^{0} ke^{-\alpha(s-t)}Mds \\
&\leq Mk\left[e^{-\alpha t}\int_{-\infty}^{t} e^{\alpha s}ds + e^{\alpha t}\int_{t}^{0} e^{-\alpha s}ds\right] \\
&\leq Mk\left[e^{-\alpha t}\lim_{B\to-\infty}\frac{1}{\alpha}e^{\alpha s}\Big|_{B}^{t} + e^{\alpha t}(-\frac{1}{\alpha})e^{-\alpha s}\Big|_{t}^{0}\right] \\
&\leq \frac{Mk}{\alpha}\left[e^{-\alpha t}(e^{\alpha t}-0) - e^{\alpha t}(1-e^{-\alpha t})\right] \\
&\leq \frac{Mk}{\alpha}\left[1-0-e^{\alpha t}+1\right] \\
&\leq \frac{Mk}{\alpha}\left[2-e^{\alpha t}\right] \\
&\leq \frac{2Mk}{\alpha}.
\end{aligned}$$

Hence $|x(t)| \leq |X(t)(I-Q)\xi| + \tilde{\beta}(t) \leq N + \dfrac{2kM}{\alpha}$ on $(-\infty, t_0]$ for some $t_0 \in \mathbb{R}_-$. We have shown that $x(t)$ remains bounded as $|t| \to \infty$.

Therefore $x \in X$ and $\dot{x}(t) = A(t)x(t) + \rho(t)$ and x is bounded, $A(t)$ is bounded and $\rho(t)$ is bounded so that $\dot{x}(t)$ is bounded.

(ii) Above we have shown that $x \in X$ satisfies $\dot{x}(t) = A(t)x(t) + \rho(t)$, that is, $(Lx)(t) = \rho(t)$ and hence $\rho \in R(L)$.

This completes the first part of the proof.

That $R(L)$ is closed is obvious since we have shown that

$$R(L) = \{\rho \in Y : \int_{-\infty}^{\infty} \psi^*(t)\rho(t)dt = 0, \text{ for all } \psi(t) \text{ satisfying } L^*x = 0\} = N(L^*)^{\perp}.$$

$$\left[\psi(t) \in N(L^*) \text{ and so } \rho(t) \in N(L^*)^{\perp}\right].$$

Since $N(L^*)^{\perp}$ is always closed the equality $R(L) = N(L^*)^{\perp}$ proves that $R(L)$ is closed. Also we showed that

$$N(L^*) = \{x \in Y^* : L^*x = 0,\ x(0) \in V^{\perp} \cap W^{\perp}\}.$$

Thus

$$\dim N(L^*) = \dim(V^\perp \cap W^\perp) < \infty \text{ and } N(L^*) = R(L)^\perp \text{ [cf. Theorem 0.1.14]}$$

and hence

$$\dim N(L^*) = \dim R(L)^\perp = \operatorname{codim} R(L) < \infty. \qquad \text{[cf. Theorem 0.1.17]}$$

So L is Fredholm.

$$\begin{aligned}
\text{Index } L &= \dim N(L) - \dim R(L)^\perp \\
&= \dim(V \cap W) - \dim(V^\perp \cap W^\perp) \\
&= \dim(V \cap W) - \dim(V + W)^\perp \text{ since } V^\perp \cap W^\perp = (W + V)^\perp \text{ [cf. Theorem 0.1.13]} \\
&= \dim(V \cap W) - [n - \dim(V + W)] \text{ since } n = \dim(V + W) + \dim(V + W)^\perp \\
&= \dim(V \cap W) - [n - \{\dim V + \dim W - \dim(V \cap W)\}] \text{ since } V \cap W \neq \emptyset \\
&= \dim V + \dim W - n.
\end{aligned}$$

This completes the proof of the lemma. □

Chapter 3. THE GENERALIZED IMPLICIT FUNCTION THEOREMS

In Chapter 2 we reformulated equation (2.1.1) to the form of equation (2.3.5), that is the equation $f(x, \mu\xi_0) = 0$. We can find bounded solutions of equation (2.1.1) by solving the equation

$$f(x, \mu\xi_0) = 0 \tag{3.1}$$

in a neighbourhood of $(0,0)$ in $E \times \mathbb{R}^N$ according to the strategy explained at the beginning of Chapter 2. In this chapter we will prove two theorems concerning the existence of an implicit solution to the equation $f(x, \mu\xi_0) = 0$, $f : E \times \mathbb{R}^N \to F$ being a C^2–map and E, F Banach spaces.

In our first theorem we wish to solve $f(x, \mu) = 0$ near $(0,0) \in E \times \mathbb{R}$. Note that in (3.1) we have $1 = \xi^0 \in \mathbb{R}^1$.

Theorem 3.1. Let E, F be Banach spaces and $f : E \times \mathbb{R} \to F$ be a C^2–mapping on a neighbourhood of $(0,0)$ in $E \times \mathbb{R}$ such that $f(0,0) = 0$ and $L \triangleq D_1 f(0,0)$ is a Fredholm operator with index $I = 0$. Let P and Q be continuous projections on E and F, respectively, such that $R(P) = N(L)$ and $R(Q) = R(L)$, respectively. Moreover assume:

(i) there is a C^2–manifold $M = \{x = \xi + \phi(\xi) : \xi \in N(L), \; \|\xi\| < \delta\}$, where $\phi : \{\xi \in N(L) \mid \|\xi\| < \delta\} \to N(P)$ is a C^2–function satisfying $\phi(0) = 0$, $D\phi(0) = 0$, such that $f(x, 0) = 0$ for all $x \in M$ and such that $N(L)$ is the tangent space of M at 0,

(ii) $Lx = -D_2 f(0,0)$ has a solution p,

(iii) If $y \in N(L)$ and $\{D_1^2 f(0,0)p + D_1 D_2 f(0,0)\}y \in R(L)$ then $y = 0$.

Then there is a neighbourhood U of 0 in E and a subinterval $J = (-\mu_0, \mu_0)$ of 0 in $\mathbb{R}$ such that for $\mu \in J$ the equation (3.1) has a solution $x(\mu) \in U$, $\mu \in J$, which is unique when $\mu \neq 0$. Moreover $x(0) = 0$, $x(\mu)$ is C^1 and when $\mu \neq 0$, $D_1 f(x(\mu), \mu)$ is invertible.

We shall not prove the theorem above since it is a Corollary of Theorem 3.2 which we shall formulate and prove below.

Remark. By hypothesis (i) $f(x, 0) = 0$ (a constant) for all x in M so that all higher order derivatives with respect to x in the direction of $N(L) = T_0 M$ are zero in M and therefore

$$\{D_1^2 f(0,0)y\}z = 0 \tag{3.2}$$

for all $y, z \in N(L)$. It follows that if (iii) holds for one solution p of $Lx = -D_2 f(0,0)$ then it holds for all such p. To see this suppose $\tilde{p}$ is another solution of $Lx = -D_2 f(0,0)$. Then

$Lp - L\tilde{p} = -D_2 f(0,0) + D_2 f(0,0) = 0 \implies p - \tilde{p} \in N(L)$. Let $p - \tilde{p} = z \in N(L)$,

so that $p = \tilde{p} + z$. Now

$$\begin{aligned}
&\{D_1^2 f(0,0)p + D_1 D_2 f(0,0)\}y \\
&= \{D_1^2 f(0,0)(\tilde{p} + z) + D_1 D_2 f(0,0)\}y \\
&= \{D_1^2 f(0,0)\tilde{p} + D_1 D_2 f(0,0)\}y + \{D_1^2 f(0,0)z\}y \\
&= \{D_1^2 f(0,0)\tilde{p} + D_1 D_2 f(0,0)\}y \qquad \text{by (3.2) above} .
\end{aligned}$$

So if (iii) holds for one solution p of $Lx = -D_2 f(0,0)$ then it holds for all such p.

We now state a more general version of Theorem 3.1. We wish to solve $f(x, \mu\xi^0) = 0$ near $(0,0) \in E \times \mathbb{R}^N$.

Theorem 3.2. Let E, F be Banach spaces and $f : E \times \mathbb{R}^N \to F$ be a C^2–mapping on a neighbourhood of $(0,0)$ in $E \times \mathbb{R}^N$ such that $f(0,0) = 0$ and $L \triangleq D_1 f(0,0)$ is a Fredholm operator with index $I \geq 0$. Let $d \triangleq \dim N(L)$ and let S be an I–dimensional subspace of E such that $N(L) = M_0 \oplus S$. Let $\widetilde{N(L)} \subset E$ such that $E = \widetilde{N(L)} \oplus N(L)$. Moreover assume:

(i) There is a C^2–manifold $M = \{x = \xi + \phi_1(\xi) + \phi_2(\xi) : \xi \in M_0, \|\xi\| < \delta\}$, where $\phi_1 : \{\xi \in M_0 \mid \|\xi\| < \delta\} \to S$ and $\phi_2 : \{\xi \in M_0 \mid \|\xi\| < \delta\} \to \widetilde{N(L)}$ are C^2–functions satisfying $\phi_1(0) = 0$, $D\phi_1(0) = 0$, $\phi_2(0) = 0$, $D\phi_2(0) = 0$, such that $\dim M = d - I$ and $f(x,0) = 0$ for all $x \in M$.

(ii) There is $\xi^0 \in \mathbb{R}^N$ such that the equation $Lx = -D_2 f(0,0)\xi^0$ has a solution $p(\xi^0)$.

(iii) If $w \in M_0$ and $\{D_1^2 f(0,0)p(\xi^0) + D_1 D_2 f(0,0)\xi^0\}w \in R(L)$, then $w = 0$.

Then there is a neighbourhood $U$of 0 in E and a subinterval $J = (-\mu_0, \mu_0)$ of 0 in $\mathbb{R}$ such that for $\mu \in J$ the equation $f(x, \mu\xi^0) = 0$ (equation (3.1)) has a solution $x = x(\gamma, \mu) \in U$, $\gamma \in S$, $\mu \in J$ which is unique when $\mu \neq 0$. Moreover $x(\gamma, 0) = 0$, $x(\gamma, \mu)$ is C^1 and when $\mu \neq 0$, $D_1 f(x(\gamma,\mu), \mu\xi^0)$ is Fredholm with index I.

Before we give a detailed proof of Theorem 3.2 we shall have some remarks and then a motivation for the proof.

Remarks.

(1) $T_0 M = M_0 \subset N(L)$. This is seen from the following observation: $T_b M = \{y | y = \frac{d}{dt} x(t) \mid_{t=0}$ for some $x : [-1,1] \to M$ such that $x \in C^1$ and $x(0) = b\}$ [cf. definition 0.2.1]. Let x be a curve satisfying the properties of the latter definition. Then by hypothesis (i) of Theorem 3.2, we can write

$$x(t) = \xi(t) + \phi_1(\xi(t)) + \phi_2(\xi(t)) \qquad t \in [-1,1],$$

where $\xi(t) \in M_0$. Moreover $\xi(t) = P_{M_0}x(t)$ where P_{M_0} is the projection on M_0. Since P_{M_0} is bounded, it follows that $\xi(t) \in C^1$. Hence

$$\begin{aligned}\dot{x}(0) &= \frac{d\xi}{dt}(0) + D\phi_1(\xi(0))\frac{d\xi}{dt}(0) + D\phi_2(\xi(0))\frac{d\xi}{dt}(0)\\ &= \frac{d\xi}{dt}(0) \qquad \text{since } D\phi_j(\xi(0)) = D\phi_j(0) = 0, \quad j = 1,2.\end{aligned}$$

Since $\xi(t) \in M_0$ for all $t \in [-1,1]$ and since M_0 is a subspace it follows that $\frac{d\xi}{dt}(0) \in M_0$. Hence $T_0M \subseteq M_0$. To show that $M_0 \subseteq T_0M$, suppose that $m \in M_0$. Choose a C^1 curve $\xi(t) \in M_0$ such that $\xi(0) = b$ and $\frac{d\xi}{dt}(0) = m$. An argument similar to the one above then shows that if $x(t) = \xi(t) + \phi_1(\xi(t)) + \phi_2(\xi(t))$ then $x(0) = b$ and $\dot{x}(0) = \frac{d\xi}{dt}(0) = m$. Since $x(t)$ is a C^1 curve on M, it follows that $m \in T_0M$. Hence $M_0 \subseteq T_0M$.

From now on we shall write $N(L) = T_0M \oplus S$.

(2) Theorem 3.1 can be deduced from Theorem 3.2 by simply taking $S = \emptyset$. Note then that $N(L) = T_0M$ in Theorem 3.1.

(3) $d - I$ is the maximum possible dimension of M in order to have hypothesis (iii). To see this write $Hw = \{D_1^2 f(0,0)p(\xi^0) + D_1D_2f(0,0)\xi^0\}w$, $H : T_0M \to F$ and suppose $v \triangleq \dim M > d - I$. Then $\dim T_0M = \dim N(H) + \dim R(H)$ and since $\dim M = \dim T_0M$ we have $v = \dim N(H) + \dim R(H)$ and $\dim R(H) \leq v$. Now if $\dim R(H) < v$ then $\dim N(H) > 0$ implying that $N(H)$ is non–trivial and so there is $w \in N(H)$, $w \neq 0$ and then hypothesis (iii) is contradicted. Therefore $\dim R(H) = v > d - I$. $I = \dim N(L) - \operatorname{codim} R(L) \implies \operatorname{codim} R(L) = \dim N(L) - I$. That is $\operatorname{codim} R(L) = d - I$ so that $\dim R(H) = v > d - I = \operatorname{codim} R(L)$. This implies that $R(H)$ either contains a complementary subspace of $R(L)$ and if so $R(L) \cap R(H) \neq \{0\}$ or $R(H)$ is entirely contained in $R(L)$ and again $R(L) \cap R(H) \neq \{0\}$ since $\dim R(H) = v > 0$. Thus the equation $Hw^0 = z^0$

has a solution $w^0 \in T_0M$, $w^0 \neq 0$ and again this contradicts hypothesis (iii). So $d - I$ must be the maximum possible dimension of M.

(4) If $d = I$ then $\dim M = d - I = I - I = 0$. So we can take $M = \{0\}$ and hence hypothesis (i) is satisfied trivially since $f(0,0) = 0$. Furthermore if $d = I$ then since $\operatorname{codim} R(L) = d - I$ we have $\operatorname{codim} R(L) = 0$, and hence $R(L) = F$ and so hypothesis (ii) is satisfied since for $x \in E$, $Lx \in R(L) = F$. $R(L) = F$ implies that $L : E \to F$ is onto so that for each $y \in F$ there is $x \in E$ such that $Lx = y$. Now take $y = -D_2f(0,0)\xi^0$. Also hypothesis (iii) is satisfied since $\dim M = \dim T_0M = \dim M_0 = d - I = I - I = 0$, so that $M_0 = \{0\}$. $[w \in M_0 \implies w = 0]$

Motivation for the proof of Theorem 3.2

We give an outline of the proof for the case $\xi^0 = 1$, $S = \emptyset$, index $(L) = 0$. This will also motivate the proof in the general case. In this motivation the function f equals the function f defined in Chapter 2.

For a C^2–map $f : E \times \mathbb{R} \to F$ we wish to solve $f(x, \mu) = 0$ in some neighbourhood of $(0,0)$ in $E \times \mathbb{R}$. We are given that $f(0,0) = 0$. In Chapter 2 we have shown that $L \triangleq D_1f(0,0)$ is Fredholm. Note that $(Lx)(t) = \dot{x}(t) - A(t)x(t)$ where $A(t) \triangleq Dg(\gamma(t))$, has a non–trivial solution $\dot{\gamma}(t)$ so that $N(L)$ is not trivial and hence L is not invertible. We then need to use an extended version of the Implicit Function Theorem. This is achieved by the Liapunov–Schmidt Reduction Method. Since L is Fredholm we have $L : E \to F$ such that

$$E = N(L) \oplus M \quad (M \text{ a closed complementary subspace of } E),$$

$$F = N \oplus R(L) \quad (N \text{ a closed complementary subspace of } F).$$

Associated with these splittings are projections. So let

$$P : E \to E \quad \text{such that } R(P) = N(L),$$
$$Q : F \to F \quad \text{such that } R(Q) = R(L).$$

Since $E = N(P) \oplus R(P)$ and $F = N(Q) \oplus R(Q)$ it follows that

$$E = N(L) \oplus N(P),$$
$$F = N(Q) \oplus R(L).$$

L is Fredholm of index zero and hence $\dim N(L) = \dim N(Q) < \infty$. One then finds a map $D_1 D_2 H(0,0) : N(L) \to N(Q)$ that is one to one and another map $L|_{N(P)} : N(P) \to R(L)$ that is one to one. The easy way of finding $D_1 D_2 H(0,0)$ and $L|_{N(P)}$ is to make use of the Liapunov–Schmidt Reduction Method since $f(x, \mu) = 0$ gets transformed into an equivalent system of equations

$$\begin{aligned} &\text{(i)} \qquad \eta = KQh(\xi + \eta, \mu) \\ &\text{(ii)} \qquad 0 = (I - Q)h(\xi + \eta, \mu), \end{aligned} \tag{3.3}$$

where h is a function to be defined in the proof of Theorem 3.2 and K is the inverse of $L|_{N(P)}$, and from (3.3) we find out that the maps $D_1 D_2 H(0,0)$ and $L|_{N(P)}$ are in fact bijections and hence invertible. The first equation defines the map $L|_{N(P)} : N(P) \to R(Q) = R(L)$ and is a bijection so that one can apply the usual Implicit Function Theorem to conclude that $\eta = \eta(\xi, \mu)$ is a solution of (3.3)(i) in a neighbourhood of $(0,0)$ in $E \times \mathbb{R}$. The second equation defines an invertible map, $D_1 D_2 H(0,0) : N(L) \to N(Q)$ so that under the conditions of the theorem, one can apply the usual Implicit Function Theorem to conclude that $\xi = \xi(\mu)$ is a solution of (3.3)(ii) in a neighbourhood of 0 in $\mathbb{R}$. Thus for $\mu \neq 0$, $\mu \in (-\mu_0, \mu_0)$, $x(\mu) = \xi(\mu) + \eta(\xi(\mu), \mu)$ is a solution of (3.1) in a neighbourhood of 0 in E.

In Theorem 3.2 f is a C^2–map $f : E \times \mathbb{R}^N \to F$ and the spaces are split such that $E = N(P) \oplus S$ where $S \subset N(L) = T_0M \oplus S$ and $F = N(Q) \oplus R(\hat{L})$ where $\hat{L} = L|_{N(P)}$. Of course $F = N(Q) \oplus R(L)$. As mentioned in the remarks above we now find that the maps $D_1D_2H(0,0) : T_0M \to N(Q)$ and $\hat{\hat{L}} : \widetilde{N(L)} \to R(\hat{L}) = R(L)$ are bijections so that eventually we get that $x(\gamma, \mu)$ is a solution of $f(x, \mu\xi^0) = 0$. Here γ is a fixed element of $S \subset N(L)$.

We now prove Theorem 3.2.

Proof of Theorem 3.2:

The first part of the proof consists of seeking maps that satisfy the conditions stipulated in the summary above. To accomplish this we need to split our spaces accordingly.

So let $P : E \to E$ be the continuous projection such that $R(P) = S$ and $N(P) = T_0M \oplus \widetilde{N(L)}$. That is

$$\begin{aligned} E &= \widetilde{N(L)} \oplus N(L) \\ &= \widetilde{N(L)} \oplus T_0M \oplus S \\ &= N(P) \oplus S. \qquad\qquad \text{[cf. definition 0.1.6]} \end{aligned}$$

[Note that $T_0M = N(L) \cap N(P)$. In the remarks above we showed that $T_0M \subset N(L)$.]

Let $T : N(P) \to N(P)$ be the continuous projection such that $R(T) = T_0M$ and $N(T) = \widetilde{N(L)}$.

Fix $\gamma \in S$ and let $g : N(P) \times \mathbb{R} \to F$ be defined by

$$g(x_0, \mu) \triangleq f(x_0 + \mu\gamma + \phi_1(Tx_0), \mu\xi^0).$$

g also depends on the $\gamma \in S$ chosen, but we will not emphasize this.

Then

$$g(0,0) = f(0 + 0\gamma + \phi_1(T0), 0\xi^0) = f(0,0) = 0.$$

Further,

$$D_1 g(x_0, \mu) = D_1 f(x_0 + \mu\gamma + \phi_1(Tx_0), \mu\xi^0) D_{(x_0)}(x_0 + \mu\gamma + \phi_1(Tx_0))$$

so that

$$D_1 g(0,0) = D_1 f(0,0) I_{N(P)}.$$

Let $\hat{L} \triangleq L|_{N(P)}$, the restriction of L to $N(P)$. Then

$$N(D_1 g(0,0) = N(\hat{L}) = T_0 M = N(L) \cap N(P)$$

and $R(\hat{L}) = R(L|_{N(P)}) = R(L)$, $\dim N(\hat{L}) = \dim T_0 M = d - I$. $I = \dim N(L) - \operatorname{codim} R(L)$ so that $\operatorname{codim} R(L) = d - I$. Therefore $\dim N(\hat{L}) = \operatorname{codim} R(L) = \operatorname{codim} R(\hat{L}) = d - I$. Thus $\hat{L}$ is Fredholm of index zero.

Also

$$\begin{aligned} D_2 g(x_0, \mu) &= D_1 f(x_0 + \mu\gamma + \phi_1(Tx_0), \mu\xi^0) D_{(\mu)}(x_0 + \mu\gamma + \phi_1(Tx_0)) \\ &\quad + D_2 f(x_0 + \mu\gamma + \phi_1(Tx_0), \mu\xi^0) D_{(\mu)}(\mu\xi^0). \end{aligned}$$

Therefore

$$\begin{aligned} D_2 g(0,0) &= D_1 f(0,0)\gamma + D_2 f(0,0)\xi^0 \\ &= D_2 f(0,0)\xi^0 \end{aligned}$$

since $L = D_1 f(0,0)$ and $\gamma \in S \subset N(L)$.

By the definition of f and g it suffices to solve the equation

$$g(x_0, \mu) = 0. \tag{3.4}$$

Write (3.4) as

$$\hat{L}x_0 = h(x_0, \mu) \tag{3.5}$$

where

$$\begin{aligned} h(x_0, \mu) &= D_1 g(0,0)x_0 - g(x_0, \mu) \\ &= \hat{L}x_0 - f(x_0 + \mu\gamma + \phi_1(Tx_0), \mu\xi^0) \end{aligned}$$

so that $h(x_0, \mu) = \hat{L}x_0$ if and only if $g(x_0, \mu) = 0$.

Let $Q : F \to F$ be the continuous projection such that $R(Q) = R(L) = R(\hat{L})$. Any $x_0 \in N(P)$ can be written as $\xi + \eta = Tx_0 + (I - T)x_0$. Now (3.5) can be written as

$$\begin{aligned} \hat{L}(\xi + \eta) &= h(\xi + \eta, \mu) \\ \hat{L}\xi + \hat{L}\eta &= h(\xi + \eta, \mu) \\ \hat{L}\eta &= h(\xi + \eta, \mu) \qquad [\xi = Tx_0 \in R(T) = T_0M = N(\hat{L})] \\ Q\hat{L}\eta &= Qh(\xi + \eta, \mu) \\ \hat{L}\eta &= Qh(\xi + \eta, \mu) \qquad [R(Q) = R(\hat{L})]. \end{aligned}$$

Let $\hat{\hat{L}} = \hat{L}|_{\widetilde{N(L)}}$. Now $\hat{\hat{L}} : \widetilde{N(L)} \to R(\hat{L}) = R(L)$ is a bijection and so has a bounded inverse K from $R(\hat{L})$ to $\widetilde{N(L)}$ since $R(\hat{L})$ is closed [cf. Theorem 0.1.20]. Thus $\eta = \hat{\hat{L}}^{-1} Qh(\xi + \eta, \mu) = KQh(\xi + \eta, \mu)$.

Further,

$$(I - Q)\hat{L}\eta = (I - Q)h(\xi + \eta, \mu)$$

so that

$$0 = (I - Q)h(\xi + \eta, \mu) \qquad [\hat{L}\eta \in R(\hat{L}) = R(Q) = N(I - Q)].$$

So (3.5) implies the system of two equations

$$\eta = KQh(\xi + \eta, \mu) \tag{3.6}$$

$$0 = (I - Q)h(\xi + \eta, \mu). \tag{3.7}$$

Conversely, if ξ, η are any solutions of these equations then $x_0 = \xi + \eta$ is a solution of (3.5), as is easily seen.

Let $\tilde{h}(\xi, \eta, \mu) = \eta - KQh(\xi + \eta, \mu)$. Then $\tilde{h}$ is a C^2–function since h is. [This follows from $KQh(\xi+\eta, \mu) = KQ(\hat{L}\eta - g(\xi+\eta, \mu) = \eta - KQg(\xi+\eta, \mu)$ and $g \in C^2$ since $f \in C^2$.]

$$\begin{aligned}\tilde{h}(0,0,0) = 0 - KQh(0+0,0) &= -KQh(0,0)\\ &= -KQ[\hat{L}0 - g(0,0)]\\ &= 0\end{aligned}$$

since $\hat{L}, K, Q$ are linear and $g(0,0) = 0$.

$$\begin{aligned}D_2\tilde{h}(0,0,0) &= D_{(\eta)}[\eta - KQh(\xi + \eta, \mu)]|_{(0,0,0)}\\ &= I_{\widetilde{N(L)}} - KQD_{(\eta)}h(\xi + \eta, \mu)]|_{(0,0,0)}\\ &= I_{\widetilde{N(L)}} - KQD_{(\eta)}[\hat{L}(\xi + \eta) - f(\xi + \eta + \mu\gamma + \phi_1(T(\xi + \eta)), \mu\xi^0)]|_{(0,0,0)}\\ &= I_{\widetilde{N(L)}} - KQD_{(\eta)}[\hat{L}\eta - f(\xi + \eta + \mu\gamma + \phi_1(\xi), \mu\xi^0)]|_{(0,0,0)}\\ &= I_{\widetilde{N(L)}} - KQ[\hat{L} - D_1 f(\xi + \eta + \mu\gamma + \phi_1(\xi), \mu\xi^0)D_{(\eta)}(\xi + \eta + \mu\gamma + \phi_1(\xi))]|_{(0,0,0)}\\ &= I_{\widetilde{N(L)}} - KQ[\hat{L} - D_1 f(0,0)]I_{\widetilde{N(L)}}\\ &= I_{\widetilde{N(L)}} - KQ(\hat{\hat{L}} - \hat{\hat{L}})\\ &= I_{\widetilde{N(L)}}.\end{aligned}$$

Therefore $D_2\tilde{h}(0,0,0)$ is invertible in $\widetilde{N(L)}$. The hypothesis of the Implicit Function Theorem on Banach spaces [cf. Theorem 0.2.2] then is satisfied and so we deduce that there is a neighbourhood U of zero in $T_0M \times J$ and a neighbourhood V of zero in $\widetilde{N(L)}$ and a unique map $\eta : U \to V$ such that $\eta(0,0) = 0$ and $\tilde{h}(\xi, \eta(\xi,\mu), \mu) = 0$ on U. The map η is a C^2–function. Thus (3.6) has a unique solution $\eta(\xi,\mu)$ in V.

Substitute $\eta(\xi,\mu)$ in (3.7) to get

$$H(\xi,\mu) \triangleq (I-Q)h(\xi + \eta(\xi,\mu),\mu). \tag{3.8}$$

Let $x_0 \in M_1 = \{x_0 \mid x_0 = \xi + \phi_2(\xi) \text{ for some } \xi \in V \subset T_0M\}$. Note that $M_1 \subset M$. Since $\xi, \phi_2(\xi) \in N(P)$, M_1 is a local C^2–manifold of $N(P)$ and for any $x_0 \in M_1$ we have

$$\begin{aligned} g(x_0, 0) &= f(\xi + \phi_2(\xi) + \phi_1(T(\xi + \phi_2(\xi))), 0) \\ &= f(\xi + \phi_2(\xi) + \phi_1(\xi), 0) \\ &= 0. \end{aligned}$$

[$T(\xi + \phi_2(\xi)) = T\xi + T\phi_2(\xi) = T\xi = \xi$ and by (i) $f(x,0) = 0$ for $x \in M_1 \subset M$.]

Thus for $\mu = 0$, $x_0 \in M_1$ satisfies (3.4) and hence (3.5). It follows then that (3.6) and (3.7) must also be satisfied for $\mu = 0$ with $\eta = \phi_2(\xi)$. If δ is sufficiently small, it follows from the uniqueness of $\eta(\xi,\mu)$ that for $\|\xi\| < \delta$, $\eta(\xi,0) = \phi_2(\xi)$ and

$$\begin{aligned} H(\xi,0) &= (I-Q)h(\xi + \eta(\xi,0),0) \\ &= (I-Q)h(\xi + \phi_2(\xi),0) \\ &= (I-Q)[\hat{L}(\xi + \phi_2(\xi)) - f(\xi + \phi_1(\xi) + \phi_2(\xi),0)] \\ &= (I-Q)[\hat{L}(\xi + \phi_2(\xi))] \quad \text{since } \ f(x,0) = 0 \text{ for } x \in M \\ &= 0 \qquad \text{since } \hat{L}(\xi + \phi_2(\xi)) \in R(\hat{L}) = R(Q) = N(I-Q). \end{aligned}$$

Hence $H(0,0) = 0$.

$$\begin{aligned}
D_1H(\xi,\mu) &= (I-Q)D_{(\xi)}h(\xi+\eta(\xi,\mu),\mu)\\
&= (I-Q)D_{(\xi)}\left[\hat{L}(\xi+\eta(\xi,\mu)) - f(\xi+\eta(\xi,\mu)+\mu\gamma+\phi_1(\xi),\mu\xi^0)\right]\\
&= (I-Q)[\hat{L}+\hat{L}D_1\eta(\xi,\mu) - D_1f(\xi+\eta(\xi,\mu)+\mu\gamma+\phi_1(\xi),\mu\xi^0)\\
&\qquad D_{(\xi)}(\xi+\eta(\xi,\mu)+\mu\gamma+\phi_1(\xi))]\\
&= (I-Q)\left[\hat{L}+\hat{L}D_1\eta(\xi,\mu) - D_1f(\xi+\eta(\xi,\mu)+\mu\gamma+\phi_1(\xi),\mu\xi^0)\right.\\
&\qquad \left.(I_{T_0M}+D_1\eta(\xi,\mu)+D\phi_1(\xi))\right].\\
D_1H(0,0) &= (I-Q)[\hat{L}+\hat{L}D_1\eta(0,0) - \hat{L}(I_{T_0M}+D_1\eta(0,0)]\\
&= (I-Q)[\hat{L}+\hat{L}D_1\eta(0,0) - \hat{L} - \hat{L}D_1\eta(0,0)]\\
&= (I-Q)(0)\\
&= 0.
\end{aligned}$$

This shows that $D_1H(0,0)$ is not invertible and so we cannot apply the Implicit Function Theorem. We shall introduce an auxilliary function such that this function vanishes for (ξ,μ), $\mu \neq 0$ if and only if the function H vanishes. This function will have a derivative with respect to ξ that is invertible. Finding the zeros of this new function then will be equivalent to finding the zeros of H.

Let

$$\tilde{H}(\xi,\mu) = \begin{cases} \frac{1}{\mu}H(\xi,\mu) & \text{for } \mu \neq 0\\ D_2H(\xi,0) & \text{for } \mu = 0\end{cases}$$

$\tilde{H}(\xi,\mu)$ is C^1 with respect to the second variable μ.

This is clear for $\mu \neq 0$. For $\mu = 0$:

$$\frac{1}{\mu}(\tilde{H}(\xi,\mu) - \tilde{H}(\xi,0)) = \frac{1}{\mu}(\frac{1}{\mu}H(\xi,\mu) - D_2H(\xi,0)).$$

By Theorems 0.2.4 and 0.2.5, since $H(\xi,\mu) \in C^2$,

$$H(\xi,\mu) = H(\xi,0) + \mu D_2H(\xi,0) + \int_0^1 (1-\alpha)D_2^2H(\xi,\alpha\mu)d\alpha.$$

Hence, since $H(\xi, 0) = 0$,

$$\begin{aligned} D_2\tilde{H}(\xi,\mu)|_{\mu=0} &= \lim_{\mu\to 0}\frac{1}{\mu}(\tilde{H}(\xi,\mu)-\tilde{H}(\xi,0)) \\ &= \lim_{\mu\to 0}\int_0^1 (1-\alpha)D_2^2H(\xi,\alpha\mu)d\alpha \\ &= \frac{1}{2}D_2^2H(\xi,0). \end{aligned} \tag{3.9}$$

Hence $\tilde{H}(\xi,\mu)$ is differentiable at $\mu = 0$. To show that the derivative is continuous: For $\mu \neq 0$ we have

$$\begin{aligned} D_2\tilde{H}(\xi,\mu) &= -\frac{1}{\mu^2}H(\xi,\mu) + \frac{1}{\mu}D_2H(\xi,\mu) \\ &= -\frac{1}{\mu}(\frac{1}{\mu}H(\xi,\mu) - D_2H(\xi,\mu)) \\ &= -\frac{1}{\mu}(D_2H(\xi,0) + \mu\int_0^1 (1-\alpha)D_2^2H(\xi,\alpha\mu)d\alpha - D_2H(\xi,\mu)). \end{aligned}$$

But since $D_2H(\xi,\mu)$ is C^1 in μ,

$$D_2H(\xi,\mu) = D_2H(\xi,0) + \mu\int_0^1 D_2^2H(\xi,\alpha'\mu)d\alpha'.$$

Hence,

$$\begin{aligned} \lim_{\mu\to 0} D_2\tilde{H}(\xi,\mu) &= \lim_{\mu\to 0} -\frac{1}{\mu}(\mu\int_0^1 (1-\alpha)D_2^2H(\xi,\alpha\mu)d\alpha \\ &\qquad - \mu\int_0^1 D_2^2H(\xi,\alpha'\mu)d\alpha') \\ &= \frac{1}{2}D_2^2H(\xi,0) \qquad \text{by (3.9)}. \end{aligned}$$

Hence $D_2\tilde{H}(\xi,\mu)$ is continuous at $\mu = 0$.

Solving $H(\xi,\mu) = 0$ is equivalent to solving

$$\tilde{H}(\xi,\mu) = 0. \tag{3.10}$$

We now show that $\tilde{H}(0,0) = 0$ and $D_1\tilde{H}(\xi,0)$ is invertible.

$$\begin{aligned}
\tilde{H}(\xi,0) &= D_2H(\xi,0) \\
&= (I-Q)D_{(\mu)}h(\xi+\eta(\xi,0),0) \\
&= (I-Q)D_{(\mu)}[\hat{L}(\xi+\eta(\xi,0)) - f(\xi+\eta(\xi,0)+\phi_1(\xi),0)] \\
&= (I-Q)\Big[\hat{L}D_2\eta(\xi,0) - D_1f(\xi+\eta(\xi,0)+\phi_1(\xi),0)D_2\eta(\xi,0) \\
&\qquad -D_2f(\xi+\eta(\xi,0)+\phi_1(\xi),0)D_{(\mu)}(\mu\xi^0)] \\
&= (I-Q)\Big[\hat{L}D_2\eta(\xi,0) - D_1f(\xi+\eta(\xi,0)+\phi_1(\xi),0)D_2\eta(\xi,0) \\
&\qquad -D_2f(\xi+\eta(\xi,0)+\phi_1(\xi),0)\xi^0\Big]. \qquad (3.11)
\end{aligned}$$

Also

$$\begin{aligned}
\tilde{H}(0,0) = D_2H(0,0) &= (I-Q)\Big[\hat{L}D_2\eta(0,0) - D_1f(0,0)D_2\eta(\xi,0) \\
&\qquad -D_2f(0,0)\xi^0\Big] \\
&= -(I-Q)D_2f(0,0)\xi^0 \\
&= 0.
\end{aligned}$$

[The second equality holds by (3.11). The last equality holds since by hypothesis (ii) $Lx = -D_2f(0,0)\xi^0$ has a solution and so $D_2f(0,0)\xi^0 \in R(L) = R(Q) = N(I-Q)$.]

Differentiate equation (3.11) with respect to ξ to get

$$\begin{aligned}
D_1\tilde{H}(\xi,0) &= D_1D_2H(\xi,0) \\
&= (I-Q)\Big[\hat{L}D_1D_2\eta(\xi,0) - (D_1^2f(\xi+\eta(\xi,0)+\phi_1(\xi),0) \\
&\qquad D_{(\xi)}(\xi+\eta(\xi,0)+\phi_1(\xi))D_2\eta(\xi,0) \\
&\quad - D_1f(\xi+\eta(\xi,0)+\phi_1(\xi),0)D_1D_2\eta(\xi,0) \\
&\quad - D_1D_2f(\xi+\eta(\xi,0)+\phi_1(\xi),0)D_{(\xi)}(\xi+\eta(\xi,0)+\phi_1(\xi))\xi^0\Big]
\end{aligned}$$

so that

$$D_1D_2H(0,0) = (I-Q)\Big[\hat{L}D_1D_2\eta(0,0) - D_1^2f(0,0)\{I_{T_0M} + D_1\eta(0,0) + D\phi_1(0)\}D_2\eta(0,0) - D_1f(0,0)D_1D_2\eta(0,0) - D_1D_2f(0,0)\{I_{T_0M} + D_1\eta(0,0) + D\phi_1(0)\}\xi^0\Big].$$

That is,

$$\begin{aligned} D_1D_2H(0,0) &= (I-Q)[\hat{L}D_1D_2\eta(0,0) - D_1^2f(0,0)D_2\eta(0,0) \\ &\quad - \hat{L}D_1D_2\eta(0,0) - D_1D_2f(0,0)\xi^0] \\ &= -(I-Q)[D_1^2f(0,0)D_2\eta(0,0) - D_1D_2f(0,0)\xi^0] \end{aligned}$$

since $D\phi_1(0) = 0$ and $D_1\eta(0,0) = D\phi_2(0) = 0$.

For $w \in T_0M$ we obtain

$$D_1D_2H(0,0)w = -(I-Q)[D_1^2f(0,0)D_2\eta(0,0) + D_1D_2f(0,0)\xi^0]w. \tag{3.12}$$

Differentiate (3.6) with respect to μ to get

$$\begin{aligned} D_2\eta(\xi,\mu) &= KQD_{(\mu)}h(\xi+\eta(\xi,\mu),\mu) \\ &= KQD_{(\mu)}[\hat{L}(\xi+\eta(\xi,\mu) - f(\xi+\eta(\xi,\mu)+\mu\gamma+\phi_1(\xi),\mu\xi^0)] \\ &= KQ[\hat{L}D_2\eta(\xi,\mu) - D_1f(\xi+\eta(\xi,\mu)+\mu\gamma+\phi_1(\xi),\mu\xi^0) \\ &\qquad D_{(\mu)}(\xi+\eta(\xi,\mu)+\mu\gamma+\phi_1(\xi)) - D_2f(\xi+\eta(\xi,\mu)+\mu\gamma+\phi_1(\xi),\mu\xi^0)D_{(\mu)}(\mu\xi^0)] \\ &= KQ[\hat{L}D_2\eta(\xi,\mu) - D_1f(\xi+\eta(\xi,\mu)+\mu\gamma+\phi_1(\xi),\mu\xi^0) \\ &\qquad \{D_2\eta(\xi,\mu)+\gamma\} - D_2f(\xi+\eta(\xi,\mu)+\mu\gamma+\phi_1(\xi),\mu\xi^0)\xi^0] \end{aligned}$$

so that

$$\begin{aligned} D_2\eta(0,0) &= KQ[\hat{L}D_2\eta(0,0) - D_1f(0,0)D_2\eta(0,0) - D_2f(0,0)\xi^0] \\ &= -KQD_2f(0,0)\xi^0 \\ &= -KD_2f(0,0)\xi^0 \end{aligned}$$

since $L\gamma = 0$ and $D_2 f(0,0)\xi^0 \in R(\hat{L}) = R(Q)$ by (ii).

Remember $K = \hat{\hat{L}}^{-1}$ so that we have

$$LD_2\eta(0,0) = \hat{\hat{L}}D_2\eta(0,0) = -D_2 f(0,0)\xi^0 \tag{3.13}$$

since $R(\hat{\hat{L}}) = R(L)$.

Now we see that $D_2\eta(0,0)$ is a solution of $Lp = -D_2 f(0,0)\xi^0$. By (3.12), (3.13), remark after Theorem 3.1 and hypothesis (iii) $D_1 D_2 H(0,0)w = 0$ only if $w = 0$.

Now use the following observations:

$$F = R(\hat{L}) \oplus N(Q),$$

$\dim T_0 M = \dim N(Q) < \infty$ since $\hat{L}$ is Fredholm of index zero. Further $D_1 D_2 H(0,0) : T_0 M \to N(Q)$, and $\dim T_0 M = d - I = \dim N(Q) < \infty$. So that the range and domain of $D_1 D_2 H(0,0)$ are both finite dimensional, with the same dimension. $N(D_1 D_2 H(0,0)) = \{0\}$ and so $D_1 D_2 H(0,0)$ is invertible.

We then use the Implicit Function Theorem [cf. Theorem 0.2.2] to deduce that there is a neighbourhood $\tilde{U}$ of zero in J and a neighbourhood $\tilde{V}$ of zero in $T_0 M$ and a unique map $\xi : \tilde{U} \to \tilde{V}$ such that $\xi(0) = 0$ and $\tilde{H}(\xi(\mu), \mu) = D_2 H(\xi(\mu), 0) = 0$ on $\tilde{U}$. The map ξ is a C^1–function. Thus (3.10) has a unique solution $\xi(\mu)$ in $\tilde{V}$. Let

$$\begin{aligned}
x &\triangleq \xi(\mu) + \eta(\xi(\mu), \mu) + \mu\gamma + \phi_1(T(\xi(\mu) + \eta(\xi(\mu), \mu))) \\
&= \xi(\mu) + \eta(\xi(\mu), \mu) + \mu\gamma + \phi_1(T\xi(\mu)) \text{ since } \eta \in \widetilde{N(L)} = N(T) \\
&= \xi(\mu) + \eta(\xi(\mu), \mu) + \mu\gamma + \phi_1(\xi(\mu)) \text{ since } \xi \in T_0 M = R(T) \\
&\triangleq x(\gamma, \mu).
\end{aligned}$$

Then $x(\gamma, \mu)$ is a solution of (3.1) and is unique in a neighbourhood of 0 in E when $\mu \neq 0$. Moveover $x(\gamma, 0) = 0$, and $x(\gamma, \mu)$ is C^1.

This completes the proof of the first part.

The proof of the second part of this theorem will consist of showing that the function K to be obtained below satisfies the hypothesis of this theorem. Before we do this we have the following remark.

Remark. If we do not assume that f is C^2 with respect to μ the above proof still goes through except that we can only conclude that $x(\gamma,\mu)$ is continuous. This follows from the following observation:
If f is C^1 with respect to μ, then so is H. Therefore by definition of $\tilde{H}$ one can show that $\tilde{H}$ is C^0. The application of the Implicit Function Theorem thus yields the fact that ξ is a C^0–function of μ and hence by the definition of x we get that $x(\gamma,\mu) \in C^0$.

We now show that when $\mu \neq 0$, $D_1 f(x(\gamma,\mu),\mu\xi^0) \triangleq L_\mu$ is an invertible Fredholm operator.

Consider

$$\begin{aligned} D_1 g(x_0,\mu) &= D_1 f(x_0+\mu\gamma+\phi_1(Tx_0),\mu\xi^0)D_{(x_0)}(x_0+\mu\gamma+\phi_1(Tx_0)) \\ &= D_1 f(x_0+\mu\gamma+\phi_1(Tx_0),\mu\xi^0)\{I_{N(P)}+D\phi_1(Tx_0)T\}. \end{aligned}$$

Substitute $x_0(\gamma,\mu)$ for x_0 where $x_0(\gamma,\mu)=\xi(\mu)+\eta(\xi,\mu)$. Then we have

$$\begin{aligned} D_1 g(x_0(\gamma,\mu),\mu) &= D_1 f(x_0(\gamma,\mu)+\mu\gamma+\phi_1(Tx_0(\gamma,\mu)),\mu\xi^0) \\ &\qquad \{I_{N(P)}+D\phi_1 Tx_0(\gamma,\mu)T\} \\ &= D_1 f(x(\gamma,\mu),\mu\xi^0)|_{N(P)}[I_{N(P)}+D\phi_1 Tx_0(\gamma,\mu)T]. \end{aligned}$$

Also note that $D\phi_1(Tx_0(\gamma,\mu))T \to D\phi_1(Tx_0(\gamma,0))T = 0$ as $\mu \to 0$ and $D_1 f(x(\gamma,\mu),$ $\mu\xi^0)|_{N(P)} \triangleq \hat{L}_\mu \to \hat{L} \triangleq D_1 f(0,0)$ in norm by the remark following Theorem 0.2.3 as $\mu \to 0$ so that there exist $\delta > 0$ such that $\left\|\hat{L}_\mu - \hat{L}\right\| < \delta$. Let $\hat{L}_\mu = \hat{L} + (\hat{L}_\mu - \hat{L}) =$

$\hat{L} + \hat{K}$. Since $\hat{K} \subset \mathcal{L}(N(P), F)$, i.e., $\hat{K}$ belongs to the Banach space of bounded linear operators that map $N(P)$ to F, perturbation theory says $i(\hat{L}_\mu) = i(\hat{L})$, that is, $\hat{L}_\mu$ and $\hat{L}$ have the same index and if $\hat{L}$ is Fredholm then so is $\hat{L}_\mu$ [cf. Theorem 0.1.19]. For $\mu \in J = (-\mu_0, \mu_0)$, $\mu \neq 0$, $\hat{L}_\mu$ is invertible if it has a trivial nullspace in $N(P)$. We need to show that

$$\begin{aligned} \hat{L}_\mu y &= D_1 g(x_0(\gamma,\mu),\mu)y \\ &= D_1 f(x(\gamma,\mu),\mu\xi^0)|_{N(P)} y = 0 \end{aligned} \tag{3.14}$$

has a unique solution $y = 0$ in $N(P)$ for $\mu \in J$, $\mu \neq 0$. $x(\gamma,\mu)$ is a solution of (3.1) so that $f(x(\gamma,\mu),\mu\xi^0) = 0$. Differentiate $f(x(\gamma,\mu),\mu\xi^0) = 0$ with respect to $\mu \in J$ to obtain

$$D_1 f(x(\gamma,\mu),\mu\xi^0) D_2 x(\gamma,\mu) + D_2 f(x(\gamma,\mu),\mu\xi^0)\xi^0 = 0.$$

At $\mu = 0$ we obtain

$$D_1 f(x(\gamma,0),0) D_2 x(\gamma,0) + D_2 f(x(\gamma,0),0)\xi^0 = 0,$$

and $x(\gamma,0) = 0$ so that

$$D_1 f(0,0) D_2 x(\gamma,0) = -D_2 f(0,0)\xi^0 = 0.$$

Hence $D_2 x(\gamma,0)$ is a solution of $Lx = -D_2 f(0,0)\xi^0$ since $L = D_1 f(0,0)$.

Let $K(y,\mu\xi^0) \triangleq D_1 f(x(\gamma,\mu),\mu\xi^0)y$. Note that $D_2^2 K$ may not exist.

$$D_2 K(y,\mu\xi^0)\xi^0 = [D_1^2 f(x(\gamma,\mu),\mu\xi^0) D_2 x(\gamma,\mu) + D_1 D_2 f(x(\gamma,\mu),\mu\xi^0)\xi^0]y.$$

Now if f is C^2, $D_2^2 K$ will involve derivatives of f of order higher than two. However, owing to the remark above we will still be able to show that (3.14) has a unique continuous solution. Let $y \in M = T_0 M$ then

$$K(y,0) = D_1 f(x(\gamma,0),0)y = D_1 f(0,0)y = Ly = 0$$

and so K satisfies hypothesis (i).

$$D_2K(0,0)\xi^0 = [D_1^2 f(0,0)D_2x(\gamma,0) + D_1D_2f(0,0)\xi^0] * 0 = 0$$

so that the equation $Lx = -D_2K(0,0)\xi^0 = 0$ has a solution $p \in T_0M$ and so K satisfies hypothesis (ii).

If $Lp = 0$ and $w \in T_0M$ we have

$$\begin{aligned} D_1K(y,\mu\xi^0) &= D_1f(x(\gamma,\mu),\mu\xi^0)|_{N(P)} = \hat{L}_\mu \\ D_1^2K(y,\mu\xi^0) &= 0 \\ D_1D_2K(y,\mu\xi^0)\xi^0 &= D_1^2 f(x(\gamma,\mu),\mu\xi^0)D_2x(\gamma,\mu) + D_1D_2f(x(\gamma,\mu),\mu\xi^0)\xi^0. \end{aligned}$$

Therefore $D_1^2K(0,0) = 0$ and

$$D_1D_2K(0,0)\xi^0 = D_1^2 f(0,0)D_2x(\gamma,0) + D_1D_2f(0,0)\xi^0$$

so that

$$\{D_1^2K(0,0)p + D_1D_2K(0,0)\xi^0\}w = \{D_1^2 f(0,0)D_2x(\gamma,0) + D_1D_2f(0,0)\xi^0\}w$$

and so K satisfies hypothesis (iii).

The conditions of Theorem 3.2 are satisfied and so we conclude that there is a neighbourhood of 0 in $N(P) \subset E$ such that for $\mu \in J$, $\mu \neq 0$ the equation $D_1f(x(\gamma,\mu),\mu\xi^0)y = 0$ has a solution $\tilde{y}(\gamma,\mu)$ that is unique in this neighbourhood. This equation is linear in y and so the unique solution must be $y = 0$ for $\mu \neq 0$, $\mu \in J$. Thus $D_1f(x(\gamma,\mu),\mu\xi^0)$ is invertible. This completes the proof of Theorem 3.2. □

In order to apply Theorem 3.2 to a system of ODE's one has to find a suitable submanifold M with the correct dimension. However, when autonomous systems are considered, one usually gets $\dim M = 1$ and $d - I$ may be greater than one. To get the correct dimension we can sometimes, in a multiparameter case, use some of the parameters to obtain M and then apply Theorem 3.2. We have the following theorem.

Theorem 3.3. Let E, F be Banach spaces, and $f : E \times \mathbb{R}^N \to F$ be a C^2–mapping on a neighbourhood of $(0,0)$ in $E \times \mathbb{R}^N$ such that $f(0,0) = 0$ and the linear operator $L \triangleq D_1 f(0,0)$ is a Fredholm operator with index $I \geq 0$. Let $d = \dim N(L)$. Let $S_0 \subset E$, $\dim S_0 = d - I$ and $S_1 \subset E$, $\dim S_1 = I$ be such that $N(L) = S_0 \oplus S_1$. Let $\widetilde{N(L)} \subset E$ be such that $E = \widetilde{N(L)} \oplus N(L)$. Moreover assume:

(i) There exists a linear subspace $W \subset \mathbb{R}^N$, $\dim W = d - I$, such that $\xi \in W$ and $D_2 f(0,0)\xi \in R(L) \implies \xi = 0$.

(ii) There exists $\xi^0 \in \mathbb{R}^N$ such that $D_2 f(0,0)\xi^0 \neq 0$ and the equation $Lx = -D_2 f(0,0)\xi^0$ has a solution $p(\xi^0) \in E$.

(iii) If $w \in S_0$ and $\{D_1^2 f(0,0)p(\xi^0) + D_1 D_2 f(0,0)\}w \in R(L)$ then $w = 0$.

Then there exists $\mu_0 > 0$, a neighbourhood U of zero in S_1, and maps $x : U \times J \to E$, $\xi : U \times J \to \mathbb{R}^N$ such that $\xi(\gamma, 0) \equiv 0$, $D_2\xi(\gamma, 0) = \xi^0$, $x(\gamma, 0) \equiv 0$, $f(x(\gamma, \mu), \xi(\gamma, \mu)) \equiv 0$, for any $(\gamma, \mu) \in U \times J$ where $J = (-\mu_0, \mu_0)$. Moreover $D_1 f(x(\gamma, \mu), \xi(\gamma, \mu)) : E \to F$ is an invertible Fredholm operator with index I, for any $\mu \neq 0$, $|\mu| < \mu_0$.

Remark. From the assumptions of the Theorem, it follows that

(1) $F = R(L) \oplus \widetilde{R(L)}$: From (i) $\xi \in W$ and $D_2 f(0,0)\xi = 0 \implies \xi = 0$. Hence since W is finite dimensional $D_2 f(0,0)|_W$ is one to one. Hence $\dim(D_2 f(0,0)|_W) = \dim W = d - I = \operatorname{codim} R(L)$. Also from (i) $R(D_2 f(0,0)|_W) \cap R(L) = \{0\}$. Let $\widetilde{R(L)} = R(D_2 f(0,0)|_W)$. Then $F = R(L) \oplus \widetilde{R(L)}$ by Theorem 0.1.21.

(2) $N > d - I$: This is seen from the following observation. Since $D_2f(0,0): \mathbb{R}^N \to R(L) \oplus \widetilde{R(L)}$ we have $R(D_2f(0,0)) \subset R(L) \oplus \widetilde{R(L)}$. But $\widetilde{R(L)} = R(D_2f(0,0)|_W)$ so that $\widetilde{R(L)} \subset R(D_2f(0,0))$ and hence $\dim R(D_2f(0,0)) \geq d - I$. Also $0 \neq D_2f(0,0)\xi^0 \in R(D_2f(0,0))$, $D_2f(0,0)\xi^0 \notin \widetilde{R(L)}$ since (i) and (ii) $\implies \xi^0 \notin W$. Hence $\dim R(D_2f(0,0)) > d - I$, and $N = \dim R(D_2f(0,0)) + \dim N(D_2f(0,0)) > d - I + \dim N(D_2f(0,0))$. Thus $N > d - I$.

Proof. We can suppose that $W = \mathbb{R}^{d-I} \times \{0\}$ and $D_2f(0,0)\xi \in R(L) \implies \xi \in \{0\} \times \mathbb{R}^{N-d+I}$, so that $\xi = (0, \xi'')$ where $\xi'' \in \mathbb{R}^{N-d+I}$.

Let us introduce a function Φ. This function will have properties that we will need later in the proof of this theorem. So let $T = \widetilde{N(L)} \times \mathbb{R}^{d-I}$ and $\Phi : S_0 \times T \to F$ be defined by $\Phi(\xi_0, (\eta, \xi')) = f(\xi_0 + \eta, (\xi', 0))$. Observe that $\xi_0 \in S_0$, $(\eta, \xi') \in T = \widetilde{N(L)} \times \mathbb{R}^{d-I}$.

Since f is C^2 so is Φ and $\Phi(0,0) = f(0{+}0, (0,0)) = f(0,0) = 0$ and $D_2\Phi(\xi_0, (\eta, \xi'))(\eta_1, \xi_1') = D_1f(\xi_0 + \eta, (\xi', 0))\eta_1 + D_2f(\xi_0 + \eta, (\xi', 0))\xi_1'$.

Therefore $D_2\Phi(0,0)(\eta_1, \xi_1') = D_1f(0,0)\eta_1 + D_2f(0,0)\xi_1'$. Since (η_1, ξ_1') is arbitrary and $D_1f(0,0) = L$ we have

$$\begin{aligned} D_2\Phi(0,0)(\eta, \xi') &= L\eta + D_2f(0,0)\xi' \\ &= L\eta \text{ since } \xi' \in W \text{ and thus } D_2f(0,0)\xi' = 0 \end{aligned}$$

$\eta \in \widetilde{N(L)}$ and $L\eta \in R(L)$. $L_{\widetilde{N(L)}} : \widetilde{N(L)} \to R(L)$ is a bijection [cf. Theorem 0.1.20] and hence $D_2\Phi(0,0)$ is invertible. The Implicit Function Theorem [cf. Theorem 0.2.2] says there exist unique maps $\eta : S_0 \to \widetilde{N(L)}$ and $\xi' : S_0 \to \mathbb{R}^{d-I}$ defined in a neighbourhood U_0 of zero in S_0 such that

$$\Phi(\xi_0, (\eta(\xi_0), \xi'(\xi_0))) = 0. \tag{3.15}$$

Also

$$\begin{aligned} D_1\Phi(\xi_0,(\eta,\xi')) &= D_1 f(\xi_0+\eta,(\xi',0)D_{(\xi_0)}(\xi_0+\eta)) \\ &= D_1 f(\xi_0+\eta,(\xi',0))I_{S_0}. \end{aligned}$$

Therefore

$$D_1\Phi(0,0)\xi_0 = D_1 f(0,0)\xi_0 = L\xi_0 = 0 \tag{3.16}$$

since $\xi_0 \in S_0 \subset N(L)$. Also $\chi(\xi_0) \triangleq \Phi(\xi_0,(\eta(\xi_0),\xi'(\xi_0))) \equiv 0$ for all ξ_0 in S_0 such that $|\xi_0| < \delta$.

$$D\chi(\xi_0) = D_1\Phi(\xi_0,(\eta(\xi_0),\xi'(\xi_0))) + D_2\Phi(\xi_0,(\eta(\xi_0),\xi'(\xi_0)))D_{(\xi_0)}(\eta(\xi_0),\xi'(\xi_0)).$$

At $\xi_0 = 0$ in the direction of any ξ_0 in S_0 we have

$$0 \equiv D\chi(0)\xi_0 = [D_1\Phi(0,0) + D_2\Phi(0,0)(D\eta(0),D\xi'(0)]\xi_0. \tag{3.17}$$

By (3.16) and (3.17) we have

$$0 = D_2\Phi(0,0)(D\eta(0),D\xi'(0))\xi_0.$$

Now $D_2\Phi(0,0)$ is a bijection as proved above and hence we must have $(D\eta(0),D\xi'(0)) = (0,0)$, that is,

$$D\eta(0) = 0, D\xi'(0) = 0. \tag{3.18}$$

Let $P : E \to E$ be the projection such that $R(P) = S_0$ and $N(P) = S_1 \oplus \widetilde{N(L)}$ and define $G : E \times \mathbb{R}^{N-d+I} \to F$ by $G(x,\xi'') = f(x,(\xi'(Px),\xi''))$. We will show that G satisfies the hypothesis of Theorem 3.2.

$G(0,0) = f(0,(\xi'(P0),0)) = f(0,(\xi'(0),0)) = f(0,(0,0)) = f(0,0) = 0$ since $P0 = 0$, $\xi'(0) = 0$ and $f(0,0) = 0$.

$$\begin{aligned} D_1G(x,\xi'')x' &= [D_1f(x,(\xi'(Px),\xi'')) + D_2f(x,(\xi'(Px),\xi''))D_{(x)}(\xi'(Px),\xi'')]x' \\ &= [D_1f(x,(\xi'(Px),\xi'')) + D_2f(x,(\xi'(Px),\xi''))(D\xi'(Px)P,\xi'')]x'. \\ D_1G(0,0)x' &= D_1f(0,0)x' + D_2f(0,0)(D\xi'(0)P,0)x' \\ &= D_1f(0,0)x' + D_2f(0,0)(D\xi'(0)Px',0) \end{aligned}$$

and since $D\xi'(0) = 0$ by (3.18) and $D_2f(0,0)$ is linear we get after replacing x' by x

$$D_1G(0,0)x = D_1f(0,0)x = Lx.$$

Now $D_1G(0,0) = L$ is Fredholm with index I, and $N(D_1G(0,0)) = N(L)$, $R(D_1G(0,0)) = R(L)$.

Let $M = \{x = \xi_0 + \eta(\xi_0) : \xi_0 \in S_0\}$. Then M is a C^2–submanifold of $S_0 \oplus \widetilde{N(L)} \subset E$. The tangent space at 0 is $T_0M = S_0$. [To see this note that $T_{\tilde{a}}M = \{y \mid y = \frac{dx}{dt}|_{t=0}$ such that $x : [-1,1] \to M$ is of class C^1, $x(0) = \tilde{a}\}$ [cf. Definition 0.2.1]. Here $\tilde{a} = 0$. We have

$$\begin{aligned} \frac{dx(t)}{dt}|_{t=0} &= \frac{d}{dt}(\xi_0(t) + \eta(\xi_0(t))|_{t=0} \\ &= \frac{d\xi_0}{dt}(0) + D\eta(\xi_0(0))\frac{d\xi_0}{dt}(0) \\ &= \frac{d\xi_0}{dt}(0) + D\eta(0)\frac{d\xi_0}{dt}(0) \\ &= \frac{d\xi_0}{dt}(0) \end{aligned}$$

since $D\eta(0) = 0$ by (3.18). S_0 is a finite dimensional space so that $\frac{d\xi_0}{dt}(0) \in S_0$ and hence $T_0M \subset S_0$ and since T_0M and S_0 have the same dimension we get $T_0M = S_0$.]

Also

$$\begin{aligned}
G(x,0)|_M &= f(\xi_0+\eta(\xi_0),(\xi'(P(\xi_0+\eta(\xi_0)),0)) \\
&= f(\xi_0+\eta(\xi_0),(\xi'(P\xi_0),0)) \text{ since } P\eta(\xi_0)=0 \\
&= f(\xi_0+\eta(\xi_0),(\xi'(\xi_0),0)) \text{ since } P\xi_0=\xi_0\in R(P)=S_0 \\
&= \Phi(\xi_0,\eta(\xi_0),\xi'(\xi_0)) \text{ by definition of } \Phi \\
&= 0 \text{ by (3.15).}
\end{aligned}$$

Therefore the function G satisfies hypothesis (i) of Theorem 3.2.

From $Lp(\xi^0)=-D_2f(0,0)\xi^0$ we get $D_2f(0,0)\xi^0\in R(L)$. Let $\xi^0=(0,\xi''^0)$ where $\xi''^0\in\mathbb{R}^{N-d+I}$. As we saw above $\xi^0\notin W$ can always be written in this form so that $\xi''^0\in\mathbb{R}^{N-d+I}$.

We now get $D_2G(x,\xi''^0)\rho=D_2f(x,(\xi'(Px),\xi''^0))D_{(\xi''^0)}(\xi'(Px),\xi''^0)\rho$ where $\rho\in\mathbb{R}^{N-d+I}$ so that

$$D_2G(0,0)\rho=D_2f(0,0)\ (0,I_{\mathbb{R}^{N-d+I}})\rho.$$

Since ρ is arbitrary we replace it by ξ''^0 to get

$$D_2G(0,0)\xi''^0=D_2f(0,0)(0,\xi''^0)=D_2f(0,0)\xi^0=-Lx=-D_1G(0,0)x.$$

Therefore $D_1G(0,0)x=-D_2G(0,0)\xi''^0$ is the same as $Lx=-D_2f(0,0)\xi^0$ and has solution $p(\xi^0)$. Therefore the function G satisfies hypothesis (ii) of Theorem 3.2.

Let $w\in T_0M=S_0$ and consider

$$[D_1^2G(0,0)p(\xi^0)+D_1D_2G(0,0)\xi''^0]w\in R(D_1G(0,0))=R(L). \tag{3.19}$$

By definition $G(x,\xi'') = f(x,(\xi'(Px),\xi''))$

so that $D_1G(x,\xi'')p(\xi^0) = \{D_1f(x,(\xi'(Px),\xi'')) + D_2f(x,(\xi'(Px),\xi''))$
$(D\xi'(Px)P,0)\}\,p(\xi^0)$

and $\{D_1^2G(x,\xi'')p(\xi^0)\}w = \{[D_1^2f(x,(\xi'(Px),\xi''))+$
$D_1D_2f(x,(\xi'(Px),\xi''))(D\xi'(Px)P,0)+$
$[D_1D_2f(x,\xi'(Px),\xi'')+$
$D_2^2f(x,\xi'(Px),\xi''))(D\xi'(Px)P,0)]$
$(D\xi'(Px)P,0) + D_2f(x,(\xi'(Px),\xi''))$
$(D^2\xi'(Px)P^2,0)]p(\xi^0)\}\,w.$

Therefore $\{D_1^2G(0,0)p(\xi^0)\}w = \{D_1^2f(0,0)p(\xi^0)\}w$ since
$D\xi'(P0) = D\xi'(0) = 0$ by (3.18)

and $Pp(\xi^0) = 0$ since $p(\xi^0) \in \widetilde{N(L)}$.

And

$$D_2G(x,\xi'')\xi''^0 = D_2f(x,(\xi'(Px),\xi''))(0,I_{\mathbb{R}^{N-d+I}})\xi''^0$$
$$= D_2f(x,(\xi'(Px),\xi''))(0,\xi''^0)$$
$$\{D_1D_2G(x,\xi'')\xi''^0\}w = \{[D_1D_2f(x,(\xi'(Px),\xi''))$$
$$+ D_2^2f(x,(\xi'(Px),\xi''))(D\xi'(Px)P,\xi'')](0,\xi''^0)\}w$$
$$\{D_1D_2G(0,0)\xi''^0\}w = \{D_1D_2f(0,0)\xi^0\}w \text{ since } \xi^0 = (0,\xi''^0)$$
$$\text{and } D\xi'(0) = 0.$$

We now see that (3.19) is equivalent to

$$\{D_1^2f(0,0)p(\xi^0) + D_1D_2f(0,0)\xi^0\}\,w \in R(L).$$

As a result of hypothesis (iii) of this theorem $w = 0$ so that G satisfies hypothesis (iii) of Theorem 3.2

We have now shown that G satisfies hypothesis (i), (ii) and (iii) of Theorem 3.2.

The first part of the conclusion of Theorem 3.3 now follows from the conclusion of the first part of Theorem 3.2.

Note that $\xi : U \times J \to \mathbb{R}^N$ is defined by $\xi(\gamma,\mu) = (\xi'(Px(\gamma,\mu)), \mu\xi''^0)$ so that

$$\xi(\gamma,0) = (\xi'(P0),0) = (\xi'(0),0) = (0,0) = 0. \tag{3.20}$$

$$\begin{aligned} D_2\xi(\gamma,\mu) &= (D\xi'(Px(\gamma,\mu))PD_2x(\gamma,\mu), D\mu\xi''^0) \\ &= (D\xi'(Px(\gamma,\mu))PD_2x(\gamma,\mu), \xi''^0). \\ D_2\xi(\gamma,0) &= (D\xi'(Px(\gamma,0))PDx(\gamma,0), \xi''^0). \end{aligned}$$

Since $x(\gamma,0) = 0$ by the conclusion of Theorem 3.2 and $D\xi'(0) = 0$ by (3.18) we have $D\xi'(Px(\gamma,0) = 0$ so that

$$D_2\xi(\gamma,0) = (0,\xi''^0) = \xi^0. \tag{3.21}$$

We now show that $D_1f(x(\gamma,\mu),\xi(\gamma,\mu)) : E \to F$ in an invertible Fredholm operator of index I.

$D_1f(x(\gamma,\mu),\xi(\gamma,\mu))$ is Fredholm of index I by perturbation theory [cf. Theorem 0.1.19]. Recall $J = (-\mu_0,\mu_0)$. Let $g : (S_0 \oplus \widetilde{N(L)}) \times J \to F$ be defined by $g(x,\gamma,\mu) \triangleq D_1f(x(\gamma,\mu),\xi(\gamma,\mu))x$. Then $g(x,\gamma,0) = D_1f(x(\gamma,0),\xi(\gamma,0))x = D_1f(0,0)x = Lx$. The second equality follows from $x(\gamma,0) = 0$ by the conclusion of Theorem 3.2 and $\xi(\gamma,0) = 0$ by (3.20).

$D_1g(0,\gamma,0) = L$ so that $N(D_1g(0,\gamma,0)) = S_0$, $R(D_1g(0,\gamma,0)) = R(L)$. Therefore $D_1g(0,\gamma,0)$ is Fredholm; index of

$$\begin{aligned} D_1g(0,\gamma,0) &= \dim N(D_1g(0,\gamma,0) - \operatorname{codim} R(D_1g(0,\gamma,0)) \\ &= \dim N(D_1g(0,\gamma,0) - \operatorname{codim} R(L) \\ &= \dim S_0 - \operatorname{codim} R(L). \end{aligned}$$

Now $\dim S_0 = d - I$ and $\operatorname{codim} R(L) = \dim N(L) - I = d - I$.

Thus $i(D_1 g(0,\gamma,0)) = (d-I) - (d-I) = 0$ so that $D_1 g(0,\gamma,0)$ is Fredholm with index zero. Let $x \in S_0 = T_0 M = M$. Then $g(x,\gamma,0)|_{S_0} = L|_{S_0} = 0$ since $S_0 \subset N(L)$.

Therefore g satisfies hypothesis (i) of Theorem 3.2.

$$D_3 g(x,\gamma,\mu) = [D_1^2 f(x(\gamma,\mu),\xi(\gamma,\mu))D_2 x(\gamma,\mu) + D_1 D_2 f(x(\gamma,\mu),\xi(\gamma,\mu))D_2\xi(\gamma,\mu)]x.$$

$$\begin{aligned} D_3 g(0,\gamma,0) &= [D_1^2 f(0,0)D_2 x(\gamma,0) + D_1 D_2 f(0,0) D_2 \xi(\gamma,0)] * 0 \\ &= 0. \end{aligned}$$

$D_1 g(0,0,0)x = Lx = 0 = -D_3 g(0,0,0)$ has a solution $p \in S_0$ so fix $p \in S_0$ to be a solution of $D_1 g(0,0,0)x = -D_3 g(0,0,0)$.

Therefore g satisfies hypothesis (ii) of Theorem 3.2.

Observe that from $f(x(\gamma,\mu),\xi(x,\mu))$ we get

$$\begin{aligned} D_{(\mu)} f(x(\gamma,\mu),\xi(x,\mu)) &= D_1 f(x(\gamma,\mu),\xi(\gamma,\mu))D_2 x(\gamma,\mu) \\ &\quad + D_2 f(x(\gamma,\mu),\xi(\gamma,\mu))D_2\xi(\gamma,\mu). \end{aligned}$$

By the conclusion of Theorem 3.2, equations (3.20) and (3.21) we have

$$D_{(\mu)} f(x(\gamma,0),\xi(x,0)) = D_1 f(x(\gamma,0),\xi(\gamma,0)D_2 x(\gamma,0) + D_2 f(x(\gamma,0),\xi(\gamma,0))\xi^0.$$

so that $0 = LD_2 x(\gamma,0) + D_2 f(0,0)\xi^0$.

Therefore $D_2 x(\gamma,0)$ satisfies the equation $Lx = -D_2 f(0,0)\xi^0$ and hence so does $D_2 x(0,0)$.

If $w \in S_0$ and $D_1 g(x,\gamma,0)p = Lp = 0$ we have

$$\{D_1^2 g(x,\gamma,0)p\}w = \{Lp\}w = 0w = 0.$$

That is $\{D_1^2 g(0,0,0)p\}w = 0$.

$$D_1 g(x,\gamma,\mu) = D_1 f(x(\gamma,\mu),\xi(\gamma,\mu)) \text{ and}$$

$$\{D_1 D_3 g(x,\gamma,\mu)\}w = [D_1^2 f(x(\gamma,\mu),\xi(\gamma,\mu))D_2 x(\gamma,\mu) + D_1 D_2 f(x(\gamma,\mu),\xi(\gamma,\mu))D_2\xi(\gamma,\mu)]w.$$

$$\text{Therefore } \{D_1 D_3 g(0,0,0)\}w = [D_1^2 f(0,0)D_2 x(0,0) + D_1 D_2 f(0,0)\xi^0]w \text{ so that}$$

$$\{D_1^2 g(0,0,0)p + D_1 D_3 g(0,0,0)\} = \{D_1^2 f(0,0)D_2 x(0,0) + D_1 D_2 f(0,0)\xi^0\}w.$$

As a result of hypothesis (iii) of this theorem $w = 0$ so that g satisfies hypothesis (iii) of Theorem 3.2.

Therefore all the hypotheses of Theorem 3.2 are satisfied and hence the equation $D_1 f(x(\gamma,\mu),\xi(\gamma,\mu))x = 0$ has a unique solution $\tilde{x}(\gamma,\mu)$ in E. Since this equation is linear in x the unique solution must be $x = 0$ for $\mu \neq 0$, $\mu \in J$.

This completes the proof of Theorem 3.3. □

Chapter 4. EXISTENCE OF BOUNDED SOLUTIONS

In this chapter we will use the theorems in Chapter 3 to show the existence of bounded solutions for systems of the form of equation (2.1.1). We state two theorems ;

(i) for the case where $I = 0$,

(ii) and for the case where $I \geq 0$.

We will prove the theorem for the case where $I \geq 0$. As we will show below the results in this chapter make use of Lemma 2.3.1 in Chapter 2 and the three theorems in Chapter 3. We intend to show that under certain conditions on the perturbative term h, equation (2.1.1) has a bounded solution $z(t,\mu)$ near $\gamma(t)$, where $\gamma(t)$ is a bounded solution of the unperturbed system $\dot{z} = g(z)$. Again in this chapter $X \triangleq C_b^1(\mathbb{R},\mathbb{R}^n)$, $Y \triangleq C_b^0(\mathbb{R},\mathbb{R}^n)$, $V = R(P)$ and $W = N(Q)$ are the stable and unstable subspaces of (2.3.3), respectively. The first main result is the following:

Theorem 4.1. Let (H1) and (H2) in Chapter 2 hold and let $1 = \xi^0 \in \mathbb{R}$ in equation (2.1.1). Let $I = \dim V + \dim W - n$. Suppose, furthermore, that

(i) the dimension of the stable and unstable subspaces has sum n,

(ii) $\dot{\gamma}(t)$ is the unique (up to a scalar multiple) bounded solution of

$$\dot{x}(t) = A(t)x(t), \quad A(t) \triangleq Dg(\gamma(t)), \tag{4.1}$$

(iii)

$$\int_{-\infty}^{\infty} \psi^*(t)h(t,\gamma(t),0)dt = 0, \int_{-\infty}^{\infty} \psi^*(t)D_1h(t,\gamma(t),0)dt \neq 0, \tag{4.2}$$

where $\psi(t)$ is the unique (up to a scalar multiple) bounded solution of the system adjoint to (4.1). Then there exists positive constants Δ, σ, $0 < |\mu| < \sigma$ such that the perturbed system (2.1.1) has a unique solution $z(t,\mu)$ satisfying $|z(t,\mu) - \gamma(t)| < \Delta$ for all t. Moreover as $\mu \to 0$, $\sup_{t \in \mathbb{R}} |z(t,\mu) - \gamma(t)| = O(\mu)$.

We do not prove Theorem 4.1 since it is a special case of Theorem 4.2. However we have the following remarks.

Remarks. (1) Note that (4.1) is the same as (2.3.3).

(2) $\psi(t)$ is the unique bounded solution of the system adjoint to (4.1): To see this note that by Lemma 2.3.1 and hypothesis (i). $L = D_1 f(0,0)$ is Fredholm of index $I = 0$. By hypothesis (ii) $N(L)$ is spanned by $\dot{\gamma}(t)$ and $\dim N(L) = 1$. In the proof of Lemma 2.3.1 we showed that $I = \dim(V \cap W) - \dim(V^{\perp} \cap W^{\perp})$ so that $\dim(V^{\perp} \cap W^{\perp}) = 1$. Let $X(t)$ be the fundamental matrix of (4.1). Then $X^{-1*}(t)$ is the fundamental matrix of the system adjoint to (4.1). Choose $\eta \in V^{\perp} \cap W^{\perp}$. Then $\psi(t) = X^{-1*}(t)\eta$ is the unique bounded solution of the system adjoint to (4.1).

(3) Consider the condition

$$\begin{aligned} \int_{-\infty}^{\infty} \psi^*(t) h(t,\gamma(t),0)\,dt &= 0, \\ \int_{-\infty}^{\infty} \psi^*(t) D_1 h(t,\gamma(t)0)\,dt &\neq 0. \end{aligned} \tag{4.4}$$

For any $t_0 \in \mathbb{R}$, $\gamma(t - t_0)$ is also a solution of $\dot{z} = g(z)$ to which we can apply the Theorem. In this case 4.4 becomes

$$\begin{aligned} \Delta(t_0) &\triangleq \int_{-\infty}^{\infty} \psi^*(t - t_0) h(t,\gamma(t - t_0),0)\,dt = 0, \\ \Delta'(t_0) &\triangleq \int_{-\infty}^{\infty} \psi^*(t - t_0) D_1 h(t,\gamma(t - t_0),0)\,dt \neq 0. \end{aligned} \tag{4.5}$$

If we let $s = t - t_0$ then $t = s + t_0$ and realizing that s is a dummy variable, then letting $s = t - t_0$ is equivalent to replacing $t - t_0$ by t and hence replacing t by $t + t_0$. Therefore the condition (4.5) implies

$$\begin{aligned} \Delta(t_0) &= \int_{-\infty}^{\infty} \psi^*(t)h(t+t_0,\gamma(t),0)dt = 0, \\ \Delta'(t_0) &= \int_{-\infty}^{\infty} \psi^*(t)D_1h(t+t_0,\gamma(t),0)dt \neq 0. \end{aligned} \tag{4.6}$$

The condition then reduces to $\Delta(t_0) = 0$, $\Delta'(t_0) \neq 0$ where $\Delta(t_0)$, $\Delta'(t_0)$ can take any one of the forms in equation (4.4), (4.5) and (4.6). In the case of (4.4) we take $t_0 = 0$.

We now extend the result above to the case where the index of L is greater than or equal to zero. In the following theorem we will make use of Lemma 2.3.1 and Theorem 3.2 of the previous chapter.

Theorem 4.2. Let $d = \dim N(L)$ and let (H1) and (H2) hold true. Let $\mathbb{R}\dot{\gamma} = M_0 \subset X$, $S \subset X$, $\dim S = d-1$ be such that $N(L) = M_0 \oplus S$. Let $I = \dim V + \dim W - n$ as in Chapter 2, and assume that $N > d - I$. Let $M = \{x_\alpha \in X : \alpha \in \mathbb{R}$, where $x_\alpha(t) = \gamma(t+\alpha) - \gamma(t)$, for all $t \in \mathbb{R}\}$. Furthermore, suppose that one of the following conditions are satisfied:

(i) $I = d$

(ii) $I = d - 1$ and $\int_{-\infty}^{\infty} \psi^*(t)h(t,\gamma(t),0)dt = 0$, $\int_{-\infty}^{\infty} \psi^*(t)D_1h(t,\gamma(t),0)dt \neq 0$, $\psi(t)$ being the unique (up to a multiplicative constant) bounded solution of the system adjoint to (4.1). Then there exists a positive constant σ, $0 < |\mu| < \sigma$, and a compact neighbourhood U of 0 in S such that the perturbed system (2.1.1) has a bounded solution $z(t,\delta,\mu)$ satisfying $|z(t,\delta,\mu) - \gamma(t)| \in U$ for all t. Moreover as $\mu \to 0$, $\sup_{t\in\mathbb{R}} |z(t,\delta,\mu) - \gamma(t)| = O(\mu)$.

Remarks. (1) $T_0M = M_0 \subset N(L)$. This is seen from the following observation: $T_0M = \{y \mid y = \frac{d}{dt}x(t)|_{t=0}$, for some C^1–function $x : [-1,1] \to M$, $x(0) = a\}$ [cf. Definition 0.2.1].

Let $\alpha \in C^1[-1,1]$, $\alpha(0) = 0$, then in this case we have $s \mapsto z(s) \in M$, $z(0) = x(0)$, $s \in [-1,1]$ where $z(s) = x_{\alpha(s)}$. Therefore

$$\begin{aligned}\frac{d}{ds}z(s)|_{s=0} &= \frac{d}{ds}\{\gamma(\cdot + \alpha(s)) - \gamma(\cdot)\}|_{s=0} \\ &= \dot{\gamma}(\cdot + \alpha(s))\dot{\alpha}(s)|_{s=0} \\ &= \dot{\gamma}(\cdot + \alpha(0))\dot{\alpha}(0) \\ &= \dot{\gamma}(\cdot)\dot{\alpha}(0).\end{aligned}$$

Since $\alpha(s) \in \mathbb{R}$ for all $s \in [-1,1]$ it follows that $\dot{\alpha}(0) \in \mathbb{R}$. Hence $T_{x(0)}M = T_0M = \mathbb{R}\dot{\gamma}$. Thus $T_0M = M_0 \subset N(L)$.

From now on we shall write $N(L) = T_0M \oplus S$.

(2) Theorem 4.1 can be deduced from Theorem 4.2 by simply taking $S = \emptyset$. Note then that $N(L) = M_0 = T_0M$ in Theorem 4.1.

Proof. Introducing a new variable x through $z(t) = \gamma(t) + x(t)$ equation (2.1.1) can be written in the form $\dot{x}(t) = g(\gamma(t) + x(t)) - g(\gamma(t)) + \mu h(t, \gamma(t) + x(t), \mu\xi^0)$. Let

$$[f(x, \mu\xi^0)](t) \triangleq \dot{x}(t) - \{g(\gamma(t) + x(t)) - g(\gamma(t)) + \mu h(t, \gamma(t) + x(t), \mu\xi^0)\}$$

so that $f : X \times \mathbb{R}^N \to Y$ is C^2 in a neighbourhood of $(0,0) \in X \times \mathbb{R}^N$ and for all $y \in X$ we have

$$[D_1 f(x, \mu\xi^0)y](t) = \dot{y}(t) - \{Dg(\gamma(t) + x(t)) + \mu D_2 h(t, \gamma(t) + x(t), \mu\xi^0)\}y(t) \quad (4.7)$$

so that $(Ly)(t) = [D_1 f(0,0)y](t) = \dot{y}(t) - Dg(\gamma(t))y(t)$. The calculations are similar to those in section 2.3 of Chapter 2.

We will show that the function f defined above satisfies the hypothesis of Theorem 3.2.

Case $I = d$. With $I = d$ hypothesis (i) of Theorem 3.2 is not needed since we can take $M = \{0\}$ and $f(0,0) = 0$. We also do not need hypothesis (ii) since $\operatorname{codim} R(L) = I - \dim N(L) = d - d = 0$ so that $R(L) = Y$. Therefore $L : X \to Y$ is onto so that for each $y \in Y$ there is $x \in X$ such that $Lx = y$. Now take $y = -D_2 f(0,0)\xi^0$ so that (ii) is satisfied with $x = p(\xi^0)$. We also do not need hypothesis (iii) since $0 = w \in T_0 M$ so that $\{D_1^2 f(0,0)p(\xi^0) + D_1 D_2 f(0,0)\xi^0\}0 = 0 \in R(L)$.

Case $I = d - 1$. By hypothesis $\dim N(L^*) = 1$ since $\psi(t)$ is the unique (up to a scalar multiple) bounded solution of the system adjoint to (4.1).

$0 \in M$ since if $\alpha = 0$ we get $\gamma(t) - \gamma(t) = 0$ so that $x(0) = 0$. $x \in M$ if and only if $x = x_\alpha$, so that

$$\begin{aligned} f(x,0)(t) &= \dot\gamma(t+\alpha) - \dot\gamma(t) - [g(\gamma(t+\alpha) - \gamma(t) + \gamma(t)) - g(\gamma(t))] \\ &= [\dot\gamma(t+\alpha) - \dot\gamma(t) - [g(\gamma(t+\alpha)) - g(\gamma(t)] \\ &= [\dot\gamma(t+\alpha) - g(\gamma(t+\alpha))] - [\dot\gamma(t) - g(\gamma(t))]. \end{aligned}$$

Now $\dot z = g(z)$ is autonomous so that if $\gamma(t)$ is a solution of this equation, then so is $\gamma(t+\alpha)$ for any $\alpha \in \mathbb{R}$ and thus $f(x,0)(t) = 0 - 0 = 0$ for all $t \implies f(x,0) = 0 \in Y$. Therefore for any $x \in M$, $f(x,0) = 0$. Also $\dim M = 1$ since $M = \Phi(\mathbb{R})$ where $\Phi : \mathbb{R} \to X$ is defined by $\Phi(\alpha) = x_\alpha$. Since $\gamma(t)$ and $\gamma(t+\alpha)$ are continuous together with their derivatives. M is a C^1–manifold and since $I = d - 1$ and $\dim M = 1$ we get $\dim M = d - I$. These facts together with the fact that $f(x,0) = 0$ for $x \in M$ imply that f satisfies hypothesis (i) of Theorem 3.2.

$$[D_2 f(x,\mu\xi^0)(t)]\xi^0 = -\{h(t,\gamma(t) + x(t),\mu\xi^0) + \mu[D_3 h(t,\gamma(t) + x(t),\mu\xi^0)\xi^0]\}$$

so that $\quad D_2 f(0,0)\xi^0 = -h(t,\gamma(t),0).$

That is, $h(t,\gamma(t),0) = -D_2 f(0,0)\xi^0 \in R(L)$ since by hypothesis $\int_{-\infty}^{\infty} \psi^*(t)h(t,\gamma(t),0)dt = 0$ and Lemma 2.3.1 says this holds if and only if $h(t,\gamma(t),0) \in R(L)$.

Now $-D_2 f(0,0)\xi^0 \in R(L)$ implies that the equation $Lx = -D_2 f(0,0)\xi^0$ has a solution $p \in X$. That is, $p(t)$ is a bounded solution of

$$\dot{p} = Dg(\gamma(t))p + h(t,\gamma(t),0) \tag{4.8}$$

or $(Lp)(t) = h(t,\gamma(t),0) = -(D_2 f(0,0)\xi^0)(t)$.

Therefore f satisfies hypothesis (ii) of Theorem 3.2.

Differentiate (4.7) with respect to x and μ to get

$$\begin{aligned} D_1 D_2 f(x,\mu\xi^0)\xi^0 &= -\left\{D_2 h(t,\gamma(t)+x(t),\mu\xi^0) + \mu\left[D_2 D_3 h(t,\gamma(t)+x(t),\mu\xi^0)\right]\xi^0\right\} \\ &= -\{D_2 h(t,\gamma(t)+x(t),\mu\xi^0) + \mu D_2 D_3 h(t,\gamma(t)+x(t),\mu\xi^0)\xi^0\} \\ D_1 D_2 f(0,0)\xi^0 &= -D_2 h(t,\gamma(t),0) \\ D_1^2 f(x,\mu\xi^0) &= -\left\{D^2 g(\gamma(t)+x(t))D_{(x)}(\gamma(t)+x(t))\right. \\ &\quad \left. + \mu D_2 h(t,\gamma(t)+x(t),\mu\xi^0)D_{(x)}(\gamma(t)+x(t))\right\} \\ &\text{so that } D_1^2 f(0,0) = -D^2 g(\gamma(t)). \end{aligned}$$

Let

$$\begin{aligned} w(t) &= [\{D_1^2 f(0,0)p + D_1 D_2 f(0,0)\xi^0\}\dot{\gamma}](t) \\ &= -\{D^2 g(\gamma(t))\dot{\gamma}(t)\}p(t) + D_2 h(t,\gamma(t),0)\dot{\gamma}(t). \end{aligned}$$

Differentiate (4.8) with respect to t

$$\begin{aligned} \frac{d\dot{p}}{dt} &= Dg(\gamma(t))\dot{p}(t) + [\{D^2 g(\gamma(t))\dot{\gamma}(t)\}p(t) + D_2 h(t,\gamma(t),0)\dot{\gamma}(t) + D_1 h(t,\gamma(t),0)] \\ &= Dg(\gamma(t))\dot{p}(t) + [-w(t) + D_1 h(t,\gamma(t),0)] \end{aligned}$$

so that $\dot{p}(t)$ is a solution of

$$\dot{x}(t) = Dg(\gamma(t))x + [-w(t) + D_1h(t, \gamma(t), 0)]$$

and hence $-w(t) + D_1h(t, \gamma(t), 0) \in R(L)$ by Lemma 2.3.1.

By hypothesis of this theorem $\int_{-\infty}^{\infty} \psi^*(t)D_1h(t, \gamma(t), 0) \neq 0$ implying that $D_1h(t, \gamma(t), 0) \notin R(L)$ by Lemma 2.3.1 and hence $w(t) \notin R(L)$.

Note that in this case we have $y = \dot{\gamma}(t) \in T_0M$ and $\dot{\gamma}(t) \neq 0$. We have shown the negation of hypothesis (iii) of Theorem 3.2 since $y = \dot{\gamma}(t) \neq 0$ implies $w(t) = [\{D_1^2f(0,0)p + D_1D_2f(0,0)\xi^0\}\dot{\gamma}](t) \notin R(L)$.

Therefore f satisfies hypothesis (iii) of Theorem 3.2.

The conditions of Theorem 3.2 are satisfied. The conclusion of that theorem quarantees the existence of a unique solution $x(t, \delta, \mu) = z(t, \delta, \mu) - \gamma(t)$ in a neighbourhood of $(0,0) \in U \times J \subset S \times \mathbb{R}$, $\delta \in U$, $\mu \in J = (-\mu_0, \mu_0)$ of the equation $f(x, \mu\xi^0) = 0$. Now $x(t, \delta, \mu) \in X$ is bounded in U so that $z(t, \delta, \mu)$ is also bounded.

We know that $x(t, \delta, 0) = 0$ and since $x(t, \delta, \mu)$ is C^1 in μ, we have

$$\begin{aligned} x(t, \delta, \mu) &= x(t, \delta, 0) + \mu \int_0^1 (1 - s)D_3x(t, \delta, s\mu)ds \quad \text{[cf. Theorems 0.2.4 and 0.2.5]} \\ &= \mu \int_0^1 (1 - s)D_3x(t, \delta, s\mu)ds \\ &= O(\mu). \end{aligned}$$

Thus

$$\sup_{t \in \mathbb{R}} |x(t, \delta, \mu)| = O(\mu).$$

This completes the proof of Theorem 4.2. □

Chapter 5: APPLICATIONS

In this chapter we give two examples that are worked out in detail to show where and how the theory in the previous chapters can be applied.

Examples 1. (The case where $I = 0$)

In this example we will apply Theorem 4.1.

Consider the forced pendulum equation

$$\ddot{x} + \sin x = \mu \sin t. \tag{5.1}$$

This equation can be written as a system of two first order equations

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -\sin x + \mu \sin t. \end{aligned} \tag{5.2}$$

Equation (5.1) has the form

$$\begin{aligned} \dot{x} &= H_y(x,y) + \mu \tilde{H}_y(t,x,y,\mu) \\ \dot{y} &= -H_x(x,y) - \mu \tilde{H}_x(t,x,y,\mu) \end{aligned} \tag{5.3}$$

where $H(x,y) = \frac{1}{2}y^2 - \cos x$, $\tilde{H}(t,x,y,\mu) = -\mu x \sin t$.

If we let

$$z = \begin{pmatrix} x \\ y \end{pmatrix}, g(z) = \begin{pmatrix} g_1(z) \\ g_2(z) \end{pmatrix} = \begin{pmatrix} H_y(x,y) \\ -H_x(x,y) \end{pmatrix}$$
$$h(t,z,\mu) = \begin{pmatrix} h_1(t,z,\mu) \\ h_2(t,z,\mu) \end{pmatrix} = \begin{pmatrix} \tilde{H}_y(t,x,y,\mu) \\ -\tilde{H}_x(t,x,y,\mu) \end{pmatrix}$$

then we see that (5.3) has the form

$$\dot{z} = g(z) + \mu h(t,z,\mu). \tag{5.4}$$

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

Now (5.4) is an equation of the form (2.1.1) with $1 = \xi^0 \in \mathbb{R}$.

Consider the unperturbed equation

$$\ddot{x} + \sin x = 0, \quad x(0) = 0, \quad \dot{x}(0) = 2, \tag{5.5}$$

or equivalently,

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= -\sin x, \quad x(0) = 0,\ y(0) = 2. \end{aligned}$$

We wish to find a solution for (5.5), so multiply by $\dot{x}$ and integrate to obtain

$$\frac{1}{2}\dot{x}^2 - \cos x = 1. \tag{5.6}$$

From (5.6) we obtain

$$\begin{aligned} \dot{x} &= \pm\sqrt{2}\,\sqrt{\cos x + 1} \\ &= \pm\sqrt{2}\,\sqrt{2\cos^2\frac{x}{2}}. \end{aligned}$$

We can therefore consider $\dot{x} = 2\cos\frac{x}{2}$. Let $x = \pi - 4z$. With this change of variables it is not difficult to show that $x(t) = \pi - 4\tan^{-1}(e^{-t})$ is a solution of (5.5). Also

$$\begin{aligned} y(t) = \dot{x}(t) &= \frac{d}{dt}[\pi - 4\tan^{-1}(e^{-t})] \\ &= -4\left[\frac{1}{1+(e^{-t})^2}(-e^{-t})\right] \\ &= \frac{4}{e^t + e^{-t}} \\ &= 2\,\mathrm{sech}\,t. \end{aligned}$$

Therefore $\gamma(t) = \begin{pmatrix} \pi - 4\tan^{-1}(e^{-t}) \\ 2\,\mathrm{sech}\,t \end{pmatrix}$ is a solution of (5.5). To find fixed points let $\dot{x} = 0$ to get $y = 0$ and $\dot{y} = 0$ to get $\sin x = 0$ so that $x = \pm n\pi$, $n \in \mathbb{Z}$. Thus, the fixed points of system (5.5) are $\begin{pmatrix} -n\pi \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} n\pi \\ 0 \end{pmatrix}$. We shall consider only

$n = -1, 1$. Note that at $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ we have $Dg\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ so that the eigenvalues are $\lambda_1 = -i$ and $\lambda_2 = i$ and hence the point $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is a stable fixed point [cf. [23], p.77]. $Dg\begin{pmatrix} -\pi \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ so that the eigenvalues are $\lambda_1 = -1 < 0$ and $\lambda_2 = 1 > 0$ so that $\begin{pmatrix} -\pi \\ 0 \end{pmatrix}$ is a hyperbolic fixed point. Since there is one negative eigenvalue and one positive eigenvalue it follows also that at $\begin{pmatrix} -\pi \\ 0 \end{pmatrix}$ we have dim (stable subspace) = dim (unstable subspace) = 1. Also $Dg\begin{pmatrix} \pi \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ so that $\lambda_1 = -1 < 0$, $\lambda_2 = 1 > 0$ and hence the hyperbolic fixed point $\begin{pmatrix} \pi \\ 0 \end{pmatrix}$ has dim (stable subspace) = dim (unstable subspace) = 1.

Observe that

$$\lim_{t \to -\infty} \gamma(t) = \lim_{t \to -\infty} \begin{pmatrix} \pi - 4\tan^{-1}(e^{-t}) \\ 2\operatorname{sech} t \end{pmatrix} = \begin{pmatrix} -\pi \\ 0 \end{pmatrix},$$

[since as $t \to -\infty$, $-4\tan^{-1}(e^{-t}) \to 2\pi$] and that

$$\lim_{t \to \infty} \gamma(t) = \lim_{t \to \infty} \begin{pmatrix} \pi - 4\tan^{-1}(e^{-t}) \\ 2\operatorname{sech} t \end{pmatrix} = \begin{pmatrix} \pi \\ 0 \end{pmatrix},$$

[since as $t \to \infty$, $-4\tan^{-1}(e^{-t}) \to -4(0) = 0$]. And $\gamma(t)$ is a bounded solution of (5.5) since the inverse of a tangent function is bounded above by $\frac{\pi}{2}$ and below by $\frac{-\pi}{2}$ and the hyperbolic secant is bounded above by 1 and below by 0.

So the solution $\gamma(t)$ is a bounded solution of (5.5) and is a heteroclinic trajectory joining the two hyperbolic fixed points, viz. $\begin{pmatrix} -\pi \\ 0 \end{pmatrix}$ and $\begin{pmatrix} \pi \\ 0 \end{pmatrix}$. Let $x_0 = \begin{pmatrix} -\pi \\ 0 \end{pmatrix}$ and $x_1 = \begin{pmatrix} \pi \\ 0 \end{pmatrix}$.

Note that (5.5) has the form $\dot{z} = g(z)$ so that the variational system along $\gamma(t)$ is $\dot{y} = Dg(\gamma(t))y$. This variational system has at least one bounded solution, namely $\dot{\gamma}(t) = \begin{pmatrix} 2\operatorname{sech} t \\ -2\operatorname{sech} t \tanh t \end{pmatrix}$. Now observe that

$$\lim_{t \to -\infty} Dg(\gamma(t)) = Dg(x_0)$$

and

$$\lim_{t\to\infty} Dg(\gamma(t)) = Dg(x_1).$$

Since the stable and unstable subspaces for the variational equation corresponding to $\gamma(t)$ must have dimension one, $\dot{\gamma}(t)$ must be the unique (up to a scalar multiple) bounded solution of it. Any other solution of $\dot{y} = Dg(\gamma(t))y$ is unbounded. A bounded solution of the system adjoint to $\dot{y} = Dg(\gamma(t))y$ is

$$\psi(t) = \begin{pmatrix} 2\operatorname{sech} t \tanh t \\ 2\operatorname{sech} t \end{pmatrix}.$$

It follows from Lemma 2.3.1 (see also Remark 2 after Theorem 4.1) and its proof that (up to a scalar multiple) the bounded solution $\psi(t)$ is unique.

The conditions on the perturbative term $\mu h(t, x, y, \mu)$ in order to continue the heteroclinic orbit $\gamma(t)$ to $\gamma(t, \mu)$ are

$$\begin{aligned}\Delta(0) &= \int_{-\infty}^{\infty} (2\operatorname{sech} t \tanh t, 2\operatorname{sech} t) \begin{pmatrix} 0 \\ \sin t \end{pmatrix} dt \\ &= \int_{-\infty}^{\infty} 2\operatorname{sech} t \sin t \; dt = 0\end{aligned}$$

and

$$\begin{aligned}\Delta'(0) &= \int_{-\infty}^{\infty} (2\operatorname{sech} t \tanh t, 2\operatorname{sech} t) \begin{pmatrix} 0 \\ \cos t \end{pmatrix} dt \\ &= \int_{-\infty}^{\infty} 2\operatorname{sech} t \cos t \; dt \neq 0.\end{aligned}$$

We now show that equation (5.1) satisfies the hypothesis of Theorem 4.1 in Chapter 4.

(i) It satisfies (H1) with $g(z) = \begin{pmatrix} g_1(z) \\ g_2(z) \end{pmatrix} = \begin{pmatrix} y \\ -\sin x \end{pmatrix}$ where $z = \begin{pmatrix} x \\ y \end{pmatrix}$ is a function of t, that is, $z(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$. The sine function is a bounded continuous function and has derivatives of all orders that are bounded and continuous so that it is in

C_b^2. The same is true for y since $y = \dot{x} = 2\cos\frac{x}{2}$. With $h(t,z,\mu) = \begin{pmatrix} 0 \\ \mu \sin t \end{pmatrix}$ we have that h is C^∞ and also bounded since $\|\sin t\| \le 1$, so that it is in C_b^2.

(ii) $\dot{z} = g(z)$ has a solution $\gamma(t) = \begin{pmatrix} \pi - 4\tan^{-1}(e^{-t}) \\ 2\,\mathrm{sech}\, t \end{pmatrix}$ and as we showed above this solution is bounded. The variational system $\dot{y} = Dg(\gamma(t))y$, that is,

$$\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\cos(\pi - 4\tan^{-1}(e^{-t})) & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix},$$

has an exponential dichotomy on both $\mathbb{R}_+$ and $\mathbb{R}_-$. To see this note that

$$\dot{\gamma}(t) = \begin{pmatrix} 2\,\mathrm{sech}\, t \\ -2\tanh t\,\mathrm{sech}\, t \end{pmatrix}$$

is bounded as observed previously and as $t \to \pm\infty$, $\dot{\gamma}(t) \to \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Now since $\mathrm{sech}\, t = \dfrac{2}{e^t + e^{-t}}$ and $\tanh t = \dfrac{e^t - e^{-t}}{e^t + e^{-t}}$ it is obvious that $\dot{\gamma}(t) \to \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ exponentially as $|t| \to \infty$. As observed in the second paragraph of the previous page $\dot{\gamma}(t)$ is the unique bounded solution of the variational system $\dot{y} = Dg(\gamma(t))y$. Also $\dot{y} = Dg(\gamma(t))y$ can be written in the form $\dot{y} = (A(t) + B(t))y$ where $A(t) = A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $B(t) = \begin{pmatrix} 0 & 0 \\ \cos(4\tan^{-1}(e^{-t}) - 1 & 0 \end{pmatrix}$. By the remark following Propostion 2.2.1, $\dot{y} = A(t)y$ has an exponential dichotomy on both $\mathbb{R}_+$ and $\mathbb{R}_-$. Let $\|B(t)\| = \max_j \sum_{k=1}^{n} |\alpha_{kj}|$ and $\delta = 1$. Then applying Proposition 2.2.2 we have $\left|\cos(\tan^{-1}(e^{-t})) - 1\right| \le 1$ so that $\dot{y} = (A(t) + B(t))y$ has an exponential dichotomy on both $\mathbb{R}_+$ and $\mathbb{R}_-$. Therefore any other solution is unbounded and thus tends to $+\infty$ in absolute value as $|t| \to \infty$.

(iii) In this case we have a 2–dimensional system, that is, $n = 2$. We have shown above that dim (stable subspace) = dim (unstable subspace) = 1 for the variational equation along $\gamma(t)$ and hence the sum of the dimension of these subspaces is 2.

Therefore this example satisfies all the hypothesis of Theorem 4.1 and we also have

$$\int_{-\infty}^{\infty} \operatorname{sech} t \sin t \, dt = 0$$

since we have a product of an odd and even function in the integrand, and

$$\int_{-\infty}^{\infty} \operatorname{sech} t \cos t \, dt \neq 0$$

since from the table of integrals we have

$$\begin{aligned}\int_{-\infty}^{\infty} \operatorname{sech} t \cos t \, dt &= 2\int_{0}^{\infty} \operatorname{sech} t \cos t \, dt \\ &= 2\left(\frac{\pi}{2} \operatorname{sech} \frac{\pi}{2}\right) = \pi \operatorname{sech} \frac{\pi}{2} \\ &= \pi \left(\frac{2}{e^{\frac{\pi}{2}} + e^{-\frac{\pi}{2}}}\right) \neq 0.\end{aligned}$$

Thus we conclude that there exists constants $K, \sigma > 0$ such that for $0 < |\mu| < \sigma$ equation (5.1) has a unique solution $\gamma(t,\mu)$ satisfying $|\gamma(t,\mu) - \gamma(t)| < K$ for all t. Since $\int_{-\infty}^{\infty} \operatorname{sech} t \sin t \, dt = 0$ and $\int_{-\infty}^{\infty} \operatorname{sech} t \cos t \, dt \neq 0$ the condition on the perturbative term h in order to continue the heteroclinic orbit $\gamma(t)$ to $\gamma(t,\mu)$ is satisfied. Moreover

$$\sup_{t\in\mathbb{R}} |\gamma(t,\mu) - \gamma(t)| = O(\mu).$$

As noted above $Dg(x_0)$ and $Dg(x_1)$ have the same number of eigenvalues with negative real parts and $\lim_{t\to-\infty} \gamma(t) = x_0$ and $\lim_{t\to\infty} \gamma(t) = x_1$. Also the perturbative term h has period 2π in t. The conditions therefore of Corollary 4.4 in [20] are satisfied and $h(t, z, \mu\xi^0) = \mu \sin t$ is 2π–periodic so (5.1) has unique 2π–periodic solutions $\gamma_0(t,\mu)$, $\gamma_1(t,\mu)$ such that for $i = 0, 1$

$$\sup_{t\in\mathbb{R}} |\gamma_i(t,\mu) - \gamma(t)| \to 0 \text{ as } \mu \to 0.$$

Moreover, $\gamma(t,\mu)$ is such that

$$|\gamma(t,\mu) - \gamma_0(t,\mu)| \to 0 \text{ as } t \to -\infty$$
$$|\gamma(t,\mu) - \gamma_1(t,\mu)| \to 0 \text{ as } t \to \infty.$$

Therefore $\gamma(t,\mu)$ is a general heteroclinic solution [cf. end of section 1.1].

Example 2. (The case where $I > 0$).

In this example we will apply Theorem 4.2.

Consider $\dot{z} = g(z)$ where $z = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $g(z) = \begin{pmatrix} g_1(z) \\ g_2(z) \\ g_3(z) \end{pmatrix} = \begin{pmatrix} x_1(x_1-1) \\ x_2(2x_1-1) \\ x_3(1-2x_1) \end{pmatrix}$ and x_i, $i = 1,2,3$ are functions of t. That is, consider the system

$$\begin{aligned} \dot{x}_1 &= x_1(x_1-1) \\ \dot{x}_2 &= x_2(2x_1-1) \\ \dot{x}_3 &= x_3(1-2x_1). \end{aligned} \tag{5.7}$$

Look for the fixed points of this system. So let $\dot{x}_1 = 0$, $\dot{x}_2 = 0$, $\dot{x}_3 = 0$, that is,

$$\begin{aligned} x_1(x_1-1) &= 0 \\ x_2(2x_1-1) &= 0 \\ x_3(1-2x_1) &= 0. \end{aligned} \tag{5.8}$$

From the first equation in (5.8) we have $x_1 = 0$ or 1, from the second equation we have $x_2 = 0$ and from the third equation we have $x_3 = 0$ so that the fixed points are $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

We then find a trajectory joining these points. $\dot{x}_1 = x_1(x_1-1)$ implies $\int \frac{dx_1}{x_1(x_1-1)} = \int dt$ and by the method of partial fractions we obtain $x_1 = \frac{1}{1-A_0e^t}$ and $A_0 < 0$

in order to have global solutions, otherwise for $A_0 \geq 1$, x_1 approaches infinity in finite time $t_0 = \ln \frac{1}{A_0}$. So choose $A_0 = -\alpha$, $\alpha > 0$.

Therefore $x_1 = \dfrac{1}{1+\alpha e^t}$ and

$$\lim_{t\to-\infty} x_1 = \lim_{t\to-\infty} \frac{1}{1+\alpha e^t} = 1.$$

Also

$$\lim_{t\to\infty} x_1 = \lim_{t\to\infty} \frac{1}{1+\alpha e^t} = 0.$$

We shall consider only one curve joining the fixed points. So we choose $\alpha = 1$. Thus $x_1 = \dfrac{1}{1+e^t} = (1+e^t)^{-1}$. In order to have such a curve, clearly $x_2 = x_3 = 0$. Thus the heteroclinic trajectory joining the two fixed points is $\gamma(t) = \begin{pmatrix} (1+e^t)^{-1} \\ 0 \\ 0 \end{pmatrix}$. It is easy to see that

$$\lim_{t\to-\infty} \gamma(t) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \lim_{t\to\infty} \gamma(t) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

as required. For $z = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ we have

$$Dg(z) = \begin{pmatrix} 2x_1 - 1 & 0 & 0 \\ 2x_2 & 2x_1 - 1 & 0 \\ -2x_3 & 0 & 1 - 2x_1 \end{pmatrix}$$

so that

$$Dg(\gamma(t)) = \begin{pmatrix} 2(1+e^t)^{-1} - 1 & 0 & 0 \\ 0 & 2(1+e^t)^{-1} - 1 & 0 \\ 0 & 0 & 1 - 2(1+e^t)^{-1} \end{pmatrix}. \tag{5.9}$$

The variational equation along $\gamma(t)$ then is

$$\begin{pmatrix} \dot{\xi}_1 \\ \dot{\xi}_2 \\ \dot{\xi}_3 \end{pmatrix} = \begin{pmatrix} \frac{1-e^t}{1+e^t} & 0 & 0 \\ 0 & \frac{1-e^t}{1+e^t} & 0 \\ 0 & 0 & -\left(\frac{1-e^t}{1+e^t}\right) \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix}. \tag{5.10}$$

That is $\dot{x} = A(t)x$, $A(t) \triangleq Dg(\gamma(t))$, $x = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix}$. $\dot{\xi}_1 = \dfrac{1-e^t}{1+e^t}\xi_1$ so that $\xi_1 = \dfrac{c_1 e^t}{(1+e^t)^2}$. Similarly $\xi_2 = \dfrac{c_2 e^t}{(1+e^t)^2}$ and $\xi_3 = \dfrac{(1+e^t)^2}{c_3 e^t}$ or $\xi_3 = 0$.

Therefore the fundamental matrix $X(t)$ satisfying $X(0) = I_{\mathbb{R}^3}$ is

$$X(t) = \begin{pmatrix} 4e^t(1+e^t)^{-2} & 0 & 0 \\ 0 & 4e^t(1+e^t)^{-2} & 0 \\ 0 & 0 & \frac{e^{-t}(1+e^t)^2}{4} \end{pmatrix} \tag{5.11}$$
$$\triangleq (X_1(t),\ X_2(t),\ X_3(t)).$$

Note that

$$\lim_{t\to-\infty} X_1(t) = \lim_{t\to-\infty} \begin{pmatrix} \frac{4e^t}{(1+e^t)^2} \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$
$$\lim_{t\to\infty} X_1(t) = \lim_{t\to\infty} \begin{pmatrix} \frac{4e^t}{(1+e^t)^2} \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

The same holds for $X_2(t)$.

$$\lim_{t\to-\infty} X_3(t) = \lim_{t\to-\infty} \begin{pmatrix} 0 \\ 0 \\ \frac{(1+e^t)^2}{4e^t} \end{pmatrix} = \infty$$
$$\lim_{t\to\infty} X_3(t) = \lim_{t\to\infty} \begin{pmatrix} 0 \\ 0 \\ \frac{(1+e^t)^2}{4e^t} \end{pmatrix} = \infty.$$

Since the bounded solutions $X_1(t)$ and $X_2(t)$ tend to 0 as t tends to $\pm\infty$ and the unbounded solution $X_3(t)$ tends to $+\infty$ in absolute value as t tends to $\pm\infty$, we have an exponential dichotomy for the variational equation on both $\mathbb{R}_+$ and $\mathbb{R}_-$. Let V be the stable subspace of (5.10) and W be the unstable subspace of (5.10). Then from the observation above we see that $\dim V = \dim W = 2$, since $X_1(t)$ and $X_2(t)$ tend to zero as $t \to \infty$ and so must have initial values in the stable subspace V and,

$X_1(t)$ and $X_2(t)$ also tend to zero as $t \to -\infty$ and so must have initial values in the unstable subspace W by the discussion in section 2.3 of Chapter 2. Also we have $N(L) = V \cap W$ so that $\dim N(L) = \dim(V \cap W)$. Now

$$\begin{aligned} I &= \dim V + \dim W - n \\ &= 2 + 2 - 3 \\ &= 1 \\ &> 0. \end{aligned}$$

Since $d = \dim N(L) = \dim(V \cap W) = 2$ we have $I = d - 1$.

We now digress from the application of Theorem 4.2 to show that the intersection of the stable and unstable manifolds of (5.7) in a neighbourhood of $\gamma(0)$ must have a particular form.

Let $\tilde{V}$ and $\tilde{W}$ be the sable and unstable manifolds of (5.7). We now show that in a neighbourhood of $\gamma(0) = (\frac{1}{2}, 0, 0)^*$ we have $\tilde{V} \cap \tilde{W} = \{(a, b, 0)^* : 0 < a < 1\}$. This is done by showing that bounded solutions with initial values in $\tilde{V} \cap \tilde{W}$ and such that they tend to zero as $|t| \to \infty$ are such that $0 < a < 1$.

Let $\gamma(t, a, b) = (\gamma_1(t, a, b), \gamma_2(t, a, b), \gamma_3(t, a, b))$. We wish $\gamma(t, a, b)$ to be a bounded solution of (5.7) satisfying $\gamma(0, a, b) = (a, b, 0)$, that is, a solution with initial condition in $\tilde{V} \cap \tilde{W}$. Then recall that $\gamma_1(t, a, b) = \dfrac{1}{1 - A_0 e^t}$ so that $\gamma_1(0, a, b) = \dfrac{1}{1 - A_0} = a$. That is, $A_0 = \dfrac{a-1}{a}$. Therefore $\gamma_1(t, a, b) = \dfrac{a}{a + (1-a)e^t}$. Also in (5.7) we have $\dot{x}_2 = x_2(2x_1 - 1)$ so that $\dot{x}_2 = \dfrac{1 + A_0 e^t}{1 - A_0 e^t} x_2$. Solving this equation yields $x_2 = \dfrac{c_2 e^t}{(1 - A_0 e^t)^2}$ so that one gets

$$\gamma_2(t, a, b) = \frac{c_2 e^t}{(1 - (\frac{a-1}{a} e^t)^2}$$

$$\gamma_2(0, a, b) = \frac{c_2}{(1 - \frac{a-1}{a})^2} \implies c_2 = \frac{b}{a^2}.$$

Thus $\gamma_2(t,a,b) = be^t[a+(1-a)e^t]^{-2}$. Also $\dot{x}_3 = x_3(1-2x_1)$ so that

$$x_3 = \frac{(1-A_0e^t)^2}{c_3e^t} = \frac{[a+(1-a)e^t]^2}{a^2c_3e^t}$$
$$= \gamma_3(t,a,b)$$

$\gamma(0,a,b) = 0 = \dfrac{1}{a^2c_3}$ and this is not possible and so we must have $\gamma(t,a,b) \equiv 0$ by the previous solution. Hence

$$\gamma_1(t,a,b) = a[a+(1-a)e^t]^{-1}$$
$$\gamma_2(t,a,b) = be^t[a+(1-a)e^t]^{-2}$$
$$\gamma_3(t,a,b) = 0.$$

Note that in order to have bounded global solutions it is necessary that $0 < a < 1$. If $a > 1$, γ_1 and γ_2 get arbitrarily large in finite time, that is, we can always find $t_0 \in \mathbb{R}$ such that for $t = t_0$ the denominator is zero. If $a = 1$, we get a global unbounded solution since $\gamma_1 = 1$, $\gamma_2 = be^t$, $\gamma_3 = 0$. If $a = 0$, we get a global unbounded solution since $\gamma_1 = 0$, $\gamma_2 = be^{-t}$, $\gamma_3 = 0$. If $a < 0$, γ_1 and γ_2 get arbirarily large in finite time, that is, we can find $t_0 \in \mathbb{R}$ such that for $t = t_0$ the denominator is zero.

Therefore in order to have bounded global solutions we must have $0 < a < 1$ since $\gamma(t,a,b) \to 0$ as $|t| \to \infty$ for $0 < a < 1$. Also if $0 < a < 1$, $|\gamma(t,a,b) - \gamma(t)| \to 0$ as $|t| \to \infty$.

The above argument therefore shows that $\tilde{V} \cap \tilde{W} = \{(a,b,0)^* : 0 < a < 1\}$.

We now turn back to the application of Theorem 4.2.

The system adjoint to (5.10) is $\dot{x} = -A^*(t)x$, that is,

$$\begin{pmatrix} \dot{\xi}_1 \\ \dot{\xi}_2 \\ \dot{\xi}_3 \end{pmatrix} = -\begin{pmatrix} \frac{1-e^t}{1+e^t} & 0 & 0 \\ 0 & \frac{1-e^t}{1+e^t} & 0 \\ 0 & 0 & -(\frac{1-e^t}{1+e^t}) \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix}$$

and as we saw above a matrix solution is

$$\psi(t) \triangleq (\psi_1(t), \psi_2(t), \psi_3(t))$$
$$= \begin{pmatrix} \frac{(1+e^t)^2}{k_1 e^t} & 0 & 0 \\ 0 & \frac{(1+e^t)^2}{k_2 e^t} & 0 \\ 0 & 0 & -(\frac{k_3 e^t}{(1+e^t)^2}) \end{pmatrix} \quad \text{or}$$
$$\tilde{\psi}(t) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -(\frac{k_3 e^t}{(1+e^t)^2}) \end{pmatrix}.$$

Also as noted above the only bounded solution (up to a multiplicative constant) is

$$\psi_3(t) = (0, 0, \frac{e^t}{(1+e^t)^2},)^*.$$

Now all the conditions of Theorem 4.2 are satisfied and the conditions on the perturbative term $\mu h(t, z, \mu\xi^0)$ in order to have a bounded solution are

$$\int_{-\infty}^{\infty} \psi^*(t) h(t, \gamma(t), 0) dt = 0,$$
$$\int_{-\infty}^{\infty} \psi^*(t) D_1 h(t, \gamma(t), 0) dt \neq 0.$$

That is,

$$\int_{-\infty}^{\infty} (0, 0, e^t(1+e^t)^{-2}) \begin{pmatrix} h_1(t, \gamma(t), 0) \\ h_2(t, \gamma(t), 0) \\ h_3(t, \gamma(t), 0) \end{pmatrix} dt = 0$$

or $$\int_{-\infty}^{\infty} e^t(1+e^t)^{-2} h_3(t, \gamma(t), 0) dt = 0$$

and similarly $$\int_{-\infty}^{\infty} e^t(1+e^t)^{-2} D_1 h_3(t, \gamma(t), 0) dt \neq 0.$$

The conclusion of Theorem 4.2 then says for $\mu \neq 0$, $\mu \in (-\mu_0, \mu_0) \subset \mathbb{R}$, $\delta \in V \subset \mathbb{R}^I$ (in this case $I = 1$) where V is a compact neighbourhood of $0 \in \mathbb{R}^I$, that is, $\delta \in V \subset \mathbb{R}$, we get bounded solutions $\gamma(t, \delta, \mu)$ of the system

$$\begin{aligned}
\dot{x}_1 &= x_1(x_1 - 1) + \mu h_1(t, z, \mu\xi^0) \\
\dot{x}_2 &= x_2(2x_1 - 1) + \mu h_2(t, z, \mu\xi^0) \\
\dot{x}_3 &= x_3(1 - 2x_1) + \mu h_3(t, z, \mu\xi^0).
\end{aligned}$$

Now one can take $\mu h(t, z, \mu\xi^0) = \begin{pmatrix} 0 \\ 0 \\ \mu \dfrac{1-e^t}{1+e^t} \end{pmatrix}$ so that the system above is

$$\begin{aligned}
\dot{x}_1 &= x_1(x_1 - 1) \\
\dot{x}_2 &= x_2(2x_1 - 1) \\
\dot{x}_3 &= x_3(1 - 2x_1) + \mu \frac{1-e^t}{1+e^t}.
\end{aligned}$$

Using integration by parts the condition above then reads

$$\int_{-\infty}^{\infty} \frac{e^t}{(1+e^t)^2} \cdot \frac{1-e^t}{1+e^t} dt = 0$$

and

$$\int_{-\infty}^{\infty} \frac{e^t}{(1+e^t)^2} \cdot \frac{-2e^t}{(1+e^t)^2} dt = -\frac{1}{6} \neq 0.$$

References

1. V.I. Arnold, "Ordinary Differential Equations", Translated from Russian and edited by R. Silverman, The MIT Press, Massechusetts, 1973.

2. D.K. Arrowsmith and C.M. Place, "An Introduction to Dynamical Systems", Cambridge University Press, Cambridge, 1990.

3. G. Bachman and L. Narici, "Functional Analysis", Academic Press, New York, 1966.

4. F. Battelli and C. Lazzari, Exponential dichotomies, heteroclinic orbits, and Melnikov functions, J. Differential Equations **86**, 342–366, 1990.

5. P. Blanchard and E. Brüning, "Variational Methods in Mathematical Physics", Springer–Verlag, 1992.

6. A.L. Brown and A. Page, "Elements of Functional Analysis", Von Nostrand Reinhold Company, London, 1970.

7. S.N. Chow, J.K. Hale, and J. Mallet–Paret, An example of bifurcation to homoclinic orbits, J. Differential Equations **37**, 351–373, 1980.

8. S.N. Chow, and J.K. Hale, "Methods of Bifurcation Theory", Springer–Verlag, New York, 1982.

9. W.A. Coppel, "Dichotomies in Stability Theory", Lecture Notes in Mathematics, Vol.629, Springer–Verlag, Berlin, 1978.

10. W.A. Coppel, "Stability and Asymptotic Behaviour of Differential Equations", D.C. Heath, Boston, 1965.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

11. J. Cronnin, "Differential Equations: Introduction and Qualilative Theory", Marcel Dekker Inc., New York, 1980.

12. K. Deimling, "Nonlinear Functional Analysis", Springer–Verlag, Heidelburg, 1985.

13. J. Gruendler, "The existance of homoclinic orbits and the method of Melnikov for systems in $\mathbb{R}^n$", SIAM J. Math. Anal. **16** (1985), 907–931.

14. J. Guckenheimer and P. Homes, "Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields", Springer–Verlag, New York, 1983.

15. J.K. Hale, "Ordinary Differential Equations", Wiley, New York, 1969.

16. J.K. Hale, Introduction to dynamic bifurcation, in "Bifurcation Theory and Applications", Lecture Notes in Mathematics, Vol. 1057, 106–151, Springer–Verlag, Berlin, 1984.

17. E. Kreyszig, "Introductory Functional Analysis with Applications", Wiley, New York, 1976.

18. S. Lang, "Real Analysis", Addison–Wesley, 1969.

19. R.D. Milne, "Applied Functional Analysis: An Introductory Treatment", Pittmann Publishing Inc., Massechusettes, 1980.

20. K.J. Palmer, Exponential dichtomies and transversal homoclinic points, J. Differential Equations **55**, 225–256, 1984.

21. K.J. Palmer, Transversal heteroclinic orbits and Cherry's example of a non–integrable hamiltonian system, J. Differential Equations **65**, (1986), 321–360.

22. L. Perko, "Differential Equations and Dynamical Systems", Springer–Verlag, New York, 1991.

23. D.A. Sánchez, "Ordinary Differential Equations and Stability Theory: An Introduction", Dover Publications, New York, 1979.

24. M. Schechter, "Principles of Functional Analysis", Academic Press, New York, 1971.

25. W. Wiggins, "Introduction to Applied Nonlinear Dynamical Systems and Chaos", Springer–Verlag, New York, 1990.

26. S. Wiggins, "Global Bifurcations and Chaos", Springer–Verlag, New York, 1988.