ANALYSIS OF SOME NONLINEAR EIGENVALUE PROBLEMS

by

Alan M. Selby

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science.

Department of Mathematics
McGill University
Montreal, Canada

January 1979.

Analysis of Some Nonlinear Eigenvalue Problems

by

Alan M. Selby

ABSTRACT

Multi-linearity and analyticity of two different operator-valued composition maps $(f,g) \to f \circ g$ imply respectively basic results in linear perturbation theory for pairs of linear operators (F,G) and the existence of stable and unstable manifolds passing through hyperbolic fixed points of analytic functions.

In a study of the branching of solutions to equations $E(x_1,\dots,x_n) = 0$ involving n Banach space variables x_j, n-dimensional Newton-Puiseux indicial sets are introduced in an extension of concepts (or least the terminology) associated with Newton-Puiseux diagrams. These indicial sets are employed to select a priori representations of solution branches. Substitution of these a priori representations into $E(x) \equiv E(x_1, \dots ,x_n)$ leads to factorizations. The vanishing of a non-trivial factor in these factorizations implicitly determines the unknown variables in the a priori representation and thus yields actual solution branches of $E(x) = 0$.

A generalization of Taylor's formula justifies the above factorizations for some differentiable functions $E(x)$. One such factorization is utilized in establishing the existence of the afore-mentionned stable and unstable manifolds.

Analyse de Quelques Problemes Non-Linéares

par

Alan M. Selby

RÉSUMÉ

Les propriétés de multilinéarité et d'analycité pour deux operations de composition sur des functions linéaires et des fonctions analytiques entraînent respectivement des résults élémentaires sur la theorie des perturbations linéaires pour une paire de fonctions linéaires (F,G) et l'existence de variétés stables et instables passant par une point fixe hyperbolique d'une fonction analytique.

Se trouve également exposée, une étude sur les branches de solution d'équations de la forme $E(x_1,\dots,x_n) = 0$ où $x_1,\dots,x_n$ prennent des valeurs dans des espaces de Banach. Des ensembles d'indices n-dimensionnels sont introduits comme une extension n-dimensionelle des diagrammes de Newton-Puiseux, a fin de choisir des representations a priori de branches de solutions de l'équation $E(x) = 0$. Le replacement de ces solutions représentées, dans l'équation initiale entraîne des factorisations. L'annulation d'un facteur non-trivial dans une telle factorisation, decrit implicitement les inconnues de la representation, et détermine de véritables solution de l'équation $E(x) = 0$.

Les factorisations dont il a été question plus haut, sont justifiée par une extension de la formule de Taylor. Un telle factorisation fut utilisée pour établir l'existence de variétes stables et instables mentionnée au premier paragraphe.

ACKNOWLEDGEMENTS.

This dissertation was written under the supervision of Professor W. F. Langford of the Department of Mathematics at McGill University.

It was typed on a typewriter borrowed from Ms Hélène LeBlanc with special typing elements borrowed from Miss Pauline Cassin and Miss Hildegarde Schroeder.

I am grateful to all of the above for their assistance and cooperation.

Foreword

The three chapters in this dissertation have arisen from my attempts to understand bifurcation theory. They have evolved from a study of three papers (1), (4) and (5) of M. G. CRANDALL and P.H. RABINOWITZ on Hopf bifurcation, bifurcation due to simple eigenvalues and linear perturbation theory. For myself, but not necessarily for the authors, the notable feature of these papers is their employment of the implicit function theorem in Banach spaces, sometimes in conjunction with a small amplitude or perturbation parameter s and an associated factorization, to derive linear and nonlinear perturbation results. Their factorization involves the division of an equation (or a function) by a power of the perturbation parameter s to obtain a new equation (or function) to which the application of the implicit function theorem is feasible at $s = 0$. A variation of this factorization method of Crandall and Rabinowitz is employed in chapter 3. Other factorization methods have been previously utilized in the work of J. DIEUDONNE (8), R.G. BARTLE (2), L. M. GRAVES (12) and D. SATHER (22) in their studies of the branching of solutions to equations. An analysis of such factorization methods is given in chapter 2.

The linear perturbation theory in chapter 1 depends on the Fréchet differentiation of the multi-linear composition maps in a (nonlinear) identity $F \circ h = G \circ h \circ c$. Here F, G, h and c are bounded linear operators.

Linear perturbation results then follow from the implicit function theorem. The perturbation of Floquet exponents and Floquet representations of disturbed linear periodic systems of ordinary differential equations is presented as an example and possible motivation for this theory. Standard eigenvalue perturbation formulas follow for single and multiple eigenvalues from differentiation of the identity $F \circ h = G \circ h \circ c$ when (h,c) are analytic operator-valued functions of (F,G). In chapter 3 a similar identity appears in which F,G,h and c are analytic rather than linear maps.

In chapter 2, factorizations and n-dimensional Newton-Puiseux indicial sets are introduced to seek sufficient conditions for families or curves of solutions $x(s)$ parameterized by a small parameter s (and possibly other variables) to issue at $s = 0$ from a branch-point of an equation $E(x) = 0$. The operator E in this equation is a Fréchet differentiable function of $n \geq 2$ Banach space variables $x = (x_1, \ldots x_n)$. When $n = 2$ and x is a 2-dimensional variable with real or complex components (x_1, x_2), the classical Newton-Puiseux diagrams or polygons are the boundaries of the convex hulls of the above Newton-Puiseux-indicial sets. For analytic maps $E(x)$, factorization of the equation $E(x) = 0$ after the substitution of an a priori representation $x = (s^{p_j} z_j)_{1 \leq j \leq n}$ imposes a sufficient geometric compatibility condition on the vector $p = (p_1, \ldots p_n)$ formed from the exponents p_j. This condition is the requirement that the vector p be a normal vector in $\mathbb{R}^n$ to a supporting hyperplane of the Newton-Puiseux-indicial set of E at $x = 0$. This permits a factorization

$$F_{p,r}(s,z) = s^{-r}E(s^{p_1}z_1,\dots s^{p_n}z_n)$$

in which the function $F_{p,r}(s,z)$ is not identically zero for $s = 0$, is analytic, and has the property that analytic extensions $z = (z_j(s))_{1\le j\le n}$ of roots $z^0 = z(0)$ of $F_{p,r}(0,z^0) = 0$, obtained with the implicit function theorem, yield via the a priori representation analytic branches

$$x(s) = (s^{p_1}z_1(s),\dots s^{p_n}z_n(s))$$

of solutions to $E(x) = 0$. These issue at a branch-point at the origin $x = 0$ when all the exponents $p_1,\dots p_n$ are positive.

For differentiable and not necessarily analytic functions, similar results follow from a factorization technique based on an extension of Taylor's Formula. In this extension, there is a remainder term containing an integrand involving a sum (in multi-index notation) of Fréchet derivatives $D^{\alpha}E(x) = D_{x_1}^{\alpha_1}D_{x_2}^{\alpha_2}\dots D_{x_n}^{\alpha_n}E(x)$ of E, with different orders $|\alpha| = \Sigma_{j=1}^{n}\alpha_j$.

The possible novel feature of chapter 2, in addition to this extension of Taylor's formula, is the introduction of the terminology Newton-Puiseux indicial-set to facilitate the description of factorization arguments involving n-Banach space arguments (In M. M. VAINBERG and V. A. TRENOGIN (27) there are similar arguments involving hyperplanes and variables $x = (x_1,\dots x_n)$ in n-dimensional Euclidean space.)

In chapter 3, a variation of the afore-mentionned factorization method of CRANDALL and RABINOWITZ, and a complex-analytic perturbation theory show the existence of stable and unstable analytic manifolds passing through hyperbolic fixed points of real and complex analytic functions. The existence follows from an application of the implicit function theorem and from the joint analyticity of certain composition maps $(f,g) \to f \circ g$ appearing in a conjugacy relation $F \circ sh = G \circ sh \circ c$. In this conjugacy relation s is a small perturbation parameter and the operators F,G,h and c are elements of Banach spaces of analytic functions. The analyticity of the composition maps in the conjugacy relation permit the preceding application of the implicit function theorem to a factorization of the conjugacy relation to yield the existence of solutions $(h,c) \neq (0,0)$ to this relation for some $s \neq 0$. This in turn for special choices of the functions (F,G) yields the existence of stable and unstable manifolds for analytic functions with hyperbolic fixed points. The complex analyticity of the composition maps in the above existence argument follows from uniform estimates provided by Cauchy's formula for analytic functions. Further details can be found in chapter 3.

Despite the links indicated above, each chapter is independent of the others. Each chapter has its own introduction which contains a chapter summary and background information. The latter is intended to supplement the results stated in the rest of its chapter. Each chapter also has its own bibliography. Finally all the chapters are exercises in Calculus for Banach spaces and chapter 3 in particular employs analytic operators defined on Banach (function) spaces.

References.

1 Crandall M. G. and P. H. Rabinowitz. Bifurcation, Perturbation of Simple Eigenvalues and Linearized Stability. Arch., Rat., Mech., Anal., 52, 1973, 161-180.

For the other references see the corresponding numbers in the Bibliography to Chapter 2, page 58.

CONTENTS. page.

page.

Chapter 1.

Some Linear Perturbation Theory based on the Persistence of Certain Invariant Subspaces of pairs of linear operators (F,G), the Implicit Function Theorem and the Identity $F \circ h = G \circ h \circ c$.

1. Introduction. Let X and Y be Banach spaces over a common field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Basic results in linear perturbation theory for pairs of bounded linear operators $F,G : X \to Y$ are derived from the analytic dependence on (F,G) of some solutions, bounded linear maps (h,c) of the conjugacy relation $F \circ h = G \circ h \circ c$. In this identity the small circle $\circ$ (and in the following, juxtaposition) indicates the composition or multiplication of operators. Here $h : N \to X$ and $c : N \to N$ are linear maps whose common domain N is a fixed Banach space over $\mathbb{K}$. A subspace S of X is an invariant subspace of the ordered pair of operators (F,G) if $F(S) \subset G(S)$. Hence with this definition of invariant subspace, the range S of h is an invariant subspace as $F(S) = G \circ h(c(N)) \subset G(S)$. Moreover by definition, this invariant subspace $S = h(N)$ is "persistent" if h depends analytically on (F,G).

Eigenvalues r of (F,G) are characterized by the presence of non-zero solutions x of the eigenvalue problem $Fx = rGx$. Eigenvalues r of (c, Id_N) similarly defined, are affiliated with eigenvalues r of (F,G) since by linearity $cn = rn$ $(n \in N)$ implies $Fx = rGx$ when $x = h(n)$. Hence if $N = \mathbb{K}$ then $(r,x) = (c(1), h(1))$ implies by linearity that $Fx = rGx$ since $c(1) = r1$. Analytic dependence of (h,c) on (F,G) in this one dimensional case reduces to an eigenvalue-eigenvector perturbation result in M. G. CRANDALL and P. H. RABINOWITZ (1). In general, dimension $(N) > 1$, and standard first-order single and multi-eigenvalue perturbation formulas follow from Fréchet differentiation of $F \circ h = G \circ h \circ c$ when (h,c) are operator-valued functions of (F,G). Such differentiation is feasible because composition of linear maps is a bounded multi-linear operation on operator-normed vector spaces of bounded linear maps.

The discussion of Floquet representations, exponents and multipliers for linear periodic systems of ordinary differential equations in section 2 provided the initial motivation for this chapter.

In T. KATO (5), $Fx = rGx$ is described as a generalized eigenvalue problem in which usually $G : X \to Y$ is a bijective operator. This latter condition permits $Fx = rGx$ to be rewritten as $G^{-1} \circ F x = rx$. In the theory below, G is required to be invertible only on the invariant subspace S. The continuity of the total eigenvalue projections in T. KATO (5) corresponds here to the persistence of some invariant subspaces S under perturbations of (F,G).

Additional motivation and a model for this chapter and chapter 3 is supplied by the non-differentiable theory for the persistence of invariant manifolds in C. HIRSCH, C. PUGH and M. SHUB (4). This "non-differentiable" theory based directly on certain linearizations and Lipschitz constants contrasts with the "differentiable" theories in this chapter, in chapter 3 and in J. MATHER (6), which depend on the Fréchet differentiation of composition maps, the implicit function theorem and the inversion of linearized operators of the form $(I - M)$ in which the spectral radius of M is less than 1.

The rest of this chapter is composed as follows. Properties of Floquet exponents are indicated in section 2. In section 3 after the definition in sub-section 3.A. of simple and semi-simple eigenvalues of (F,G) and after the introduction in sub-section 3.B. of notation for spectral radii and for Banach spaces of bounded linear operators, two preliminary lemmas 1 and 2 are given. The second of these lemmas requires one of six alternative, but not equivalent, hypotheses H1 to H6. The one and only invariant-subspace persistence-perturbation result of this chapter is given in theorem 1 in section 4. Lemmas 1 and 2 are required in lemmas 3 and 4 to indicate sufficient conditions for the hypotheses of theorem 1 to hold, but for the proof of theorem 1 and the perturbation of semi-simple and simple eigenvalues in theorem 2 of section 5 these 4 lemmas 1,2,3 and 4 are not needed. In section 6 these lemmas are employed to show the persistence under perturbations of Floquet representations. In the last section, section 7, standard first-order perturbation formulas follow from the Fréchet differentiation of the identity $F \circ h = G \circ h \circ c$ with respect to variations in (F,G) when the linear maps (h,c) depend analytically on (F,G).

2. Properties of Floquet Exponents.

The properties of Floquet exponents and multipliers described below are intended to partially motivate the derivation of linear perturbation results in this chapter. The question of continuously varying Floquet representations for continuously disturbed, periodic systems of linear ordinary differential equations is answered in section 6. A reference for the Floquet theory described in this section is J. K. HALE'S textbook (3).

Let $A(t)$ be a T-periodic $m \times m$ continuous matrix-valued function of t in $\mathbb{R}$. Then every fundamental $m \times m$ matrix solution $Z(t)$ of the matrix differential equation

(1) $$\frac{dZ}{dt}(t) = A(t)Z(t)$$

has a non-unique Floquet representation

(2) $$Z(t) = H(t)e^{-Ct}$$

in which C is an $m \times m$ matrix and $H(t)$ is a $m \times m$ matrix-valued T-periodic differentiable function. Substitution of the Floquet representation (2) in the differential equation (1) yields

(3) $$\frac{dH}{dt}(t) - A(t)H(t) = H(t)C$$

after cancellation of the exponential term e^{-Ct}.

From (3) observe whenever η in $\mathbb{C}^m$ is an eigenvector or pseudo-eigenvector of the $m \times m$ matrix C that $H(t)\eta$ is an eigenfunction or pseudo-eigenfunction of the operator $\frac{d}{dt} - A(t)$ acting on the space of T-periodic differentiable functions valued in $\mathbb{C}^m$. Thus the eigenvalues of the matrix C are eigenvalues, or by definition negative Floquet exponents, of the operator $\frac{d}{dt} - A(t)$ with respect to the identity-inclusion operator I embedding the vector space of T-periodic differentiable functions valued in $\mathbb{C}^m$ into the vector space of T-periodic continuous functions valued in $\mathbb{C}^m$. The columns $H(t)e_j$ of $H(t)$ span an m-dimensional vector space which, because of (3), is mapped by the differential operator $F = \frac{d}{dt} - A(t)$ into itself. In the language of the introduction, $G = I$ the identity inclusion-map, $hn = H(t)n$, $cn = Cn$ and $N = \mathbb{K}^m = \mathbb{R}^m$ or $\mathbb{C}^m$.

From the Floquet representation (2), $Z(0) = P(0)$ and hence

$$e^{-CT} = Z(0)^{-1}Z(T)$$

since $Z(T) = H(T)\exp(-CT) = Z(0)\exp(-CT)$. Thus $(-CT)$ is the log of an invertible matrix. But the imaginary part of this log function is multi-valued. Consequently, the real parts of $-CT$ and hence of C are uniquely determined, while the imaginary parts of the eigenvalues of C associated with the invariant subspaces of $Z(0)^{-1}Z(T)$ are determined modulo $\frac{2\pi i}{T}$. The eigenvalues of $\exp(-CT)$ are called Floquet multipliers. The importance of the Floquet representation is that the signs of the real parts of the eigenvalues of $-C$ (i.e. the Floquet exponents) govern the asymptotic behaviour of the solutions $x(t)$ of the equation

$$\frac{dx}{dt} - A(t)x = 0$$

as $t \to \infty$. There is exponential decay if these real parts are all negative.

In the above discussion one may have tacitly assumed that T was the least period of the function $A(t)$, or that this matrix-valued function was not constant, but neither assumption is necessary. The Floquet exponents of $\frac{d}{dt} + A(t)$ when $A(t)$ is constant are just the eigenvalues of $-A(0)$, modulo $\frac{2\pi i}{T}$. For instance, suppose $B(t,u)$ is an $m \times m$ continuous matrix function of t which converges uniformly as $u \to 0$ to $A(0)$, and which has period T in t. Then

(i) the Floquet exponents of the differential operator $\frac{d}{dt} + B(t,u)$ should approach those of $\frac{d}{dt} + A(t)$ as $u \to 0$;

(ii) there should be a family $Z_u(t)$ of fundamental solutions of the differential equations

$$\frac{dZ}{dt}(t) - B(t,u)Z(t) = 0$$

with Floquet representations

$$Z_u(t) = H_u(t)e^{-C(u)t}$$

in which the $m \times m$ matrix $C(u)$ and the $m \times m$ matrix-valued differentiable T-periodic function $H_u(t)$ satisfy

$$\lim_{u \to 0} C(u) = A(0)$$

and

$$\lim_{u \to 0} H_u(t) = I_{m \times m} \qquad \text{(uniformly in t.)}$$

Here the limiting Floquet representation

$$Z_o(t) = I_{m \times m} e^{-A(0)t}$$

is a fundamental matrix for the limit differential equation

$$\frac{dx}{dt} - A(t)x = 0.$$

The continuity properties (i) and (ii) of the Floquet exponents and Floquet representations will follow from the more general considerations of section 6.

3.A. Definition of G-Simple and G-semi-simple Eigenvalues of F.

Let $F,G : X \to Y$ be a pair of linear operators with domains in X. The following definition appears in M. G. CRANDALL and P. H. RABINOWITZ (1).

Definition 3.1. G-Eigenvalues of F. A complex number r is a G-eigenvalue of F if there is a non-zero element x in X with $Fx = rGx$. The elements of the nullspace of $F - rG$ are the G-eigenvectors of F associated with r. A G-eigenvalue r of F is G-semi-simple if the dimension of this nullspace is finite and if $Fx \neq rGx$ implies $(F - rG)x \notin G(\text{kernel}(F - rG))$. A G-semi-simple eigenvalue r of F is G-simple if the nullity of $F - rG$ is 1.

Note below G-eigenvalues of F will also be called eigenvalues of (F,G) or eigenvalues of F with respect to G.

3.B. Notation for Spaces of Bounded Linear Operators and for Spectral Radius.

For any pair X and Y of Banach spaces over $\mathbb{K}$, let $B(X,Y)$ denote the Banach space of linear operators $L : X \to Y$ with bounded operator-norm

$$\|L\|_{B(X,Y)} = \sup_{x \in X, \|x\|_X = 1} \|Lx\|_Y$$

If $X = Y$, let $B(X) = B(X,Y)$. For L in $B(X)$ define the spectral raduis of L by

$$r(L) = \limsup_{j \to \infty} \left(\|L^j\|_{B(X)} \right)^{1/j}$$

Note if $r(L) < 1$ then by the root test

$$(I - L)^{-1} = \Sigma_{j=0}^{\infty} L^j$$

is an element of the Banach space $B(X)$.

3.C. Two Preliminary Lemmas.

Remark. Lemmas 1 and 2 in this sub-section and lemmas 3 and 4 in section 4 are not employed until section 6.

Let N and V be Banach spaces over $\mathbb{K}$. Let $a \in B(V)$ and $b \in B(N)$.

Lemma 1. For $d : N \to V$ put $Md = adb$. Then $M : B(N,V) \to B(N,V)$ is linear, bounded and has spectral raduis $r(M) < r(a)r(b)$, so that $r(a)r(b) < 1$ implies $(I - M)^{-1} : B(N,V) \to B(N,V)$ is a bounded linear operator. Further

$$(I - M)d = d - adb.$$

Proof. For $j > 1$, $M^j d = a^j d b^j$. Therefore for $\|d\| = 1$, $\|M^j d\| < \|a^j\| \cdot \|b^j\|$ and hence

$$\|M^j\| < \|a^j\| \, \|b^j\|.$$

Q.E.D.

Lemma 2. If any of the six hypotheses listed below hold then the linear map $L : B(N,V) \to B(N,V)$ defined by

$$Ld = ad - db$$

for d in $B(N,V)$, has an inverse L^{-1} in $B(B(N,V))$. i.e. L is surjective and bounded below.

The six hypotheses H1, H2, H3, H4, H5 and H6 occur in dual pairs (H1,H2), (H3,H4) and (H5,H6) (see proof for the effects of this duality.) The hypotheses H1 and H3 are special instances of H5, and H2 and H4 are special instances of H6.

H1. $a^{-1} \in B(V)$, $b \in B(N)$ and $r(a^{-1})r(b) < 1$.

H2. $b^{-1} \in B(N)$, $a \in B(V)$ and $r(b^{-1})r(a) < 1$.

H3. N is finite dimensional and it has a Jordan Canonical basis $e_1, \dots e_n$ for $b : N \to N$. There are n scalars r_j in $\mathbb{K}$ such that $b e_1 = r_1 e_1$, $be_j \in \{r_j e_j, r_j e_j + e_{j-1}\}$ for $2 < j < n$, and $(a - r_j I_V)^{-1} \in B(V)$ for $1 < j < n$. $(n > 1.)$

H4. V is finite dimensional and it has a Jordan Canonical basis $e_1, \dots e_n$ for $a : V \to V$. There are n scalars r_j in $\mathbb{K}$ such that $a_n e_n = r_n e_n$, $ae_j \in \{r_j e_j, r_j e_j + e_{j+1}\}$ for $2 < j < n$, and $(b - r_j I_N)^{-1} \in B(N)$ for $1 < j < n$. $(n > 1.)$

H5 V is a direct sum of closed subspaces $W_1, \dots W_n$ $(n > 1.)$ There are n scalars r_j in $\mathbb{K}$, $h \in (0,1) \subset \mathbb{R}$, and n continuous projections $P_j : V \to W_j$ with $(P_j)^2 = P_j$ and $P_j a P_m = \delta_{jm} a P_m$ for $1 < j,m < n$ such that for $1 < j < n$ at least one of the following holds:

(i) $(a - r_j I_V)P_j \in B(W_j)$, $b^{-1} \in B(Y)$ and $r((a - r_j I_V)P_j)r(b^{-1}) < h < 1$

(ii) $[(a - r_j I_V)P_j]^{-1} \in B(W_j)$, $b \in B(Y)$ and $r([(a - r_j I_V)P_j]^{-1})r(b) < h < 1$.

H6. N is a direct sum of closed subspaces $Z_1, \dots Z_n$ $(n > 1.)$ There are n scalars r_j in $\mathbb{K}$, $h \in (0,1) \subset \mathbb{R}$, and n continuous projections $Q_j : N \to Z_j$ with $(Q_j)^2 = Q_j$ and $Q_j b Q_j = bQ_j$ for $1 < j < n$. such that for $1 < j < n$ at least one of the following holds:

(i) $a^{-1} \in B(V)$, $(b - r_j I_N)Q_j \in B(Z_j)$ and $r(a^{-1})r((b - r_j I_N)Q_j) < h < 1$

(ii) $a \in B(V)$, $[(b - r_j I_N)Q_j]^{-1} \in B(Z_j)$ and $r(a)r([(b - r_j I_N)Q_j]^{-1}) < h < 1$.

Proof of lemma 2.

For (H1), write $Ld = a(d - a^{-1}db) = a\hat{L}d$ where by lemma 1, $\hat{L}d = d - a^{-1}db$ has an inverse. Therefore $L^{-1}d = (\hat{L})^{-1}a^{-1}d$, is a bounded linear operator on $B(N,V)$.

For (H2), write $Ld = -(d - adb^{-1})b = (-\hat{L}d)b$ where by lemma 1, $\hat{L}d = d - adb^{-1}$ has an inverse. Therefore $L^{-1}d = -(\hat{L})^{-1}(db^{-1})$ is a bounded linear map on $B(N,V)$.

For (H3), $ad - db = Ld = c \in B(N,V)$ iff

$$ce_1 = (ad - db)e_1 = (a - r_1 I_V)de_1$$

and

$$ce_j = (ad - db)e_j = \begin{cases} (a - r_j I_V)de_j & \text{if } be_j = r_j e_j, \\ (a - r_j I_V)de_j + e_{j-1} & \text{if } be_j = r_j e_j + e_{j-1}, \end{cases}$$

iff the effect of d on the basis $e_1, \ldots e_n$ of N is given by

$$(1) \qquad de_j = \begin{cases} (a - r_j I_V)^{-1}ce_j & , \text{ if } be_j = r_j e_j, \\ (a - r_j I_V)^{-1}(ce_j - de_{j-1}) & , \text{ if } be_j = r_j e_j + e_{j-1}. \end{cases}$$

By induction, starting with $j = 1$, the expressions in (1) recursively determine each de_j, and hence $d : N \to V$, as bounded linear functions of c for which $Ld = ad - db = c$.

For (H4), let $e^1, \ldots e^n$ in the dual space V^* of V be the dual basis of V^* to the basis $e_1, \ldots e_n$ of V. Then $ad - db = Ld = c \in B(N,V)$ iff for $1 < j < n$

$$e^j c = e^j(ad - db) = e^j(a(\Sigma_{m=1}^n e_m e^m d)) - e^j db$$

since $I_V = \Sigma_{m=1}^n e_m e^m(\cdot)$. Therefore $Ld = c$ iff for $1 < j < n$

$$e^j d = \begin{cases} e^j d(r_j I_N - b) & , \text{ if } j = 1 \text{ or if } be_{j-1} = r_{j-1}e_{j-1}, \\ e^j d(r_j I_N - b) + e^{j-1}d & , \text{ if } be_{j-1} = r_{j-1}e_{j-1} + e_j . \end{cases}$$

Hence, by induction on j, starting with j = 1, the formulas

$$e^j d = \begin{cases} e^j d(r_j I_N - b)^{-1}, & \text{if } j = 1 \text{ or if } ae_{j-1} = r_{j-1}e_{j-1}, \\ (e^j c - e^{j-1}d)(r_j I_N - b)^{-1}, & \text{if } 1 < j < n \text{ and } ae_{j-1} = r_{j-1}e_{j-1} + e_j, \end{cases}$$

recursively determines each $e^j d$, and hence $d = \Sigma_{j=1}^n e_m(e^j d)$, as bounded linear functions of $c \in B(N,V)$ such that d is the only solution in B(N,V) of $Ld = ad - db = c$.

For (H5), $Ld = ad - db = c \in B(N,V)$ iff for $1 < j < n$

$$P_j c = P_j(ad - db) = P_j(a\Sigma_{m=1}^n P_m d) - P_j db$$

$$= (\Sigma_{m=1}^n P_j(P_m a P_m d)) - P_j db = aP_j d - P_j db$$

$$= (a - r_j I_V)P_j d - P_j d(b - r_j I_N)$$

since $I_X = \Sigma_{m=1}^n P_j$ and $P_j a P_m = \delta_{jm} a P_m$. But by the previous arguments for (H1) and (H2), the last equality in the above expression uniquely determines each $P_j d$, and hence $d = \Sigma_{j=1}^n (P_j d)$, as bounded linear functions of c, for which $ad - db = c$ has d as its only solution in B(N,V).

Finally, for (H6), $ad - db = c \in B(N,V)$ iff for $1 < j < n$

$$cQ_j = (ad - db)Q_j \quad adQ_j - dQ_j bQ_j$$

$$= (a - r_j I_V)dQ_j - dQ_j(b - r_j I_N)Q_j$$

since $(Q_j)^2 = Q_j$ and $Q_j b Q_j = bQ_j$. So again by the arguments for (H1) and (H2), $L : B(N,V) \to B(N,V)$ is surjective and it has a bounded inverse.

Q.E.D.

4. On The Persistence of Invariant Subspaces. The Identity $F \circ h = G \circ h \circ c$.

Let X and Y be vector spaces over $\mathbb{K}$. Let F,G be a pair of linear operators mapping X into Y.

Definition 4.1. Invariant Subspace of (F,G). A subspace N of X is an invariant subspace of (F,G) if

$$F(N) \subseteq G(N).$$

Observe if the restriction $G_N : N \to Y$ of G to N is injective and if F_N denotes the restriction of F to N, then

$$c = G_N^{-1} F_N : N \to N \text{ and } h = Id_N : N \to X$$

satisfy the identity

$$F \circ h = G \circ h \circ c. \tag{1}$$

Conversely, if $c : N \to N$ and $h : N \to X$ ((h,c) arbitrary) satisfy (1) and if $S = \text{range}(h) = h(N)$ then $F(S) = G(h(c(N)) \subseteq G(S)$ and S is an invariant subspace of (F,G).

Now let X be a normed vector space with norm $\|x\|$.

Definition 4.2. A Gap between Subspaces. Let S and $\hat{S}$ be subspaces of X and let $d_H(S,\hat{S}) = \sup_{\substack{s \in S \\ \|s\| \le 1}} \inf_{\substack{\hat{s} \in \hat{S} \\ \|\hat{s}\| \le 1}} \|s - \hat{s}\| \le 1$ measure the gap between the unit balls of S and $\hat{S}$.

Note $d_H(S,\hat{S}) = 0$ if the closures of S and $\hat{S}$ in X are equal. Further if $P : S \to \hat{S}$ is a projection which is surjective and bounded below (with $\|P\| \le 1$) then for all s in S

$$\inf_{\hat{s} \in \hat{S}, \|s\| \le 1} \|s - \hat{s}\| \le \|s - Ps\| \le (\|I - P\|)\|s\|.$$

and hence $d_H(S,\hat{S}) \le \|I - P\|$.

Theorem 1 and lemma 4 and 5 below contain sufficient conditions for the persistence of invariant subspaces S of a pair of operators (F,G).

Theorem 1. On the Persistence of Invariant Subspaces. Let X and Y be the Banach spaces over $\mathbb{K}$. Let X be a direct sum of closed complementary subspace N and V. Let Y be a direct sum of closed complementary subspaces Y_1 and Y_2. Let $P : Y \to Y_2$ be the continuous projection of Y onto Y_2 given by $P(y_1 + y_2) = y_2$ when $y_1 \in Y_1$, $y_2 \in Y_2$. Let F,G be a pair of bounded linear maps of X into Y. Suppose N is an invariant subspace of (F,G) with $F(N) \subseteq G(N) = Y_1$. Let $G_N : N \to Y_1$ have a bounded inverse. Assume $F_N : N \to Y_1$ is bounded. Set $b = G_N^{-1}F_N : N \to N$. Finally assume the linear map $L : B(N,V) \to B(N,Y_2)$ defined for $v : N \to V$ in B(N,V) by

(1) $$Lv = P(Fv - Gvb)$$

is surjective and has a bounded inverse. (Lemmas 3 and 4 below, and theorem 2 and its corollary in the next section indicate sufficient conditions for this.)

Then there is a neighbourhood W of (F,G) in $B(X,Y)^2$ and operator-valued functions h(f,g) and c(f,g) valued in B(N,X) and B(N) respectively, which depend analytically on (f,g) in W, and which satisfy (i) the identity

$$fh(f,g) = gh(f,g)c(f,g)$$

in which juxtaposition indicates the multiplication of linear operators; and (ii) the "initial conditions"

$$h(F,G) = I_N : N \to X,$$

$$c(F,G) = G_N^{-1}F_N : N \to X.$$

Proof of Theorem 1. For (f,g,c,v) in $B(X,Y)^2 \times B(N) \times B(N,V)$, set

$$A(f,g,c,v) = f(I_N + v) - g(I_N + v)c$$

Then A is a $B(N,Y)$ valued analytic function of its arguments because it is the difference of multi-linear bounded operator-valued functions on $B(X,Y)^2 \times B(N) \times B(N,V)$. Further

$$A(F,G,b,0) = FI_N - GI_N b = 0$$

since $b = G_N^{-1}F_N = (GI_N)^{-1}(FI_N)$. The partial (c,v) Fréchet derivative of A at $(F,G,b,0)$ is the $B(N,Y)$ valued linear map

(1) $$Q(c,v) = (Fv - Gvb) - Gc$$
$$= Lv + [(I - P)(Fv - Gvb) - G_N c]$$

since $Lv = P(Fv - Gvb)$. Thus $Q(c,v) = y \in B(N,Y)$ implies

(2) $$v = L^{-1}Py$$
$$c = G_N^{-1}(I - P)[(Fv - Gvb) - y]$$

These formulas show $Q : B(N) \times B(N,V) \to B(N,Y)$ to be surjective and bounded below. Hence by the implicit function theorem there is a product neigbourhood $W \times U$ of $(F, G, b, 0)$ in $B(X,Y)^2 \times (B(N) \times B(N,V))$, in which the solutions of $A(f,g,c,v) = 0$ form the graph of analytic operator-valued functions $(c(f,g), v(f,g))$ mapping the neighbourhood W of (F,G) into the neighbourhood U of $(b,0)$ in $B(N) \times B(N,V)$ with

$$(c(F,G), v(F,G)) = (b,0) = (G_N^{-1}F_N, 0).$$

The conclusions (i) and (ii) are satisfied by setting

$$(h(f,g), c(f,g)) = (I_N + v(f,g), c(f,g)).$$

Q.E.D.

<u>Corollary 1.1. (Change of Parameters.)</u> Let $\hat{N}$ be a Banach space over $\mathbb{K}$. Suppose there is an isomorphism $\phi : \hat{N} \to N$. Then for (f,g) in W the operators

$$\hat{h}(f,g) = h(f,g)\phi : \hat{N} \to X$$

$$\hat{c}(f,g) = \phi^{-1}c(f,g)\phi : \hat{N} \to \hat{N}$$

satisfy (i) the identity

$$f\hat{h}(f,g) = g\hat{h}(f,g)\hat{c}(f,g)$$

and (ii) the "initial conditions"

$$\hat{h}(f,g) = \phi : \hat{N} \to X$$

$$\hat{c}(f,g) = \phi^{-1}G_N^{-1}F_N\phi = \phi^{-1}b\phi \quad : \hat{N} \to \hat{N}$$

<u>Proof:</u> This is obvious by substitution.

Q.E.D.

Lemma 3. Suppose $PF_V : V \to Y$ is bounded below and

$$r((PF_V)^{-1}PG_V)r(b) < 1 \qquad (b = G_N^{-1}F_N : N \to N)$$

then $Lv = P(Fv - Gvb)$ is surjective and bounded below.

Proof. By lemma 1, $Lv = (PF_V)[v - (PF_V)^{-1}(PG_V)vb]$ is a product L_1L_2 of surjective and bounded below operators defined by

$$v \in B(N,V) \xrightarrow{L_1} (PF_V)v$$
$$v \in B(N,V) \xrightarrow{L_2} v - (PF_V)^{-1}(PG_V)vb.$$

Thus L is surjective and bounded below since $L^{-1} = L_2^{-1}L_1^{-1}$.

Q.E.D

Lemma 4. Suppose $PG_V : V \to Y_2$ is bounded below and the map of v in B(N,V) to

$$av - vb = (PG_V)^{-1}(PF_V)v - vG_N^{-1}F_N$$

in B(N,V) is surjective and bounded below (see the sufficient conditions in Hypotheses H1 to H6 of lemma 2) then $Lv = P(Fv - Gvb)$ is surjective and bounded below.

Proof. Again $Lv = (PG_V)(av - vb)$ is a product of invertible maps which are bounded below.

Q.E.D

Note $PG_V : V \to Y_2$ is bounded below iff $G : X \to Y$ is bounded below since G can be identified with the triangular matrix in

$$\begin{pmatrix} G_N & , 0 \\ (I - P)G_N, & PG_V \end{pmatrix} \begin{pmatrix} n \\ v \end{pmatrix} = \begin{pmatrix} G_N n \\ (I - P)Gn + PGv \end{pmatrix} \in Y_1 \times Y_2$$

and since 2 x 2 triangular matrices are surjective and bounded below iff their diagonal elements are surjective and bounded below and their off-diagonal elements are bounded.

5. Perturbation of Simple and Semi-Simple Eigenvalues.

In theorem 1, if $b = (G_N)^{-1}F_N = rI_N$ then r is an eigenvalue of (F,G) and kernel(F - rG) = N. Further r is a simple or semi-simple eigenvalue if the dimension of N is finite. The general case dimension(N) > 1 and the special case dimension(N) = 1 are treated in theorem 2 and corollary 2.1. below.

Theorem 2. Perturbation of Semi-Simple Eigenvalues. Let $F,G : X \to Y$ be bounded linear operators. Let r^0 be a simple or semi-simple G-eigenvalue of F. Suppose $N = \text{kernel}(F - r^0G)$ has a closed complement V in X. Further assume $G_N : N \to Y$ is bounded below, and $Y_2 := \text{range}(F - r^0G)$ is a closed complement of $Y_1 = G(N)$ in Y. Then there is a neighbourhood W of (F,G) in $B(X,Y)^2$ and analytic operator-valued functions (h(f,g),c(f,g)) defined on W and valued in $B(N,X) \times B(N)$ such that

$$fh(f,g) = gh(f,g)c(f,g) \qquad (\text{for } (f,g) \text{ in } W)$$

$$h(F,G) = I_N : N \to X$$

and

$$c(F,G) = r^0 I_N : N \to N$$

Proof. The operator $Lv = (F_V - r^0G)v = Fv - Gvb \ (v \in B(N,V))$ in theorem 1 is surjective and bounded below.

Q.E.D.

Corollary 2.1. Perturbation of Simple Eigenvalues. In theorem 2 if dimension(N) = 1 then there are analytic functions (x(f,g),r(f,g)) defined on a neigbourhood W of (F,G) in $B(X,Y)^2$ and valued in $X \times \mathbb{K}$ such that

$$fx(f,g) = r(f,g)gx(f,g) \qquad (\text{for } (f,g) \text{ in } W)$$

$$x(F,G) \neq 0$$

and

$$r(F,G) = r^0.$$

Proof. Let $(h(f,g),c(f,g))$ be as in the conclusion of theorem 2. Now identify N with $\mathbb{K}$ and put $x(f,g) = h(f,g)(1)$ and $r(f,g) = c(f,g)(1)$.

Q.E.D.

6. Perturbation of Floquet Exponents and Floquet Representations.

6.A. Floquet Theory. Fix $T > 0$. Let $K(t)$ be a continuously differentiable, $m \times m$ matrix-valued function with period T in the real-variable t and coefficients in $\mathbb{K}$. Let $A(t)$ be a continuous, T-periodic, $m \times m$ matrix function of t, again with coefficients in $\mathbb{K}$. Let

(1) $$Fx = K(t)\frac{dx}{dt} - A(t)x, \qquad Gx = K(t)x$$

when $x = x(t)$ is a continuously differentiable function valued in $\mathbb{K}^m (m>1.)$ Let $Z(t)$ be an $m \times m$ fundamental matrix for the differential equation $Fx = 0$. i.e. $Z(t)$ satisfies the differential equation

(2) $$FZ(t) = K(t)\frac{dZ(t)}{dt} - A(t)Z(t) = 0$$

with a non-singular initial value $Z(0)$. Then $Z(t)$ has a (first) Floquet representation

$$Z(t) = H(t)e^{-Ct} \qquad (\, H(t) \;\; m \times m, \; C \;\; m \times m.)$$

in which C is defined by $\exp(-CT) = Z(0)^{-1}Z(T)$. Necessarily $H(0) = H(T)$ since $Z(T) = H(T)Z(0)^{-1}Z(T)$ and since $Z(0) = H(0)$. By induction, it follows that $H(t)$ is a T periodic continuously differentiable function.

Substitution of the first Floquet representation into equation (2) yields

(3) $$FH(t) = K(t)\frac{dH(t)}{dt} - A(t)H(t) = K(t)H(t)C = G(\,H(t)C\,).$$

Similarly Z(t) has a second Floquet representation

(4) $$Z(t) = e^{-Rt}Q(t) \qquad (Q(t)\ \ m \times m,\ R\ \ m \times m)$$

in which $\exp(-RT) = Z(T)Z(0)^{-1}$. Hence $Q(T) = Q(0)$ and consequently $Q(t)$ is a T-periodic continuously differentiable function since by the definition of R, $Z(T) = e^{-RT}Q(T) = Z(T)Z(0)^{-1}Q(T)$ and $Z(0) = Q(0)$. Here $Y(t) = K^*(t)^{-1}Q^*(t)^{-1}\exp(-(-R^*t))$ will be the Floquet representation of a fundamental matrix of an adjoint differential operator F^* to F.

6.B Periodic Function Spaces, Adjoint Operators and Fredholm's Alternative.

Let Y be the Banach space of continuous T-periodic functions $y(t)$ with bounded sup-norms $\|y\|_Y = \max\{\,\|y(t)\|_{\mathbb{K}^m} : 0 < t < T\}$. Let $X \subset Y$ be the Banach space of continuously differentiable functions $x(t)$ valued in $\mathbb{K}^m$ with a finite norm $\|x\|_X = \|x\|_Y + \|\frac{dx}{dt}\|_Y$. Define an inner product on Y and on $X \subset Y$ by

$$(f,g)_T = \int_0^T \langle f(t),g(t)\rangle_{\mathbb{K}^m}\, dt = \int_0^T f(t)\, g(t)\, dt. \quad (f,g \in X \text{ or } Y.)$$

Note the restrictions of F and G to X are continuous linear operators valued in the Banach space Y.

For the T-periodic, differentiable functions $x(t), y(t)$ in X , integration by parts and the periodicity imply

$$(y,Fx)_T = (y,K(t)\frac{dx}{dt} - A(t)x)_T = \left(\left[-K^*(t)\frac{dy}{dt} - (\frac{dK^*}{dt}(t) + A^*(t))y\right],x\right)_T$$

$$= (F^*y,x)_T ,$$

if the adjoint operator F^* is defined by

$$F^*y = -K^*(t)\frac{dy}{dt} - (\frac{dK^*}{dt}(t) + A^*(t))y.$$

The same computation indicates $(F^*)^*x = Fx$.

Let $Y(t)$ be a $m \times m$ fundamental matrix solution of the adjoint equation $F^*y = 0$ which satisfies the initial condition $Y(0) = Z^*(0)^{-1}K^*(0)^{-1}$. Then for all t in $\mathbb{R}$

$$\frac{d}{dt}(Y^*(t)K(t)Z(t)) = \left(\frac{dK^*}{dt}Y + K^*\frac{dY}{dt}\right)^*Z + Y^*K\frac{dZ}{dt} = 0 ,$$

but $Y^*(0)K(0)Z(0) = I_{m\times m}$. Therefore

(5) $$Y(t) = K^*(t)^{-1}Z^*(t)^{-1}$$

is a fundamental solution of the adjoint equation $F^*y = 0$. From the second Floquet representation (4) of $Z(t)$, a first Floquet representation of $Y(t)$ is

$$Y(t) = K(t)^{*-1}(e^{-Rt}Q(t))^{*-1} = (K(t)^{*-1}Q(t)^{*-1})\, e^{R^*t}.$$

Fredholm's Alternative. Variation of Parameters.

For the inhomogeneous differential equation

$$Fx = K(t)\frac{dx}{dt} - A(t)x = g(t)$$

in which $g(t)$ is a continuous function of t, variation of parameters yields

$$x(t) = Z(t)\left(c_0 + \int_0^t Z(s)^{-1}K(s)^{-1}g(s)ds\right)$$

as the general solution which satisfies the initial condition $x(0) = Z(0)c_0$. Observe from (5) that the collection of functions $\{K(T)^{*-1}Z(T)^{*-1}y_0 : y_0 \in \mathbb{K}^m\}$ is the solution space of the adjoint equation $F^*y = 0$. For simplicity, suppose $Z(0) = I_{m\times m}$ the identity matrix. Then for T-periodic functions $g(t)$ in Y the periodic differential equation $Fx = g$ has a periodic solution $x(t)$ in X

iff
$$x(0) = Z(T)\left(x(0) + \int_0^T Z(s)^{-1}K(s)^{-1}\, g(s)ds\right)$$

iff
$$(I - Z(T))x(0) = \int_0^T Z(s)^{-1}K(s)^{-1}\, g(s)ds$$

iff for all y_0 in kernel$((I-Z(T))^*)$ = kernel$(K(T)^{*-1}(Z(T)^{*-1}-I))$ = kernel$(Y(T) - Y(0))$,

$$0 = \int_0^T y_0^* Z(s)^{-1}K(s)^{-1}\, g(s)\, ds = \int_0^T \langle Y(s)y_0, g(s)\rangle_{\mathbb{K}^m} ds = (Y(t)y_0, g(t))_T$$

iff for all $y = Y(t)y_0$ in $\text{kernel}_X(F^*)$

$$(y(t),g(t))_T = 0$$

Therefore in Y

(6) $\text{range}(F_X) = \{\text{kernel}(F_X{}^*)\}^\perp = \{y \in Y : x \in \text{kernel}_X(F_X) \text{ implies}(y,x) = 0\}$

Likewise, in Y

(7) $$\text{range}(F_X{}^*) = \{\text{kernel}(F_X)\}^\perp$$

Finally

$$\begin{aligned}\text{nullity}_X(F) &= \text{nullity}(Z(T) - I_{m \times m}) \\ &= \text{nullity}(Y(T) - Y(0)) \\ &= \text{co-dimension}_Y(F_X)\end{aligned}$$

Hence

(8) $$\dim \text{kernel}_X(F) = \text{co-dimension}_Y(\text{range}(F_X))$$

and F has Fredholm index zero. Similarly

(9) $$\dim \text{kernel}_X(F^*) = \text{co-dimension}(\text{range}(F_X))$$

6.C. Generalized Eigenvalue Problems. Floquet Exponents.

Henceforth restrict the operators F and G so that

$$X = \text{domain}(F) = \text{domain}(G).$$

By the preceding argument each of the operators $F - rG$ (r in $\mathbb{K}$) has Fredholm index zero. Put $G^*y = K(t)^*y$. Then the adjoint of the differential operator $(F - rG)^*$ (computed as above) is $F^* - r^*G^*$. Further the Fredholm index of these operators on X is zero. Hence for each eigenvalue r of (F,G) there is an eigenvalue r* of(F*,G*) with the same geometric multiplicity given as (or defined by)

$$\text{nullity}(F - rG) = \text{nullity}(F^* - r^*G^*).$$

The negatives -r of the eigenvalues r of (F,G) are called Floquet exponents.

From the Jordan canonical form of the matrix C in the first Floquet representation $Z(t) = H(t)e^{-Ct}$, there are q orthogonal projections (P_j) of $\mathbb{C}^m$ onto subspaces E_j of $\mathbb{C}^m$ such that

$$C = \Sigma_{j=1}^{q} (\lambda_j I + M_j)P_j$$

for some q nilpotent or zero operators M_j on E_j and some $q \le m$ distinct eigenvalues λ_j, of C. Here each E_j is spanned by the eigenvectors and pseudo-eigenvectors of C associated with the eigenvalue λ_j. The dimension of E_j is the algebraic multiplicity of the eigenvalue λ_j of C.

Note each invariant subspace E_j of $(C, I_{m \times m})$ is a direct sum of $n = \text{nullity}(F - \lambda_j G)$ invariant subspaces which contain, modulo the complex numbers, only one eigenvector of (F,G). Let $Q_1, \ldots Q_w$, $w = \Sigma_{j=1}^{q} r_j$, be projections onto these invariant subspaces. Other Floquet representations of Z(t) are given by

(10) $$Z(t) = (H(t)\exp(\Delta Ct))e^{-(C + \Delta C)t}$$

whenever $\Delta C = \Sigma_{j=1}^{W} \frac{2\pi\alpha_j i Q_j}{T}$ for some integers α_j. There cannot be any more Floquet representations (or Floquet exponents) other than those given in (10) since the equation

$$K(t)\frac{dx}{dt} - A(t)x = 0$$

has at most m linearly independent solutions. In particular modulo $\frac{2\pi i}{T}$, each eigenvalue r of (F,G) is equal to one of the eigenvalues λ_j of the operator C.

Note some bounded linear perturbations of (F,G) in $B(X,Y)^2$ are given by bounded perturbations in X^m and Y^m respectively of the columns of the matrices K(t) and A(t). Here the elements of the product Banach spaces X^m and Y^m are being identified with m x m matrix-valued T-periodic functions of t.

6.D.1. Perturbation of Simple and Semi-Simple Floquet Exponents.

Fix j $(1 \leq j \leq q.)$ Suppose $r = \lambda_j$ is a simple or semi-simple eigenvalue of the matrix C. Put $N = N_j = H(t)E_j$ and let

$$V = N^{\perp} = \{x \in X : (x,y)_T = 0 \text{ for all } y \text{ in } N\}.$$

Then V is a closed complement of N and for x = n + v in X, with n in N, v in V

$$0 \neq (F - rG)v \notin G(\text{kernel}(F - rG)) = G(N)$$

as otherwise there would be an $\hat{n}$ in $N = N_j$ and a v in V such that

$$K(t)\frac{dv}{dt} - A(t)v - rK(t)v = K(t)H(t)\hat{n}.$$

The latter contradicts the simplicity or semi-simplicity of the Floquet exponent r of (F,G) as it implies the existence of a Floquet representation not of the form given in (10).

Let

$$Y_2 = \text{range}(F - rG) = \text{kernel}((F - rG)^*)^{\perp}$$

be a closed subspace of Y. From nullity($(F - rG^*)$ = nullity$(F - rG)$ = dimension $G(N)$ and $G(N) \cap Y_2 = \{0\}$, it follows that $Y_1 = G(N)$ is a closed complement of Y_2 in Y. Thus the hypotheses of theorem 2 are satisfied if dimension $N \geq 1$ and those of its corollary 2.1. if dimension$(N) = 1$.

6.D.2. General Case. Perturbation of a non-semi-simple Eigenvalue.

Again fix j ($1 \leq j \leq q$) and let $r = \lambda_j$ be an eigenvalue of C. Put $N = N_j = H(t)E_j$ and

$$V = N^{\perp}(\text{in } X) = \{x \in X : (x,y)_T = 0 \text{ for all } y \text{ in } N\}.$$

N and V are then closed complementary subspaces of X. Put $Y_2 = (F - rG)(V)$ and $Y_1 = G(N)$. Then r is not an eigenvalue of (F_V, G_V) and $Y_1 \cap Y_2 = \{0\}$ as otherwise the dimension of $N = N_j$ would increase. Put

$$N_0 = \{x \in N : (x,y)_T = 0 \text{ for } y \text{ in kernel}(F - rG)\}.$$

Then

$$N_0 + V = \{\text{kernel}((F - rG))\}^{\perp} \qquad (\perp \text{in } X)$$

and

$$(F - rG)(N_0 + V) = \text{range}(F - rG) = \{\text{kernel}((F - rG)^*)\}^{\perp} \qquad (\perp \text{in } Y.)$$

Thus $(F - rG)N_0$ is a finite dimensional complement of $Y_2 = (F - rG)(V)$ in the closed subspace kernel$((F - rG)^*)^{\perp}$ of Y. Therefore Y_2 is closed. Hence $(F - rG)_V : V \to Y_2$ is surjective and bounded below since kernel$(F - rG)_V = \{0\}$. Further Y_2 has co-dimension

$$\text{dimension}(N_0) + \text{nullity}(F - rG)^*_X = \text{dimension}(G(N)).$$

But

$$Y_1 \cap Y_2 = \{0\}.$$

Hence $Y_1 = G(N)$ is a closed complement of Y_2 with dimension $(Y_1) =$ co-dimension$_Y(Y_2)$. Therefore Y_1 and Y_2 are closed complementary subspaces of Y.

Let P be a continuous projection of $Y_1 = G(N)$ onto $Y_2 = (F - rG)V$. In theorem 1 for maps $v : N \to V$ in $B(N,V)$,

$$Lv = P(Fv - Gvb),\ b = G_N^{-1} F_N = rI_{N_j} + M_j$$

where M_j is a nilpotent or zero linear operator. Hence hypothesis (H3) of lemma 2 is satisfied with $r_j = r$ for $1 \le j \le n \equiv \text{dimension}(E_j)$ and with $a = (PG|_V)^{-1}(PF_V) : V \to V$. Note r is not an eigenvalue of $a : V \to V$ since $av = \lambda v$ $(v \in V)$ implies

$$0 = P(F - \lambda G)v = P(F - rG)v + (r - \lambda)PGv$$
$$= (F - rG)v + (r - \lambda)PGv$$

and this implies $\lambda \ne r$ because r is not an eigenvalue of (F_V, G_V). Hence by lemma 4, the persistence theorem 1 is applicable.

By corollary 1.1 with $\hat{N} = E_j$ and the change of parameters $\phi = H(t)P_j : E_j \to N$ there are operator-valued functions

$$\hat{h}_j(f,g)(\cdot) : E_j \to X \qquad \hat{c}_j(f,g)(\cdot) : E_j \to E_j$$

which (i) depend analytically in $B(E_j, X)$ and $B(E_j)$, respectively, on perturbations (f,g) in a neigbourhood W_j of (F,G) in $B(X,Y)^2$, and (ii) satisfy

(11) $\hat{h}(F,G) = H(t)P_j : E_j \to N \subset X, \quad \hat{c}(F,G) = \lambda_j I_j + M_j : E_j \to E_j \subset \mathbb{K}^m$

and

(12) $$f\hat{h}_j(f,g) = g\hat{h}_j(f,g)\hat{c}_j(f,g).$$

6.D.3. Perturbation of Floquet Representations.

For $1<j<q$, the argument in subsection 6.D.2 yields operators $(\hat{h}_j f,g),\hat{c}_j(f,g))$ in $B(E_j,X) \times B(E_j)$ which depend analytically on (f,g) in a neighbourhood $W=\bigcap_{1\le s\le q} W_s$ of (F,G) in $B(X,Y)^2$ and which satisfy the last two equations of subsection 6.D.2. Now recall that $\Sigma_{j=1}^q P_j = I : \mathbb{K}^m \to \mathbb{K}^m$; that $\Sigma_{j=1}^q (I_j + M_j)P_j = C : \mathbb{K}^m \to \mathbb{K}^m$; and that $B(\mathbb{K}^m,X)$ is isomorphic to X^m.

Therefore

$$\hat{H}(f,g) \overset{\text{def'n}}{=} \Sigma_{j=1}^q \hat{h}_j(f,g)P_j = \hat{H}(f,g)(t) : \mathbb{K}^m \to X^m$$

satisfies

$$\hat{H}(F,G) = \Sigma_{j=1}^q H(t)P_j = H(t) \quad : \mathbb{K}^m \to X^m.$$

Likewise

$$\hat{C}(f,g) \overset{\text{def'n}}{=} \Sigma_{j=1}^q \hat{c}_j(f,g)P_j \overset{\text{at }(F,G)}{=} \Sigma_{j=1}^q (I_j+M_j)P_j = C : \mathbb{K}^m \to \mathbb{K}^m.$$

Therefore

$$f\hat{H}(f,g) = g\hat{H}(f,g)\hat{C}(f,g)$$

because $\hat{c}_j(f,g) : E_j \to E_j = \text{range}(P_j) \subset \text{kernel}(P_s)$ if $s \neq j$.

Now if $(k(t),a(t))$ in $Y^m \times Y^m$ are $m \times m$ periodic matrices such that $g = k(t) : X \to Y$ and the differential operator $f = k(t)\frac{d}{dt} + a(t) : X \to Y$ satisfy $(f,g) \in W$ then

$$(k(t)\tfrac{d}{dt} + a(t))\,\hat{H}(f,g)(t) = k(t)\,\hat{H}(f,g)(t)\,\hat{C}(f,g).$$

Thus a Floquet representation of a fundamental matrix of the differential equation

$$k(t)\tfrac{d}{dt}x(t) + a(t)x(t) = 0$$

is given by

$$\hat{Z}(f,g)(t) = \hat{H}(f,g)(t)\exp(-t\hat{C}(f,g)).$$

This represents an analytic continuation of the Floquet representation $Z(t) = H(t)e^{-tC} = \hat{Z}(F,G)(t)$ of the fundamental matrix-solution $Z(t)$ of

$$Fx = K(t)\tfrac{d}{dt}x(t) + A(t)x(t) = 0 \qquad (G = K(t) : X \to Y)$$

since the operators $f = k(t)\frac{d}{dt} + a(t)$ and $g = k(t)$ in $B(X,Y)$ depend linearly, and hence analytically on bounded perturbations $(k(t),a(t))$ of $(K(t),A(t))$ in $Y^m \times Y^m$.

7. Derivation of Standard Perturbation Formulas.

Assume the hypotheses of theorem 1. Let $h = I_N + v$, $v \in B(N,V)$ and

$$A(f,g,c,v) = fh - ghc$$

be as in the proof of theorem 1. Let $h(f,g) = I_N + v(f,g) : N \to X$ and $c(f,g) : N \to N$ be the analytic operator-valued maps appearing in the proof and conclusion of theorem 1. Linear perturbation formulas will be obtained for the first-order changes in $(h(f,g),c(f,g))$ from Fréchet differentiation of the equation $A(f,g,c,v) = 0$ with respect to variations in (f,g) in $B(X,Y)^2$.

Let (df,dg) denote elements of $B(X,Y)$. By the chain-rule

$$0 = DA(F,G,b,0)(df,dg)$$

$$= A_{(f,g)}(F,G,b,0)(df,dg) + A_{(c,v)}(F,G,b,0)(Dh(F,G),Dc(F,G))(df,dg)$$

$$= (df)I_N - (dg)b + Q(Dh(F,G)(df,dg),Dc(F,G)(df,dg))$$

where $Q : B(N,V) \times B(N) \to B(N,Y)$ is the (c,v)-partial Fréchet derivative of $A(f,g,c,v)$ evaluated at $(F,G,b,0)$. Put

$$y = -((df)I_N - (dg)b).$$

From the inverse of Q calculated in the proof of theorem 1, if $v = -L^{-1}Py$ then

$$Dh(F,G)(df,dg) = -L^{-1}Py = v$$

$$Dc(F,G)(df,dg) = (G_N)^{-1}(I - P)(Fv - Gvb - y)$$

Therefore (i) if $b = 0$ or if $G(V) \cup F(V) \subset Y_2 = \text{range}(P)$ then

$$Dc(F,G)(df,dg) = -G_N^{-1}(I - P)(df\, I_N - dg\, b) \; : \; N \to N$$

and (ii) if $b = 0$ then $Lv = P(Fv - Gvb)$ implies

$$Dh(F,G)(df,dg) = (PF_V)^{-1}P(df\, I_N) \; : \; N \to V.$$

The remainder of the calculations in this section concern the matrix representation of $Dc(F,G)(df,dg)$ when N is finite dimensional.

Assume N has a basis $n_1,\dots,n_m$ Then $G(N) \subset Y$ has a basis $Gn_1,\dots,Gn_m$ since G_N is one-to-one on N. By assumption $(I - P) : Y \to Y$ is a continuous projection. Therefore there are m linear functionals $z_j : Y \to \mathbb{K}$ such that

$$(I - P)y = \Sigma_{j=1}^{m} \langle z_j, y\rangle Gn_j.$$

Hence $\langle z_j, Gn_k\rangle = \delta_{jk}$. Further

$$G_N^{-1}(I - P)(df\, I_N - dg\, b)n_k = G_N^{-1}\Sigma_{j=1}^{m} \langle z_j, (df - dg\, b)\, n_k\rangle Gn_j$$

$$= \Sigma_{j=1}^{m} \langle z_j, (df - dg\, b)\, n_k\rangle n_j$$

Therefore if $b = 0$ or if $G(V) + F(V) = \text{range}(P)$ then a matrix representation of $Dc(F,G)(df,dg) : N \to N$ with respect to the above basis of N is given by

$$Dc(F,G)(df,dg) = -(\langle z_j, (df - dg\, b)\, n_k\rangle)_{1\le j,k\le m} :$$

An immediate consequence of this matrix representation is the formula:

$$\text{trace}(Dc(F,G)(df,dg)) = \Sigma_{1\le j\le m} \langle z_j, (df - dg\, b)n_j\rangle$$

for the trace of the first-order change in $c(F+df,G+dg)$. This trace equals the first-order change in the mean-value of the eigenvalues of $c(f,g)$ at (F,G). When $m = 1$ and $b = r^0$ the matrix and trace formulas reduce to the common expression and standard perturbation formula

$$Dr(F,G)(df,dg) = -\langle z, (df - r^0 dg)n\rangle$$

for a simple eigenvalue r^0 of (F,G). Note the special case $r^0 dg = 0$.

The perturbation formulas derived in this section for perturbations of $c(F,G)$, its trace or a simple eigenvalue r^0 give the linear terms in the Taylor expansions of their extensions $c(F+df,G+dg)$, $\text{trace}(c(F+df,G+dg))$ and $r(F+df,G+dg)$ in the presence of perturbations (df,dg) of (F,G). This information can sometimes be useful in the indication of the linearized stability of certain nonlinear problems when Lyapunov's Stability criterion holds.

8. Bibliography for Chapter 1.

1 Crandall M. G. and P. H. Rabinowitz. Bifurcation, Perturbation of Simple Eigenvalues and Linearized Stability. Arch., Rat., Mech., Anal., 52, 1973, 161-180.

2 Dunford N. and J. T. Schwartz. Linear Operators, vols., I (1958), II (1963), III (1971). John Wiley and Sons Inc.

3 Hale J.K. Ordinary Differential Equations, 1969. Vol., XXI in Pure and Applied Math Series, Wiley-Interscience.

4 Hirsch M. W., C. C. Pugh, and M. Shub. Invariant Manifolds: Lecture Notes in Mathematics # 583. Springer-Verlag. Berlin, Heidelberg, New York, 1977.

5 Kato T. Perturbation Theory for Linear Operators. Springer-Verlag. New York, 1966. Die Grundlehren Der Mathematischen Wissen Schaffen In Ein Zeldarsellungen, vol., 132.

6 Mather J. Appendix I, pp 792-5, on the persistence of Anosov-diffeomorphisms under perturbations in S. Smale's paper: Differentiable Dynamical Systems. Bulletin American Math Society. 73 (1967) pp 747-817.

7 Nirenberg L. Topics in Nonlinear Functional Analysis: Lecture Notes of the Courant Institutes of Mathematical Sciences. 251 Mercier Street, New York, New York 10012.

8 Reid W. T., Ordinary Differential Equations. John Wiley and Sons, Inc., 1971. New York, London, Sydney Toronto.

Chapter 2.

Factorizations, A priori representations, Newton-Puiseux diagrams and polygons, and Multi-dimensional Newton-Puiseux Indicial sets.

1. Introduction. Traditionally, two dimensional Newton-Puiseux diagrams and polygons have been employed to determine rational exponents $\mu = p/q$ in fractional power series expansions, valid as $x \to 0$ or as $x \to \infty$, of solutions

(1) $$y(x) = \Sigma_{j\geq 0}\, a_j x^{j\mu}$$

to analytic or algebraic equations

(2) $$0 = E(x,y) = \Sigma_{j\geq 0}\Sigma_{k\geq 0}\, E_{jk}x^j y^k = \Sigma_{j\geq 0}\Sigma_{k\geq 0}\, \frac{1}{(j!)(k!)} \cdot D_x^j D_y^k E(0,0)x^j y^k$$

and to analytic or algebraic differential equations of the form

(3) $$Q(x,y)\frac{dy}{dx} = P(x,y).$$

See for instance the books B.A. FUCHS and V.I. LEVIN(10) (Russian 1949, English translation 1961), E.L. INCE(15) (1925), R.J. WALKER(28) (1949). In the power series expansion (1) the substitution $x = s^q$ yields a parametric representation

$$(x,y) = (s^q, \Sigma_{j\geq 0}\, s^{jp} \ )$$

of a curve of solutions of equation (2) or (3) which issues at $s = 0$ from a branch point at the origin (0,0) or at infinity depending on the signs of p and q with say $q > 0$ iff $x \to 0$.

In O. PERRON(20) (1961) another application of Newton-Puiseux diagrams is given for m^{th} order, analytic, scalar, ordinary differential equations not of the form in equation (3). The role of Newton-Puiseux diagrams in the analysis of asymptotic behaviour and of branching of solutions to equations $E(x,y) = 0$ has been investigated by many, recently including J. DIEUDONNE(8) (1949), R.G. BARTLE(1) (1953), L.M. GRAVES(12) (1955), D. SATHER(22) (1970) and (23) (1973), and M.M. VAINBERG and V.A. TRENOGIN in (27) (1960). In J. DIEUDONNE(8), $E(x,y)$ is permitted to possess an exponential-logarithmic expansion at $(x,y) = (0,0) \in \mathbb{R}^2$ of a type defined in G.H. HARDY(13). In R.G. BARTLE(1), L.M. GRAVES(12) and D. SATHER(22),and (23) the variable y in $E(x,y) = 0$ is allowed to be an m-dimensional element of $\mathbb{R}^m$ or $\mathbb{C}^m$. In this multi-dimensional case $(m>1)$ the coefficients $E_{jk}x^j$ of y^k in the Taylor expansion in equation (2) of $E(x,y)$ at the origin (0,0) of $\mathbb{R}^2$ or C^2 are k-linear functionals on $\mathbb{R}^m$ or C^m, respectively. The article (23) is a survey which compares these real and complex cases for the equation $E(x,y) = 0$. The textbook,review monograph (27) by M.M. VAINBERG and V.A. TRENOGIN is oriented towards the branching of solutions of nonlinear integral equations depending on one parameter, and it considers after treating Newton-Puiseux diagrams, related topics in divisibility theory.

In brief for the Taylor series expansion of $E(x,y)$ in expression (2) above, the Newton-Puiseux diagram is the boundary of the convex hull in $\mathbb{R}^2$ of the indical set $I_E = \{(j,k) : E_{jk} \neq 0\}$. This boundary is a Newton-Puiseux polygon when E is a polynomial. The rational slopes of the sides or support lines to the Newton-Puiseux diagram or polygon provide admissible values of the negative reciprocal $(-1/\mu)$ of the exponent μ

in the fractional power series expansion (1). For variables (x,y) in $\mathbb{R}^2$ each continuous curve of solutions (x,y(x)) of E(x,y) = 0 with y(0) = 0 is shown in J.DIEUDONNE(8) to satisfy

$$0 = \lim_{x\to 0} y(x)x^{-\mu}$$

for some exponent $\mu > 0$ whose negative reciprocal $(-1/\mu)$ is the rational slope to a support line of the Newton-Puiseux diagram, containing at least two points of the indicial set I_E.

In this chapter for equations

$$0 = E(x) = E(x_1, \dots x_n)$$

involving a Fréchet differentiable function $E(x_1, \dots x_n)$ of n-Banach space variables $x = (x_1, \dots x_n)$ in a product Banach space $X = \Pi_{j=1}^n X_j$, the n-dimensional Newton-Puiseux indicial set of the Taylor expansion of E at x = 0 is defined as the set I_E of multi-indices $\alpha = (\alpha_1, \dots \alpha_n)$ with non-negative integer components, for which the corresponding Fréchet derivative

$$D^\alpha E(0) = (\Pi_{j=1}^n D_{x_j}^{\alpha_j}) E(0) = D_{x_1}^{\alpha_1} D_{x_2}^{\alpha_2} \dots D_{x_n}^{\alpha_n} E(0)$$

at x = 0 exists independently of the order of differentiation, and does not vanish. In brief

$$I_E = \{ \alpha \geq 0 : D^\alpha E(0) \neq 0 \}.$$

This set is finite if E(x) is a polynomial in $x = (x_1, \dots, x_n)$. Normal vectors $p = (p_1, \dots, p_n)$ to supporting hyperplanes in $\mathbb{R}^n$ of the indicial set I_E, for which $\alpha \in I_E$ implies $\langle \alpha, p \rangle_{\mathbb{R}^n} > r$ for some r in $\mathbb{R}$, provide admissible values of exponents p_j in a priori representations

$$x_p(s,z) = (s^{p_1} z_1, \dots, s^{p_n} z_n) = (s^{p_j} z_j)_{1\leq j\leq n} \quad (s>0; z = (z_1, \dots, z_n) \in X)$$

of solutions to $E(x) = 0$. The variable s in the a priori representation $x_p(s,z)$ is allowed to assume non-positive values in $\mathbb{R}$ (or $\mathbb{C}$) if all the exponents p_j are integers and X is a real (respectively, complex) Banach space. For functions E(x) analytic at the origin 0 of X, the admissible values of $p = (p_1, \ldots, p_n)$ permit a factorization

$$E(x_p(s,z)) = s^r F_{p,r}(s,z)$$

in which $F_{p,r}(s,z)$ is a continoūs or differentiable function with

$$F_{p,r}(0,z) = \Sigma_{\substack{\alpha \\ <\alpha,p> \geq 0}} \frac{D^\alpha E(0) z^\alpha}{(\alpha!)} \qquad (\alpha! = (\alpha_1!)(\alpha_2!)\cdots(\alpha_n!)\ ).$$

This factorization is justified by the local absolute convergence of the Taylor expansion $\Sigma_{\alpha\geq 0} \frac{D^\alpha E(0)}{\alpha!} x^\alpha$, and a subsequent re-arrangement when $x = x_p(s,z)$ belongs to domain(E). The most obvious property of this factorization is that $F_{p,r}(s,z) = 0$ implies $E(x_p(s,z)) = 0$ whenever $r \geq 0$ or $s \neq 0$. Hence for a fixed r and p, the satisfaction of the hypothesis of an implicit function theorem at a root $(0,z^0)$ of $F_{p,r}(0,z) = 0$ supplies sufficient conditions for an s-dependent curve of solutions

$$x_p(s,z(s)) = (s^{p_1} z_1(s), \ldots, s^{p_n} z_n(s)), \qquad z(0) = z^0$$

to issue at $s = 0$ from a branch point of $E(x) = 0$. This branch point is at the origin if all the p_j are non-negative with $z_j^0 = 0$ if $p_j = 0$, and it is at "infinity" if for some $p_j < 0$, $z_j^0 \neq 0$ or

$$\lim_{s\to 0} \| s^{p_j} z_j(s) \|_{X_j} = \infty.$$

Further details and justifications in this analytic situation will be given in section 2 below.

For differentiable and not necessarily analytic functions similiar results follow in section 3 from a generalization of Taylor's formula

$$(4) \qquad E(x) = \Big(\sum_{\alpha \in I \setminus \partial I} \frac{D^\alpha E(0) x^\alpha}{\alpha!}\Big) + R_{\partial I}(x)$$

in which the remainder has an integral representation

$$R_{\partial I}(x) = \sum_{\alpha \in \partial I} \int_0^1 C_I^\alpha \cdot \frac{(1-t)^{|\alpha|-1}}{(|\alpha|-1)!} D^\alpha E(tx) x^\alpha dt \qquad (|\alpha| = \Sigma_{j=1}^n \alpha_j),$$

involving partial Fréchet derivatives of different orders $|\alpha|$ and constants C_I^α with combinatorial significiance. In particular factorization is possible if for some $p = (p_1, \ldots, p_n)$ and some $r > 0$, $D^\alpha E(0) = 0$ when $\alpha \in I \setminus \partial I$ and $\langle\alpha, p\rangle > r$ when $\alpha \in \partial I$. In this case, the first summation in (4) vanishes and the remainder term is $O(s^r)$ when $x = x_p(s,z)$ since $\alpha \in \partial I$ implies $D^\alpha E(tx_p(s,z))(x_p(s,z))^\alpha = s^{\langle\alpha,p\rangle} D^\alpha E(tx_p(s,z)) z^\alpha = O(s^r)$. Derivation of the extension (4) of Taylor's formula is given in section 4. Here I is a set of multi-indices with certain properties to be specified.

In section 5 an attempt is made to indicate the optimal or desirable properties of factorizations corresponding to different admissible selections of the exponents p_j.

Finally in section 6 some subjects related to Newton-Puiseux diagrams and their associated factorizations are mentioned.

2. Factorization of Analytic Functions.

Let $\mathbb{K}$ denote either one of the fields of real or complex numbers. Let $X = \Pi_{j=1}^{n} X_j$ be a product of Banach spaces X_j over $\mathbb{K}$. Let Y be another Banach space over $\mathbb{K}$.

Suppose $E : X \to Y$ has a Taylor series expansion

$$(1) \qquad E(x) = E(x_1, \dots, x_n) = \Sigma_{\alpha \geq 0} E_\alpha x^\alpha = \Sigma_{\alpha\geq 0} \frac{D^\alpha E(0) x^\alpha}{\alpha!} \neq 0$$

which converges absolutely on an open convex neighbourhood X_0 of the origin 0 in X. Put domain$(E) = X_0$.

Definition 2.1. The n-dimensional Newton-Puiseux Indical Set of the analytic map $E : X_0 \to Y$ at the origin 0 in X is the set of lattice points (multi-indices) with non-negative integer components in $\mathbb{R}^n$ given by

$$I_E = \{ \alpha = (\alpha_1, \dots, \alpha_n) : D^\alpha E(0) \neq 0 \}.$$

Observe I_E is finite or countable. It has finite cardinality iff E is a Y-valued polynomial in $x = (x_1, \dots, x_n)$.

For $p = (p_1, \dots, p_n)$ in $\mathbb{R}^n$, let

$$(2) \qquad S_p \doteq \begin{cases} \mathbb{K} & \text{if all the } p_j \text{ are non-negative integers,} \\ \mathbb{K}\setminus\{0\} & \text{if the } p_j \text{ are integers with some being negative,} \\ \mathbb{R}_+ & \text{if the } p_j \text{ are non-negative, but not all integers,} \\ \mathbb{R}_+\setminus\{0\} & \text{if none of the above i.e. otherwise.} \end{cases}$$

Then for $z = (z_1, \dots, z_n)$ in X and s in S_p set

$$x_p(s,z) = (s^{p_1} z_1, \dots, s^{p_n} z_n) = (s^{p_j} z_j)_{1\leq j\leq n} \in X.$$

(Note the notation S_p will be employed in the theorems below.)

Choose a normal vector $p = (p_1, \dots, p_n)$ in $\mathbb{R}^n$ to a hyperplane $\langle\alpha,p\rangle = r$ such that $\alpha \in I_E$ implies $\langle\alpha,p\rangle \geq r$. If equality holds for some α in I_E then $\langle\alpha,p\rangle = r$ is a supporting hyperplane to I_E. Further, if equality holds in $\langle\alpha,p\rangle \geq r$ for some α in I_E and I_E is not contained in the supporting hyperplane $\langle\alpha,p\rangle = r$ then, by definition, p is an inward pointing normal to I_E - or rather its convex hull

$$\{ \Sigma_{\alpha\in I_E} t_\alpha \alpha : \Sigma_{\alpha\in I_E} t_\alpha = 1 \text{ and } t_\alpha \geq 0 \text{ for all } \alpha \text{ in } I_E \} \subset \mathbb{R}^n_+.$$

As was previously noted when $n = 2$, the boundary of this convex hull is the Newton-Puiseux diagram or polygon of the Taylor expansion of E at 0.

For the above choice of $p = (p_1, \dots, p_n)$, the indicial set I_E is a union of disjoint subsets

$$I_j = \{ \alpha\in I_E : \langle\alpha,p\rangle = r_j \}$$

corresponding to a finite or countable sequence of real numbers $r_j \geq r$. Hence, for s in S_p and $x_p(s,z) = (s^{p_j} z_j)_{1\leq j\leq n}$ in the region X_o of absolute convergence of the Taylor expansion in (1) of E at $x = 0$,

$$\begin{aligned} E(x_p(s,z)) &= E(s^{p_1} z_1, \dots, s^{p_n} z_n) \\ &= \Sigma_{\alpha\in I_E} E_\alpha z^\alpha s^{\langle\alpha,p\rangle} \\ &= \Sigma_{j\geq 1}(\Sigma_{\alpha\in I_j} E_\alpha z^\alpha s^{\langle\alpha,p\rangle}) \\ &= \Sigma_{j\geq 1}(\Sigma_{\alpha\in I_j} E_\alpha z^\alpha)\, s^{r_j} \\ &= s^r F_{p,r}(s,z) \end{aligned}$$

where

(3) $$F_{p,r}(s,z) = \Sigma_{j\geq 1}(\Sigma_{\alpha\in I_j} E_\alpha z^\alpha) s^{(r_j - r)}$$

since $\alpha \in I_j$ implies $\langle\alpha,p\rangle \geq r_j$. The re-arrangement of the series in the preceeding calculation is justified by the absolute convergence of $\Sigma_{\alpha\epsilon I_E} E_\alpha z^\alpha s^{<\alpha,p>}$. The expression in (3) for $F_{p,r}(s,z)$ also converges absolutely for $z = (z_1, \dots, z_n)$ in X_o and s in $S_p \cup \{0\}$ with $|s| < 1$, by comparison with $\Sigma_{\alpha\geq 0}|E_\alpha z^\alpha|$ as $\Sigma_\alpha E_\alpha z^\alpha$ converges absolutely for z in X_o.

The above argument and some of its immediate consequences are summarized in the next two theorems.

Theorem 1. On Factorization of Analytic Maps.

Let $E:X=\Pi_{j=1}^n X_j \to Y$ be analytic at the origin of X. Let I_E be the Newton-Puiseux Indical Set for the Taylor expansion $\Sigma_{\alpha\geq 0} E_\alpha x^\alpha$ of E at 0 (Here $E_\alpha = (D^\alpha E(0)/\alpha!) \neq 0$ iff $\alpha \in I_E$.) Suppose for $p = (p_1, \dots, p_n)$ in $\mathbb{R}^n$ and r in $\mathbb{R}$, α in I_E implies $<\alpha,p> \geq r$, Then there is a function

(4) $$F_{p,r}(s,z) = \Sigma_{\alpha\epsilon I_E} E_\alpha z^\alpha s^{<\alpha,p>-r}, \quad z = (z_1, \dots, z_n)$$

defined on a neighbourhood of (0,0) in $(\{0\} \cup S_p) \times X$ for which the expansion on the right hand side of (4) converges absolutely, and such that

$$(s,x_p(s,z)) = (s,(s^{p_1} z_1, \dots, s^{p_n} z_n)) \in S_p \times X_0$$

implies a factorization

(5) $$E(x_p(s,z)) = s^r F_{p,r}(s,z).$$

The function $F_{p,r}(0,z) = \Sigma_{\alpha\epsilon I_E, <\alpha,p>=r} E_\alpha z^\alpha$ is not identically zero if there is an α such that $D^\alpha E(0)=(\alpha!)E_\alpha \neq 0$ and $<\alpha,p>=r$. If the exponents

p_j are integers, the map $F_{p,r}$: $\mathbb{K} \times K \to Y$ is analytic on its domain. Otherwise, it is a continuous function of (s,z) in $\mathbb{R}_+ \times X_s$, is analytic in z when s is fixed, is continuously differentiable with respect to z and is infinitely Fréchet differentiable with respect to (s,z) when s > 0.

For integer values of the exponents p_j in the a priori representation $x_p(s,z)$ the proof of thereom 1 follows from the discussion preceeding it. For non-integer values of these exponents, see the following digression on "hybrid" Banach spaces, but note on first reading it may be best to restrict the exponents p_j to integer values only. The statement of theorem 2 after the digression is a consequence of the analytic factorization result in theorem 1 and the ordinary implicit function theorem for Banach spaces, in say reference (19) of the biblography, when the p_j are integers.

On Hybrid Banach Spaces. (Digression.)

When $\mathbb{K}$ represents the complex numbers and the exponents p_j are not all integers, combinatorial problems associated with defining the domains of the functions $s \to s^{p_j}$ in the a priori representation $x_p(s,z) = (s^{p_j} z_j)$ and of the maps $s \to s^{\langle\alpha,p\rangle - r)}$ and $s \to s^{\langle\alpha,p\rangle}$ in the expansions of $E(x_p(s,z))$ and $F_{p,r}(s,z)$, are avoided by restricting the variable s to positive or non-negative real values. In this situation there results a Hybrid Banach space $\mathbb{R} \times X \supset (\{0\} \cup S_p) \times X$ in which the first factor $\mathbb{R}$ is a real Banach space and the second factor X is a complex Banach space[1]. The partial s-Fréchet derivative of $F_{p,r}(s,z)$ is real-linear in $\Delta s \in \mathbb{R}$, while the partial z- (or x-) Fréchet derivative of $F_{p,r}(s,z)$ is complex-linear in $\Delta z \in X$. From Cauchy formulas for analytic functions, $F_{p,r}(s,z)$ is infinitely and continuously Fréchet differentiable with respect to z in X for (s,z) in the interior of its domain in $\mathbb{R} \times X$. From these same formulas, $F_{p,r}(s,z)$ and its partial derivatives $D_z^j F_{p,r}(s,z)$ are infinitely and continuously differentiable with respect to s when $s > 0$. The latter follows from the (locally) uniform convergence of the defining expression (4) for $F_{p,r}(s,z)$. The remarks in this digression can be justified by methods found in references (2),(7),(16) and (19). See also Chapter 3.

[1] The norm on the Hybrid Banach space $\mathbb{R} \times X$ is the same as that given to the product Banach space $\mathbb{R} \times X$ when X is treated as a real Banach space i.e. when scalar multiplication in X is restricted to the real numbers $\mathbb{R} \subset \mathbb{K}$.

Theorem 2. On the Existence of Branches of Solutions.

Let E, X, Y, p, r and $F_{p,r}(s,z)$ be as in Theorem 1. Suppose $z^0 = (z_j^0)_{1\le j\le n}$ in X is a root of the equation

$$0 = F_{p,r}(0,z) = \Sigma_{\alpha\ge 0,\ \langle\alpha,p\rangle = r} \frac{D^\alpha E(0) z^\alpha}{\alpha!}$$

Suppose $X = U + V$ where U and V are subspaces of X with V closed and such that the partial v-Fréchet derivative ($v \in V$), denoted by

$$D_v F_{p,r}(0,z) = D_z F_{p,r}(0,z)|_V : V \to Y,$$

in the direction of the V-subspace is surjective and bounded below at $z^0 = u^0 + v^0$. Then

A) There are continuous functions $v(s,u)$, $z(s,u) = u + v(s,u)$ and $x(s;u) = x_p(s,u) = (s^{p_j} z_j(s,u))_{1\le j\le n}$ defined on a neighbourhood A_1 of $(0,u^0)$ in $(\{0\}\times S_p)\times U$ with values in V, X and X, respectively, which satisfy

$$v(0,u^0) = 0,\ z(0,u^0) = u^0 + v^0 = z^0,$$

$$F_{p,r}(s,\ u + v(s,u)) = 0$$

and

$$E(x(s,u)) = s^r F_{p,r}(s,\ u + v(s,u)) = 0 \quad (\text{if } s \ne 0 \text{ or } r > 0)$$

for all (s,u) in the common domain A_1.

B) $x(0,u^0) = 0$ if all the p_j are non-negative and $z_j(0,u^0) = 0$ when $p_j = 0$.

C) If all the p_j's are integers then $v(s,u)$, $z(s,u)$ and $x(s,u)$ are analytic functions of (s,u) in a neighbourhood of $(0,u^0)$ in $\mathbb{K}\times U$, with $x(s,u)$ possessing a pole of order at most $s^{\min(p_j : 1\le j\le n)}$ at $s = 0$ when u is fixed. If the p_j are not all integers, these functions are analytic in u for

$s = 0$, and infinitely and continuously differentiable with respect to $s > 0$ together with all their higher order partial u-Fréchet derivatives.

D) There is a neighbourhood A_2 of $(0,u^0,v^0)$ in $(\{0\} \times S_p) \times U \times V$ such that all solutions (s,u,v) in A_2 of $F_{p,r}(s,u+v) = 0$ must satisfy $v = v(s,u)$ for some (s,u) in A_1. Moreover it follows (from this) that every non-zero solution x of $E(x) = 0$ in the set

$$x_p(A_2) = \{x_p(s,u+v) : (s,u,v) \in A_2\}$$

must satisfy $x = x(s,u)$ for some (s,u) in A_1 with $s \neq 0$.

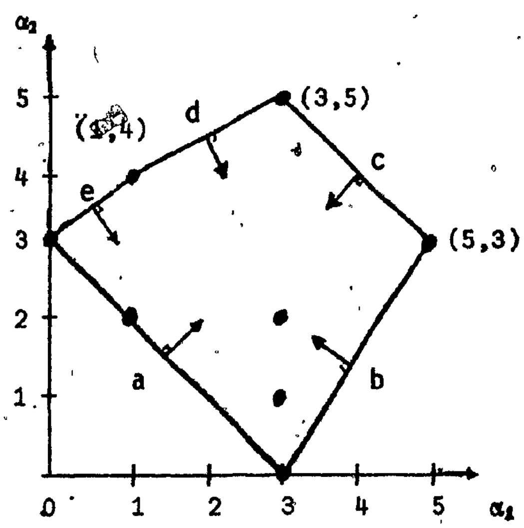
Proof: First consider the part of the conclusions as they apply to the equation $F_{p,r}(s,u+v) = 0$. This part follows immediately from implicit function theorems and results on the differentiablity of implicitly defined functions in references (2),(7),(16) and (19). From (16), differentiability properties of the implicitly defined function $v(s,u)$ correspond at a point $(\hat{s},\hat{u})$ in its domain A to those of $F_{p,r}(s,u+v)$ in a neighbourhood of $(\hat{s},\hat{u},v(\hat{s},\hat{u}))$ in its domain. The rest of the above assertion should be self-evident.

Q.E.D.

<u>Example 1.</u> Let (x,y) be two real or two complex variables. Put

$$E(x,y) = x^0y^3 + x(y^2 + y^4) + x^3(1 + y + y^2 + y^5) - x^5y^3.$$

The Newton-Puiseux polygon for E is indicated in Figure 1.



<u>Figure 1.</u> Newton-Puiseux Polygon for $E(x_1,x_2) = 0$.

Inward pointing normals to the sides a,b,c,d and e are respectively $(1,1)$, $(-3,2)$, $(-1,1)$, $(2,-1)$ and $(1,-1)$. For side a

$$E(sx,sy) = s^3(x^3 + y^3) + s^4x^3y + s^5(x^3y^2 + xy^4) + s^8(x^5y^3 + x^3y^5)$$

Thus

$$F(s,x,y) = E(sx,sy) \div s^3 = x^3 + y^3 + sx^3y + s^2(x^3y^2 + xy^4) + s^5(x^5y^3 + x^3y^5)$$

For $j = 0,1,2$, $(x^0, y_j^0) = (1, \exp \frac{(2j+1)\pi i}{3})$ is a root of

$$0 = F(0,x,y) = x^3 + y^3.$$

From $F_y(0,1,y_j^0) = 3(y_j^0)^2 \neq 0$, the implicit function theorem implies there are analytic functions $y_j(s)$ such that $F(s,1,y_j(s)) = 0$ and $y_j(0) = y_j^0$ for

$j = 0$, 1 or 2. From

$$\frac{dy_j}{ds}(0) = -\frac{F_s(0,1,y_j^0)}{F_y(0,1,y_j^0)} = -\frac{+1^3 y_j^0}{3\bar{y}_j^0} = -\frac{1}{3},$$

the first order expansion of $y_j(s)$ is

$$y_j(s) = y_j^0 - \frac{1}{3}s + O(s^2).$$

Therefore,

$$(s, sy_j(s)) = (s, s\exp(\frac{(2j+1)\pi i}{3}) - \frac{1}{3}s^2 + O(s^3))$$

is the second order asymptotic expansion of three complex curves of solutions corresponding to $j = 0$, 1 and 2, passing through the origin at $s = 0$ in $\mathbb{C}$. Note only the branch given by $j = 1$ is real when s is restricted to real values.

Example 2. Let $f(x)$ be an analytic function of a variable x in $\mathbb{K}$ with an m^{th} order zero at $x = 0$ in $\mathbb{K}$. Thus $f(x) = x^m h(x)$ for some analytic function h with $h(0) \neq 0$. Put $E(x,y) = y - x^m h(x)$. Then

$$I_E = \{(j,k) : D_x^j D_y^k E(0) \neq 0\} \subset \{(0,1)\} \cup \{(j,0) : j \geq m\}.$$

A superset of I_E and a support line (side a) to both I_E and the superset are indicated in figure 2, below. Generically for analytic functions f with

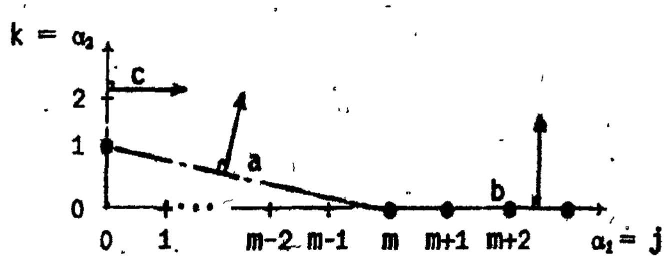


Figure 2. Generic Newton-Puiseux diagram for $E(x_1,x_2) = x_2 - x_1^m h(x_1)$ at $(0,0)$.

m^{th} order zeroes at 0, in a normed vector space of maps defined on an open ball B in the complex plane with $\|f\| = \sup\{|f(x)| : x \in B\}$, the above superset of I_E equals I_E. In this generic case the vectors (1,m), (0,1) and (1,0) are normals to the sides a, b and c, respectively, of the Newton-Puiseux diagram in figure 2 which point into the convex hull of I_E. For the a priori representation $(sx, s^m y)$ corresponding to the vector (1,m),

$$E(sx, s^m y) = s^m F(s,x,y)$$

where

$$F(s,x,y) = y - x^m h(sx).$$

Let $x_j^o = \exp(2\pi ij/m)$, $1 \leq j \leq m$, represent the m m^{th} roots of unity. Then

$$F(0, x_j^o, h(0)) = 0$$

and

$$F_x(0, x_j^o, h(0)) = m(x_j^o)^{m-1} h(0) = \frac{mh(0)}{x_j^o} \neq 0.$$

Therefore the implicit function theorem implies the existence of m curves $(sx_j(s), s^m h(0))$ satisfying the initial conditions $x_j(0) = x_j^o$ and

$$E(sx_j(s), s^m h(0)) = s^m F(s, x_j(s), h(0)) = 0$$

or

$$s^m h(0) = (x_j(s))^m h(x_j(s)) = f(x_j(s)).$$

for $1 \leq j \leq m$. Here in the real case $\mathbb{K} = \mathbb{R}$, at most two of the curves corresponding to $j = \frac{1}{2}m$ (m even) and $j = m$ (m even or odd) yield real solution branches $(sx_j(s), s^m h(0))$.

3. Factorization of Differentiable Functions. Extension of Taylor's Formula.

Let $X = \Pi_{j=1}^{n} X_j$ and Y be Banach spaces over $\mathbb{K}$. Let $E : X \to Y$ be a continuous map defined on a convex neighbourhood X_0 of the origin in X.

Definition 3.1. Newton-Puiseux Indical Set for Differentiable Functions. Let I be a collection of lattice points (multi-indices) α in $\mathbb{R}_+^n$ for which $0 < \beta < \alpha$ implies $\beta \in I$ when $\alpha \in I$, and for which

$$D^{\alpha}E(x) = (\Pi_{j=1}^{n} D_{x_j}^{\alpha_j})E(x)$$

exists independently of the order of partial Frêchet differentiation, and is continuous on the convex neighbourhood X_0 of the origin 0 in X. Then the n-dimensional Newton-Puiseux Indical Set in I for E at $x = 0$ is

$$I_E = \{ \alpha \in I : D^{\alpha}E(0) = 0 \}.$$

Remark. The collection of functions $E : X_0 \to Y$ which satisfy the smoothness properties indicated in definition 3.1. form a smoothness class closed under scalar multiplication and point-wise addition.

<u>Theorem 3. An Extension of Taylor's Formula.</u> Let $X = \Pi_{j=1}^{n} X_j$ and Y be Banach spaces over $\mathbb{K}$ $(n > 1.)$ Let $p = (p_1, \ldots, p_n)$ with positive components p_j. Let $r > 0$. Let e_j be the j^{th} element of a standard basis of $\mathbb{R}^n$ with 1 as its j^{th} component and zeroes elsewhere. Define sets of multi-indices by

$$I = \{\alpha > 0 : (\langle\alpha,p\rangle < r) \text{ or } \exists j (1 \le j \le n) \text{ such that } \alpha - e_j > 0 \text{ and } \langle\alpha - e_j, p\rangle \ge 0\},$$

$$\partial I = \{\alpha \in I : \langle\alpha,p\rangle > r\},$$

$$\text{Int}(I) = I \setminus \partial I = \{\alpha \in I : \langle\alpha,p\rangle < r\}.$$

Suppose $E : X \to Y$ is a function defined and continuous together with its Fréchet derivatives $D^\alpha E(x)$ when $\alpha \in I$, on a convex neighbourhood X_0 of the origin 0 in X. Then for x in X_0

$$E(x) = \left(\sum_{\alpha \in I \setminus \partial I} \frac{D^\alpha E(0) x^\alpha}{\alpha!}\right) + R_{\partial I}(x)$$

where the remainder term

$$R_{\partial I}(x) = \sum_{\alpha \in \partial I} \int_0^1 C_I^\alpha \frac{(1-t)^{|\alpha|-1}}{(|\alpha|-1)!} . D^\alpha E(tx) x^\alpha dt.$$

Note for future convenience

$$I = \{\alpha + \beta : \alpha \in \text{Int}(I) \text{ and } |\beta| = 1\} \cup \{0\}$$

and

$$\partial I = \{\alpha \in I : \exists e_j \text{ with } \alpha - e_j \in \text{Int}(I)\}.$$

The proof of theorem 3 is given in section 4. The following corollaries of the above extension of Taylor's formula provide differentiable analogues of the analytic factorization result in theorem 1.

Corollary 3.1. Factorization of Differentiable Functions I. Suppose $\alpha \in \mathrm{Int}(I)$ implies $D^{\alpha}E(0) = 0$. Then

$$E(x_p(s,z)) = s^r F_{p,r}(s,z)$$

where for $x_p(s,z) = (s^{p_j} z_j)_{1 \le j \le n}$ in X_0 and s in $S_p \cup \{0\}$

$$F_{p,r}(s,z) = \sum_{\alpha \in \partial I} \left(C_I^{\alpha} \int_0^1 \frac{(1-t)^{|\alpha|-1}}{(|\alpha|-1)!} D^{\alpha}E(tx_p(s,z))z^{\alpha}dt\right)s^{<\alpha,p>-r} .$$

(Note: $\alpha \in \partial I$ implies $<\alpha,p> \geq r$.)

Proof: In theorem 3, only the remainder term appears and

$$D^{\alpha}E(x_p(s,z))\,(x_p(s,z))^{\alpha} = (D^{\alpha}E(x_p(s,z))z^{\alpha})s^{<\alpha,p>} .$$

Q.E.D.

In the next corollary 3.2. identify $\mathbb{R}^n$ with the subspace $\mathbb{R}^n \times \{0\}$ of $\mathbb{R}^{n+1}$ Thus in particular "$p = (p_1, \dots, p_n) = (p_1, \dots, p_n, 0)$" and the elements α of I, $\mathrm{Int}(I)$ and ∂I are identified with the corresponding elements $(\alpha,0)$ in the subsets $I \times \{0\}$, $\mathrm{Int}(I) \times \{0\}$ and $\partial I \times \{0\}$ in $\mathbb{R}^{n+1}$.

Corollary 3.2. Factorization of Differentiable Functions II. Let $X = \Pi_{j=1}^{n+1} X_j$ and Y be Banach spaces. Let $E : X \to Y$ be a function defined and continuous, together with its Fréchet derivatives $D^{(\alpha,0)}E(x)$ when $\alpha \in I$, on a convex neighbourhood X_0 of the origin in X. Further suppose $\alpha \in \mathrm{Int}(I)$ implies $D^{(\alpha,0)}E(0,x_{n+1}) = 0$ for all $(0,x_{n+1})$ in X_0. Then

$$E(x_{(p,0)}(s,z)) = s^r F_{(p,0),r}(s,z)$$

where for $x_{(p,0)}(s,z) = ((s^{p_j} z_j)_{1 \le j \le n}, z_{n+1})$ in X_0 and s in $S_{(p,0)} = S_p$,

$$F_{(p,0),r}(s,z) = \sum_{\alpha \in \partial I} C_I^{\alpha} \int_0^1 \frac{(1-t)^{|\alpha|-1}}{(|\alpha|-1)!} . D^{(\alpha,0)}E(tx_{(p,0)}(s,z))z^{(\alpha,0)}dt\ s^{<\alpha,p>-r} .$$

<u>Proof</u>: For fixed x_{n+1}, apply corollary 3.1.

Q.E.D.

Note the differentiability properties of the functions $F_{p,r}$ and $F_{(p,0),r}$ correspond at $s = 0$ to those of the Fréchet derivatives $D^{\alpha}E(x)$ appearing in their definitions in corollaries 3.1. and 3.2. , while for $s \neq 0$ they correspond to those of E. These differentiability properties are in part inherited by functions implicitly defined by the equations

$$0 = F_{p,r}(s,z)$$

and

$$0 = F_{(p,0),r}(s,z).$$

The statement of branch-point results for differentiable functions, analogous to the branch-point result in theorem 2 for analytic functions, and based on the factorizations in corollaries 3.1. and 3.2. , and based on implicit function theorems will be omitted.

In section 4 below, the proof of the extension in theorem 3 of Taylor's formula is given in the first subsection 4.A. In subsection 4.B , there is a digression on the evaluation of the constants C_I^{α} appearing in theorem 3. The constants C_I^{α} depend on p and r. An example in which these constants are determined is given in subsection 4.C. In subsection 4.D , there is another digression. This second digression concerns a further generalization of Taylor's formula. The effect of different choices of p and r in $x_p(s,z)$ and in associated factorizations is discussed in section 5.

4. A. Proof of Theorem 3.

First suppose $D^{\alpha}E(0) = 0$ for all α in $\mathrm{Int}(I) = I\backslash\partial I$. The general case, as will be shown below, follows from this special one.

By induction on $m \geq 1$, there are integers C_I^{α} such that

$$(m) \quad E(x) = \Big(\sum_{\substack{\alpha \in \partial I \\ 1 \leq \alpha \leq m}} + \sum_{\substack{\alpha \in \mathrm{Int}(I) \\ |\alpha| = m}}\Big) C_I^{\alpha} \int_0^1 \frac{(1-t)^{\alpha-1}}{(|\alpha|-1)!} . D^{\alpha}E(tx)x^{\alpha}dt$$

whenever (i) $E : X \to Y$ is a map defined and continuous together with all its Fréchet derivatives $D^{\alpha}E(x)$ $(\alpha \in I)$ on a convex neighbourhood X_0 of the origin 0 in X; and (ii) $D^{\alpha}E(x) = 0$ for all α in Int(I). The inductive proof of (m) is as follows.

For $m = 1$, statement (m) is true with $C_I^{\alpha} = 1$ since $|\alpha| = 1$ implies $\alpha \in I$, and since

$$E(x) = E(0) + \int_0^1 \Big(\frac{d}{dt} E(tx)\Big)dt = \sum_{\alpha \in I, |\alpha| = 1} \int_0^1 D^{\alpha}E(tx)x^{\alpha}dt.$$

Now as an inductive hypothesis, assume statement (m) holds for some $m > 1$. For γ in Int(I) with $|\gamma| = m$, integration by parts implies

$$\int_0^1 \frac{(1-t)^{|\gamma|-1}}{(|\gamma|-1)!} . D^{\gamma}E(tx)x^{\gamma}dt = \int_0^1 \frac{(1-t)^{|\gamma|}}{|\gamma|!} . \frac{d}{dt}(D^{\gamma}E(tx)x^{\gamma})dt$$

$$= \sum_{\substack{\beta \\ \alpha = \gamma+\beta, |\beta| = 1}} \int_0^1 \frac{(1-t)^{|\alpha|-1}}{(|\alpha|-1)!} . D^{\alpha}E(tx)x^{\alpha}dt.$$

Note

$$\{\alpha + \beta : \alpha \in \mathrm{Int}(I),\ |\alpha| = m,\ |\beta| = 1\} = \{\alpha \in I : |\alpha| = m\}$$

$$= \{\alpha \in \partial I : |\alpha| = m\} \cup \{\alpha \in \mathrm{Int}(I) : |\alpha| = m\}$$

Therefore the second summation in (m) is

$$\Sigma_{\substack{\gamma \in Int(I) \\ |\gamma|=m}} C_I^\gamma \int_0^1 \frac{(1-t)^{|\gamma|-1}}{(|\gamma|-1)!} \cdot D^\gamma E(tx)x^\gamma dt$$

$$= \Sigma_{\substack{\gamma \in Int(I) \\ |\gamma|=m}} \Big(\Sigma_{\substack{\alpha=\gamma+\beta \\ |\beta|=1}} C_I^\alpha \int_0^1 \frac{(1-t)^{|\alpha|-1}}{(|\alpha|-1)!} \cdot D^\alpha E(tx)x^\alpha dt \Big)$$

$$= \Big(\Sigma_{\substack{\alpha \in \partial I \\ |\alpha|=m+1}} + \Sigma_{\substack{\alpha \in Int(I) \\ |\alpha|=m+1}}\Big) C_I^\alpha \int_0^1 \frac{(1-t)^{|\alpha|-1}}{(|\alpha|-1)!} \cdot D^\alpha E(tx)x^\alpha dt$$

where the integer coefficients C_I^α are inductively defined for $|\alpha| = m+1$ by

(1) $$C_I^\alpha = \Sigma_{(\gamma,\beta),\ \gamma \in Int(I),\ |\beta|=1,\ \alpha=\gamma+\beta,\ |\gamma|=m} (C_I^\gamma)$$

in terms of coefficients C_I^γ affiliated with lower-order multi-indices γ with $|\gamma| = m$. Substitution of the last expression for the second summation in (m) back into (m) yields the corresponding formula for $m+1$ in place of m. This substitution completes the inductive step and establishes (m) for $m > 1$.

For $m > \max\{|\alpha|: \alpha \in I\}$ the set $\{ \alpha \in Int(I) : |\alpha| = m \}$ is void, and the set $\{ \alpha \in I : |\alpha| = m \} \subseteq \partial I$. Thus the Theorem holds in the special case of $D^\alpha E(0) = 0$ for all $\alpha \in Int(I)$ since for sufficiently large m in the preceeding inductive arguement,

(2) $$E(x) = \Sigma_{\alpha \in \partial I} C_I^\alpha \int_0^1 \frac{(1-t)^{|\alpha|-1}}{(|\alpha|-1)!} \cdot D^\alpha E(tx)x^\alpha dt.$$

For the general case let E satisfy the hypotheses of the theorem without the special restrictions imposed above in the special case, and set

$$F(x) = E(x) - \Sigma_{\alpha \in Int(I)} \frac{D^\alpha E(0)x^\alpha}{\alpha!}$$

But

$$D^\beta(D^\alpha E(0)x^\alpha) = \begin{cases} 0 & \text{if } \alpha_j < \beta_j \text{ for some } j \text{ i.e. } \alpha \not\geq \beta, \\ \frac{\alpha!}{(\alpha-\beta)!} D^\alpha E(0)x^{\alpha-\beta} & \text{if } \alpha \geq \beta. \end{cases}$$

Therefore $D^\alpha F(0) = 0$ for all α in Int(I). Further $D^\alpha F(x) = D^\alpha E(x)$ for all α in ∂I because $\beta \in \mathrm{Int}(I)$ implies $<\alpha,p> \geq r > <\beta,p>$, so that for all α in ∂I and all β in Int(I) $\alpha_j > \beta_j$ for some j, and hence $D^\alpha(D^\beta E(0)x^\beta) = 0$. The general case now follows from the application of the special case formula (2) to F(x).

Q.E.D.

4.B. On the Evaluation of the C_I^α Coefficients in Theorem 3.

As can be seen from an induction arguement involving the recursive formula (1) defining the coefficients C_I^α in the above proof of theorem 3, each coefficient C_I^α is equal to the number of lattice-point sequences or paths $(\beta^j)_{1\leq j\leq|\alpha|}$ with $0 \lneq \beta^j < \beta^{j+1} \leq \alpha$ for $1 < j < |\alpha|$, which are contained in Int(I) except possibly for the end-point $\alpha = \beta^{|\alpha|}$. For α in Int(I), $0 < \beta < \alpha$ implies $\beta \in \mathrm{Int}(I)$, so that the number of the above paths $(\beta^j)_{1\leq j\leq|\alpha|}$ is the multi-nomial coefficient

(3) $$C_I^\alpha = \binom{|\alpha|}{\alpha} = \frac{|\alpha|!}{\alpha!}.$$

Formula (3) is also valid for those α in ∂I for which $0 < \beta \lneq \alpha$ implies $\beta \in \mathrm{Int}(I)$.

This evaluation of the C_I^α agrees with a result given by Taylor's formula when $r = N$ and all the $p_j = 1$. For then $I = \{\alpha \geq 0 : |\alpha| = N\}$ $\partial I = \{\alpha > 0 : |\alpha| = N\}$ and $\mathrm{Int}(I) = \{\alpha > 0 : |\alpha| = N\}$; and further by Liebnitz's rule, Taylor's formula

$$E(x) = \Sigma_{j=1}^{N-1} \frac{D^j E(0)x^j}{j!} + \int_0^1 \frac{(1-t)^{N-1}}{(N-1)!} \cdot \frac{D^N E(tx)x^N}{N!} dt$$

$$= \Sigma_{j=1}^{N-1} \Sigma_{|\alpha|=j} \frac{D^\alpha E(0)x^\alpha}{\alpha!} + \int_0^1 \Sigma_{|\alpha|=N} \binom{|\alpha|}{\alpha} \frac{(1-t)^{|\alpha|-1}}{(|\alpha|-1)!} \cdot D^\alpha E(tx)x^\alpha dt$$

$$= \Sigma_{\alpha \in Int(I)} \frac{D^\alpha E(0)x^\alpha}{\alpha!} + R_{\partial I}(x)$$

coincides with the formula in Theorem 3, as expected, because of the evaluation formula (3).

4.C. An Example in which the Coefficients C_I^α are determined.

Let $p = (2,1)$ and $r = 4$. The elements of I are indicated by the dots in Figure 3. Here $\partial I = \{(2,0),(2,1),(1,2),(1,3),(0,4)\}$ and

$$C_I^{(2,0)} = \binom{2}{2,0} = 1, \quad C_I^{(1,2)} = \binom{3}{1,2} = 3, \quad C_I^{(0,4)} = \binom{4}{0,4} = 1$$

$$C_I^{(2,1)} = C_I^{(2,0)} + C_I^{(1,1)} = \binom{2}{2,1} + \binom{2}{1,1} = 3 \quad \text{and}$$

$$C_I^{(1,3)} = C_I^{(1,2)} + C_I^{(0,3)} = \binom{3}{1,2} + \binom{3}{0,3} = 3 + 1 = 4.$$

Thus when $D^\alpha E(0) = 0$ for α in $Int(I) = I \setminus \partial I$,

$$E(x,y) = \int_0^1 \Big(\frac{(1-t)^2}{2!} D_x^2 E(tx,ty)x^2 + 3\frac{(1-t)^3}{3!} D_x^2 D_y^1 E(tx,ty)x^2 y^1 +$$

$$3\frac{(1-t)^3}{3!} D_x^1 D_y^2 E(tx,ty)x^1 y^2 + 4\frac{(1-t)^4}{4!} D_x^1 D_y^3 E(tx,ty)x^1 y^3$$

$$+ 1\frac{(1-t)^4}{4!} D_x^4 E(tx,ty)x^4 \Big) dt$$

at (x,y) in $\mathbb{R}^2$.

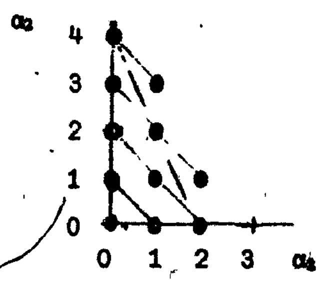


Figure 3. The set I in Theorem 3 corresponding to $p = (2,1)$ and $r = 4$.

The downward sloping 45° lines in Figure 3 passing through the points (0,1),(0,2),(0,3) and (0,4) represent the inductive steps $m = 1,2,3$ and 4 in the proof of theorem 3. The other line segment is contained in the support line $\langle \alpha, p \rangle = r$ of the Newton-Puiseux indical set I_E in I of the function E. In corollary 3.1. , only the Fréchet derivatives $D^\alpha E(0)$ corresponding to the multi-indices α in this line segment appear in the integral representation of $F_{p,r}(s,z)$ when $s = 0$.

4.D. A Further Extension of Taylor's Formula.

Let I be a collection of multi-indices α in $\mathbb{R}^n$, which is the union of two disjoint subsets Int(I) and ∂I with the properties

(i) $0 \in Int(I)$,

(ii) $\alpha \in \partial I$ iff $\alpha - e_j$ Int(I) for some j $(1 < j < n)$, and

(iii) $\alpha \in I$ implies there is a path $(\beta^j)_{1 \leq j \leq |\alpha|}$ such that $\beta^j \in Int(I)$ and $0 \lneqq \beta^j \lneqq \beta^{j+1} < \alpha$ for $1 < j < |\alpha|$ and such that $\beta^{|\alpha|} = \alpha$.

In this situation, the conclusions and proofs of theorem 3 and the corollaries 3.1. and 3.2. remain valid as well as the comments in 4.B. concerning the evaluation of the coefficients C_I^α. Figure 4. indicates a simple example of a set I with the above properties. The lines in Fig., 4 form a Newton-Puiseux diagram for a function $E(x_1, x_2)$, in the special case $I_E = \partial I$ (See the the definition 3.1 of I_E in I.)

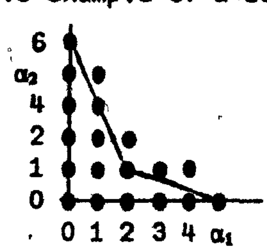


Figure 4. Elements of I are denoted by • .

5. Comparison of the a priori representations.

Let I_E be the Newton-Puiseux Indical Set of a function $E : (\Pi_{j=1}^n X_j) \to Y$. Suppose $\langle\alpha,p\rangle \geq r$ for all α in I_E with equality for some α in I_E. Assume a factorization $E(x_p(s,z)) = s^r F_{p,r}(s,z)$ in which $F_{p,r}(s,z)$ is a continuous or differentiable function. Observe.

(A) If the set $\{ \alpha \in I_E : \langle\alpha,p\rangle = r \}$ spans $\mathbb{R}^n$ then the components of p are uniquely determined as rational multiples of r.

(B) If $\varepsilon > 0$ then $\langle\varepsilon\alpha,\varepsilon p\rangle \geq \varepsilon r$ for all α in I_E. Further

$$E(x_{\varepsilon p}(s,z)) = s^{\varepsilon r} F_{p,r}(s^\varepsilon,z)$$

since $x_{\varepsilon p}(s,z) = (s^{\varepsilon p_j} z_j)_{1\leq j\leq n} = x_p(s^\varepsilon,z)$. Hence

$$F_{p,r}(s^\varepsilon,z) = F_{\varepsilon p,\varepsilon r}(s,z)$$

whenever both sides are defined. Therefore $F_{p,r}(s,z)$ satisfies the hypotheses of theorem 2 (or an non-analytic analogue) at $s = 0$ iff $F_{\varepsilon p,\varepsilon r}(s,z)$ satisfies these same conditions (Note however that S_p and $S_{\varepsilon p}$ need not be equal.) If now the hypotheses of theorem 2 are satisfied and $z_p(s,u)$ and $z_{\varepsilon p}(s,u)$ are the functions appearing in its conclusion for (p,r) and $(\varepsilon p,\varepsilon r)$ respectively, then

$$z_{\varepsilon p}(s,u) = z_p(s^\varepsilon,u)$$

and

$$x_{\varepsilon p}(s,z_{\varepsilon p}(s,u)) = x_p(s^\varepsilon,z_p(s^\varepsilon,u)).$$

whenever both sides of these equations are defined. From this it follows that the $(\varepsilon p,\varepsilon r)$ factorization yields branches obtainable with the (p,r) when the components of εp are not all integers. In particular if the components ρ_j

are rational multiples of $r \neq 0$, then there is a unique $\eta > 0$ such that all the components ηp_j of ηp and ηr are integers with greatest common divisor 1. For this η, $S_{\eta p} \cup \{0\} = \mathbb{K}$ and every branch $x_{\varepsilon p}(s, z_{\varepsilon p}(s,u))$ of solutions x obtained from a factorization associated with $(\varepsilon p, \varepsilon r)$ for some $\varepsilon > 0$, with the aid of theorem 2, can also be produced in the same manner from the $(\eta p, \eta r)$ factorization after a change of parameterization $s \to s^{(\varepsilon/\eta)}$ i.e. for all s in $S_{\varepsilon p}$, $z_{\varepsilon p}(s,u) = z_{\eta p}(s^{(\varepsilon/\eta)},u)$ and $x_{\varepsilon p}(s, z_{\varepsilon p}(s,u)) = x_{\eta p}(s^{(\varepsilon/\eta)}, z_{\eta p}(s^{(\varepsilon/\eta)},u))$. Moreover in this circumstance where all the p_j are rational multiples of r, $(\eta p, \eta r)$ is the only multiple of (p,r) such that the solution branches $x_{\eta p}(s, z_{\eta p}(s,u))$ of $E(x) = 0$ coming from the equation

$$F_{\eta p, \eta r}(s,z) = 0$$

upon the application of an implicit function theorem duplicate for any $\varepsilon > 0$ the solution branches $x_{\varepsilon p}(\hat{s}, z_{\varepsilon p}(\hat{s},u))$ coming from the equation

$$F_{\varepsilon p, \varepsilon r}(\hat{s},z) = 0$$

in the same manner, after a change of parameterization $\hat{s} = s^{(\varepsilon/\eta)}$.

(C) If $p < q \in \mathbb{R}^n$ and $\langle \alpha, q \rangle = r$ defines a supporting hyperplane to I_E then the existence of α in I_E with $\alpha_j > 0$ for some j and $\langle \alpha, q \rangle = r$ implies for such j that $q_j = p_j$ (else the contradiction $r = \langle \alpha, q \rangle > \langle \alpha, p \rangle \geq r$.) Further the relation

$$x_q(s,z) = (s^{q_j} z_j)_{1 \le j \le n} = x_p(s, (s^{q_j - p_j} z_j)_{1 \le j \le n})$$

implies

$$F_{q,r}(s,z) = s^{-r} E(x_q(s,z)) = F_{p,r}(s, (s^{q_j - p_j} z_j)_{1 \le j \le n}).$$

Therefore solution branches obtainable from the (q,r) factorization with theorem 2 are also obtainable from the (p,r) factorization. In particular note the case where $(p,r) = (0,0) < (q,r)$.

6. Related Topics.

A. A. SESTOKOV in (26) has investigated the role of indical sets and their supporting hyperplanes $\langle \alpha, p \rangle = r$ in the solution of singular first-order systems of autonomous, analytic, ordinary differential equations. For further usage of a priori representations in the analysis of singularities of differential equations see the Frobenius-Fuchs theory as described in E. L. INCE (15), E. A. CODDINGTON and N. LEVINSON (3), and B. A. FUCHS and V. I. LEVIN (10).

There should be a formal analogue of the analytic theory for the existence of branches of formal solutions to equations $E(x_1,\ldots,x_n) = 0$ in which the function E has a finite or infinite asymptotic expansion. As noted in the introduction J. DIEUDONNE in (8) has defined Newton-Puiseux diagrams for asymptotic expansions other than those given by Taylor series expansions.

J. DIEUDONNE has also shown in the same article (8) that all continuous branches of solutions to $E(x_1,x_2) = 0$ issuing from a branch point at the origin of a real-valued analytic function $E(x_1,x_2)$ of two real variables can be obtained from the negatively sloping sides of the Newton-Puiseux diagram of E at the origin of $\mathbb{R}$ (This remark was also noted in the introduction and in fact J. DIEUDONNE uniqueness result is a little stronger than that indicated here.) In other situations the question of determining all solutions in a neighbourhood of a given branch point appears to require a case by case study. In bifurcation due to simple eigenvalues and in Hopf bifurcation of periodic solutions for ordinary differential equations in finite dimensions, all solutions sufficiently near the bifurcation point in these theories lie on an initially given branch of solutions or on a second bifurcating branch of solutions whose existence is demonstrated (see for example the paper's (4), (5) and (6) of M.G. CRANDALL and P.H. RABINOWITZ.) In an extension of bifurcation due to simple eigenvalues, a generalized Morse lemma is employed in L. NIRENBERG (19) in a situation where the range of a linearized operator has co-dimension 1, to show that all solutions in a neighbourhood of a given branch point belong to one of two branches found with this lemma. The use of this Morse lemma provides a connection with the analysis of singularities and cusps in algebraic geometry and in catastrophe theory as in say M. GOLUBITSKY and V. GUILLEMIN (11) and E. C. ZEEMAN (29).

7. Bibliography for Chapter 2.

1 Bartle R. G. Singular Points of Functional Equations. American Math., Trans., vol., 75, 1953, 366-384.

2 Berger, M. S. Nonlinearity and Functional Analysis: Lectures on Nonlinear Problems in Mathematical Analysis. Pure and Applied Math., Series Academic Press, New York, 1977.

3 Coddington E. A. and N. Levinson. Theory of Ordinary Differential Equations. International Series in Pure and Applied Mathematics. McGraw-Hill Book Company, 1955.

4 Crandall M. G. and P. H. Rabinowitz. Bifurcation from Simple Eigenvalues. Journal of Functional Analysis 8, 321-340, 1971.

5 ________ The Hopf Bifurcation Theorem. Mathematical Research Center Technical Report 1604. University of Wisconsin, Madison 53706. (Pre-print of (6) which includes an exposition of the finite dimensional case in $\mathbb{R}^n$.)

6 ________ The Hopf Bifurcation Theorem in Infinite Dimensions. Arch., Rat., Mech., and Anal., 6 #2, 31 XII. 1977. pp 53-72.

7 Dieudonné J. Foundations of Modern Analysis. Pure and Applied Mathematics Series, vol., X. Academic Press. New York - London. 1960.

8 ________ Sur le Polygon de Newton. Arch., of Math., vol., 2 1949-50. Reproduced with some additonal examples in the following textbook.

9 Dieudonné J. Infinitesimal Calculus. William Clowes & Son's Ltd., London, Beccles and Colchester, Herman Paris 1971. Translated from Calcul Infinitésimal, published by Herman Paris 1968. (Appendix to chapter III on Newton's Polygon and Puiseux Developments.)

10 Fuchs B. A. and V. I. Levin. Functions of a Complex Variable and some of their applications, vol., II, English Translation 1961, Pergamon Press. Addison-Wesley, Reading, Massachussetts. (originally in Russian in 1949.)

11 Golubitsky M. and V. Guillemin. Stable Mappings and Their Singularities. Graduate text in Mathematics # 14. Springer-Verlag, New York. Heidelberg. Berlin 1973.

12 Graves L. M. Remarks on Singular Points of Functional Equations. Transactions American Math., Soc., 1955, 150-157.

13 Hardy G. H. Orders of Infinity. Cambridge Tracks no. 12. Cambridge 1924. Archiv der Mathematik. Bd 2. Heft 1.

14 Hukahara M., T. Kimura, M[me] T. Matuda. Equations Différentielles Ordinaires du Premier Ordre dans le Champ Complexe. Publication of the Mathematical Society of Japan #7, 1961.

15 Ince E. L. Ordinary Differential Equations. Longmans, Green and Co., 1926. Dover Reprint 1956.

16 Loomis L. H. and S. Sternberg. Advanced Calculus. Addison-Wesley Series in Mathematics 1968.

17 Moser J. K. The Order of a Singularity in Fuch's Theory. Math., Zeitschr. 72,379-398, 1960.

18 Nayfeh A. H. Perturbation Methods. Wiley Interscience Series of Texts, Monographs, and Tracts. Pure and Applied Mathematics. John Wiley and Sons, New York, London, Sydney, Toronto.

19 Nirenberg L. Topics in Nonlinear Functional Analysis 1973-4. Lecture Notes of the Courant Institute of Mathematical Sciences. 251 Mercier Street, New York, New York, 10012.

20 Perron O. Uber Lineare Differenzglechung und eine Anwendung auf lineare Differentialglechung mit Polynemkoeffizienten. Math., Zeitscher 72, 16-24, 1959.

21 Pimbley G. H. Eigenfunction Branches of Nonlinear Operators and their Bifurcations. Springer-Verlag Lecture Notes in Mathematics # 104.

22 Sather D. Branching of Solutions of an Equation in Hilbert Space Arch., Rat., Mech., Anal., 1970, 47-64.

23 ________ Branching of Solutions of Nonlinear Equations. Rocky Mountain Journal of Mathematics, vol., 3, # 2, spring 1973, 203-249.

24 Sattinger D. H. Group Representation Theory and Branch Points of Nonlinear Functional Equations. SIAM J. Math., Anal., vol., 8, # 2, 1977.

25 Stakgold I. Branching of Solutions of Nonlinear Equations. SIAM Review, vol., 13, # 3, 1971, 281-332.

26 Sestakov A. A. Asymptotic Behaviour of Multi-Dimensional Systems of Ordinary Differential Equations with a Singularity of Higher Order. Soviet Math., Dolk I, 1960, 394-397.

27 Vainberg M. M. and V. A. Trenogin. Theory of Branching of Solutions of Nonlinear Equations. English Translation: Noordhoff International Publishing, Leyden 1974. Original Title: "Teoriya volvleniya reshnic nelineinykh uravenii." Moscow 1969.

28 Walker R. J. Algebraic Curves. Princeton Univ., 1950. Dover reprint 1962.

29 Zeeman E. C. Catastrophe Theory (selected papers, 1972-1977) Addison-Wesley Publishing Company 1977.

Chapter 3.

A Complex-Analytic Perturbation Theory for the Existence of Stable and Unstable Manifolds of Analytic Functions.

1.Introduction. Passing through a hyperbolic fixed point of a C^k or analytic function, when $1<k<\infty$, there are stable and unstable C^k or analytic manifolds respectively. The stable manifold is mapped into itself by the function while the unstable manifold is mapped onto a superset of itself. The tangent spaces of these two manifolds are closed complements of each other at the hyperbolic fixed point. In a neighbourhood of the fixed point the stable manifold contains, and is characterized as, the set of points whose images approach the fixed point under forward iterates of the function. If the function is locally invertible at the fixed point, then in another neighbourhood of the fixed point, the unstable manifold contains and is characterized as the set of points which approach the fixed point under backward iterates of the function.

The existence of these just described invariant manifolds and their characterization in terms of backward and forward iterates of functions is well-known. See for instance S. SMALE(21) and I.C. IRWIN(12) and (13) for general and particular cases of the results asserted above. The articles (12) and (13) by I.C. IRWIN establish the existence and characterization of the stable manifold for Lipshitz continuous, C^k and analytic functions by employing sequence spaces, a graph transform(11), and the

Fréchet differentiability of a composition map. The use of these sequence spaces is a discrete analogue of the arguements in J. HALE(9) and in P. HARTMAN(10) showing the presence of stable and unstable manifolds passing through saddle points of autonomous differential equations. More generally, center, center-stable and center-unstable manifolds have been shown to pass through certain non-hyperbolic fixed points of some functions in O.E. LANFORD(15,appendix A) and M. HIRSCH, C. PUGH and M. SHUB(11); and to pass through generalized saddle-points of differential equations in R. ABRAHAM and J. ROBBIN(2, appendix C written by AL. KELLEY.)

A demonstration of the existence of the stable and unstable manifolds for real- and complex-analytic functions defined on real and complex Banach spaces respectively is presented below. Their existence here is a consequence of the ordinary implicit function thereom for Banach spaces in say L. NIRENBERG(18), the persistence under linearization of a root of a linearized conjugacy equation and the complex-analyticity of a composition map $(f,g) \rightarrow f \circ g$. The implicit function thereom is utilized to determine an operator-valued mapping. The existence of the invariant manifolds is proven first for complex-analytic functions and then by specialization for real-analytic maps. For convenience, real-analytic functions defined on real Banach spaces will be identified below with local extensions defined on complexifications of the real Banach spaces. Any originality in this chapter stems not from the statement of its results, but from their derivation with elementary calculus in Banach spaces.

The two theorems in section 5 below indicate an analytic dependence of the above invariant manifolds on certain perturbations of functions with hyperbolic fixed points.

The outline of the remainder of this chapter is as follows. In the next section there is a definition of hyperbolic fixed point and a statement of the main results in thereoms 1 and 2. Theorem 1 is given a proof assuming the statement of theorem 2. Preliminaries for a proof of theorem 2 are presented in section 3. These preliminaries include A) definition of notation for certain Banach function spaces; B) definition and demonstration of complex-analyticity of a composition map; C) a change of norm lemma; 'and D) a discussion of real-analyticity. In section 4, a proof of theorem 2 is divided into two parts: the complex-analytic case and the real-analytic case. The complex-analytic part does not require part D of the preliminaries. Section 5 concerns the afore-mentioned analytic dependence of the invariant manifolds on certain perturbations. Finally in section 6 the technique for constructing the invariant manifolds as suggested by the use of the ordinary implicit function theorem in the proof of theorem 2 is compared to other techniques.

2. Definition of Hyperbolic Fixed Points and Statement of Main Theorems.

Definition. Hyperbolic Fixed Point. Let W be an open subset of a Banach space X. A fixed point p of a C^1 function f mapping W into X is hyperbolic if there are closed complementary subspaces X_1 and X_2 in X with $X_1 \cap X_2 = \{0\}$ and $X_1 + X_2 = X$ such that (1)

$$Df(p)X_1 \subseteq X_1 \quad \text{and} \quad Df(p)X_2 \supseteq X_2 ,$$

(2) the restriction $Df(p)|_{X_2}$ of the Fréchet derivative $Df(p): X \to X$ of f is invertible, and (3) the spectral radii of the linear maps

$$Df(p)|_{X_1} : X_1 \to X_1 \quad \text{and} \quad (Df(p)|_{X_2})^{-1} : X_2 \to X_2$$

are both strictly less than 1.

Condition (3) in the above definition is a spectral separation constraint on Df(p) since it requires the spectrum of $Df(p)|_{X_1}$ to lie inside a disk of radius $k \leq 1$ and the spectrum of $Df(p)|_{X_2}$ to lie outside of a disk of radius $(1/k)$ in the complex plane for some positive real number k.

The existence of the stable and unstable manifolds passing through a hyperbolic fixed point of an analytic function defined on a real or complex Banach space is asserted in theorem 1. This theorem is a special case of theorem 2.

<u>Theorem 1. On the Existence of Stable and Unstable Analytic Manifolds.</u>

Let W be an open neighbourhood of the origin 0 in a Banach space X. Suppose 0 is a hyperbolic fixed point of an analytic function f mapping W into X. Let X_1 and X_2 be as in the preceding definition of hyperbolic fixed point with $p = 0$. Then

A) There is a neighbourhood U_1 of the origin in X_1, and analytic maps $h_1 : U_1 \to W \subset X$ and $c_1 : U_1 \to U_1 \subset X$ with $c_1(0) = 0$, $h_1(0) = 0$ and $Dh_1(0) = I_{X_1}$ such that $f \circ h_1 = h_1 \circ c_1$, i.e. the following diagram commutes

$$\begin{array}{ccc} U_1 & \xrightarrow{\;h_1\;} & W \subset X \\ {\scriptstyle c_1}\downarrow & & \downarrow{\scriptstyle f} \\ U_1 & \xrightarrow{\;h_1\;} & W \subset X. \end{array}$$

Moreover, the range S_1 of h_1 is an analytic manifold with $f(S_1) \subset S_1$ and a one chart atlas provided by (h_1, U_1); and, possibly after a change of norm on X_1, the map c_1 is a strict contraction on U_1.

B) Likewise, there is a neighbourhood U_2 of the origin in X_2, and analytic maps $h_2 : U_2 \to W \subset X$ and $c_2 : U_2 \to U_2 \subset X$ with $c_2(0) = 0$, $h_2(0) = 0$ and $Dh_2(0) = I_{X_2}$ such that $f \circ h_2 \circ c_2 = h_2$ i.e. the following diagram commutes

$$\begin{array}{ccc} U_2 & \xrightarrow{\;h_2\;} & W \subset X \\ {\scriptstyle c_2}\downarrow & & \uparrow{\scriptstyle f} \\ U_2 & \xrightarrow{\;h_2\;} & W \subset X \end{array}$$

The range S_2 of h_2 is an analytic manifold with $f(S_2) \supset S_2$ and a one chart atlas supplied by (h_2, U_2); and, possibly after a change of norm on X_2, the map c_2 is a strict contraction on U_2, and its inverse is a strict expansion on $c_2(U_2)$.

Note in the above theorem as c_1 is a strict contraction mapping the range S_1 of h_1 is a submanifold of the previously described stable manifold of f passing through the hyperbolic fixed point at 0. Since the tangent space X_1 of S_1 at 0 is also the tangent space of the stable manifold at the same point 0, it follows that S_1 contains a neighbourhood of 0 in the stable manifold. Hence by the characterization in (12) and (21), the stable manifold is contained in the set

$$\{ x \in W \mid f^{j}(x) \in S_1 \text{ for some integer } j \geq 1 \}.$$

Similarly the set

$$\{ x \in W \mid f^{-j}(x) \in S_2 \text{ for some integer } j \geq 1 \}$$

contains the unstable manifold passing through the fixed point 0 of f.

Also note X induces topologies on X_1 and X_2. Because X_1 and X_2 are closed complementary subspaces of X, the open mapping thereom(20) allows the topology on X to be identified with the product topology on $X_1 x X_2$. If $\|x\|_j$ are norms on $X_j (j=1,2)$ then the product topology on $X_1 x X_2$ is induced by the norm $\|(x_1,x_2)\| = \max(\|x_1\|_1, \|x_2\|_2)$. Therefore replacing given norms on X_1 and X_2 by equivalent norms does not affect their topologies nor the product topology on $X_1 x X_2$. Further, it induces a replacement of the original norm on X by an equivalent norm. This is the type of norm change referred to in theorem 1 above, and also in theorem 2 below.

<u>Theorem 2. An Extension of Theorem 1.</u> Let W be an open neighbourhood of the origin O in a Banach space X. Let $F,G : W \subset X \to Y$ be analytic maps valued in another Banach space Y. Suppose there are closed complementary subspaces N and V of X such that

A) $F(0) = G(0)$.

B) $DG(0)V \subset DF(0)V$ and $DF(0)N \subset DG(0)N$.

C) $DF(0)|_V : V \to DF(0)V = Y_1$ and $DG(0)|_N : N \to DG(0)N = Y_2$ are both bounded below.

D) Y_1 and Y_2 are closed complementary subspaces of Y.

E) the spectral radii $r(a)$ and $r(b)$ of the linear maps

$$a = (DF(0)|_V)^{-1}DG(0)|_V : V \to V$$

$$b = (DG(0)|_N)^{-1}DG(0)|_N : N \to N$$

are both $\leq k < 1$.

Then there is a neighbourhood U of the origin O in N, and analytic maps $h : U \to W \subset X$ and $c : U \to U \subset N$ which satisfy

$$F \circ h = G \circ h \circ c,\ h(0) = 0,\ c(0) = 0 \text{ and } Dh(0) = I_N : N \to X.$$

Further the range S of h is an analytic manifold with a one chart atlas provided by (h,U) which satisfies $F(S) \subset G(S)$; and, after possibly a change of norm on N or X, the function $c : U \to U$ is a strict contraction mapping.

Proof of Theorem 1 (assuming theorem 2.)

Part A of thereom 1 follows from thereom 2 by setting

$Y = X$, $(F,G) = (f, I_X)$, $N = X_1$, $V = X_2$ and $(h_1, c_1) = (h,c)$

for then $F \circ h = G \circ h \circ c$ and $F(S) \subset G(S)$ become respectively

$$f \circ h_1 = h_1 \circ c_1 \qquad \text{and} \qquad f(S_1) \subset S_2.$$

Similiarly, part B is obtained from letting

$Y = X$, $(F,G) = (I_X, f)$, $N = X_2$, $V = X_1$ and $(h_2, c_2) = (h,c)$

for then $G \circ h \circ c = F \circ h$ and $G(S) \supset F(S)$ become respectively

$$f \circ h_2 \circ c_2 = h_2 \qquad \text{and} \qquad f(S_2) \supset S_2.$$

The invertibility of $c_2 : U_2 \to U_2$ on a sufficiently small neighbourhood U follows from the inverse function theorem since

$$Df(0)\, Dh_2(0)\, Dc_2(0) = Dh_2(0) = I_{X_2}$$

implies $Dc_2(0) : X_2 \quad X_2$ has a bounded inverse $(Df(0)I_{X_2})^{-1}$.

Q.E.D.

Note the characterization of the invariant manifolds S_1 and S_2 in terms of the asymptotic behaviour of sequences generated by iterates of f or its inverse is not included in the statement of theorem 1. A proof of theorem 2 and preliminaries towards this proof are presented in the next two sections.

3. Preliminaries for a proof of Theorem 2.

A. Definition of Function Spaces $C^\omega(W_1,Z_2)$, $C^\omega(W_1,W_2)$ and $B^j(Z_1,Z_2)$.

For any pair (Z_1,Z_2) of Banach spaces over $\mathbb{R}$ or $\mathbb{C}$ and any subset W_1 of Z_1 which is the closure of its own interior, let $C^\omega(Z_1,Z_2)$ be the complex Banach space of functions $f : W_1 \subset Z_1 \to Z_2$ which are analytic in the interior of W_1, continuous on W_1 and possess finite sup-norms

$$\|f\| = \|f\|_{C^\omega(Z_1,Z_2)} = \sup_{z_1 \in W_1} \|f(z_1)\|_{Z_2}.$$

Subscripts on norms and elements of Banach spaces will be omitted whenever their affiliations should be self-evident from their context.
For $W_2 \subset Z$, let $C^\omega(W_1,W_2)$ denote the subset of W_2-valued maps in $C^\omega(W_1,Z_2)$. Elementary properties and equivalent definitions of complex-analytic maps are given in J. DIEUDONNE(8,Chapter IX.) and M.S. BERGER(3, pp 84-88.) . Also, let $B^j(Z_1,Z_2)$ denote the Banach space of bounded j-linear operators $L : Z_1^j \to Z_2$. Finally, let $D(s,Z)$ denote the closed ball of raduis $s > 0$ in a Banach space Z.

B.1. The Complex-Analyticity of a Composition Map.

Let A be a set. Let X and Y be complex Banach spaces. Denote by C(A,X) and C(A,Y) respectively the complex Banach spaces of functions $f : A \to X$ and $\hat{f} : A \to Y$ with finite norms

$$\|f\|_{C(A,X)} = \sup_{a \in A} \|f(a)\|_X \text{ and } \|\hat{f}\|_{C(A,Y)} = \sup_{a \in A} \|\hat{f}(a)\|_Y .$$

Lemma 1. Let A be a set. Let X and Y be complex Banach spaces. For $0 < s < r$, the composition map

$$(f,g) \in C^{\omega}(D(r,X),Y) \times D(s,C(A,X)) \to f\circ g \in C(A,Y)$$

is a complex-analytic operator-valued function, whose Frèchet derivative at (f,g) is given by the linear map

$$(\Delta f,\Delta g) \to \Delta f\circ g + (Df\circ g)(\Delta g)$$

when $(\Delta f,\Delta g) \in C^{\omega}(D(r,X),Y) \times C(A,X)$.

(Here $(Df\circ g)(\Delta g)(x) = Df(g(x))(\Delta g(x)) \in Y$.)

Proof: Part 1. Derivation of Cauchy Estimates.

Let w and z denote complex numbers. For (x,t) in $D(s,X) \times D(1,X)$ and $|z| < (r-s)/s$, Cauchy's formula implies

$$f(x+tz) = \oint_{|w|=\frac{r-s}{s}} \frac{f(x+tw)}{w-z}\,dw.$$

Thus

$$\frac{1}{(j!)}\frac{\partial^j f}{\partial z^j}(x+tz) = \oint_{|w|=\frac{r-s}{s}} \frac{f(x+tw)}{(w-z)^{j+1}}\,dw.$$

Therefore $\|t\|_X = 1$ and $z = 0$ imply the Cauchy estimate

$$\|(D_X^j f)(x)\, t^j\|_Y < 2\pi(j!)\left(\frac{r-s}{s}\right)^{j-1}\|f\|_{C^{\omega}(D(r,X),Y)}.$$

So that for $\|x\|_X < s$ in X

$$\| D_X^j f(x)\|_{B^j(X,Y)} < 2\pi(j!)\left(\frac{r-s}{s}\right)^{j-1}\|f\|_{C^{\omega}(D(r,X),Y)}$$

and hence the restriction of $D^j f$ to the disk $D(s,X)$ is a bounded linear map of $C^{\omega}(D(r,X),Y)$ into $C^{\omega}(D(s,X),B^j(X,Y))$.

Part 2. Computation of the Frèchet Derivative.

For (f,g) and $(\hat{f},\hat{g})$ in $C^{\omega}(D(s,X),Y) \times D(t,C(A,X))$

$$\hat{f}\circ\hat{g} - f\circ g - ((\hat{f} - f)\circ g + (Df\circ g)(\hat{g} - g))$$
$$= \hat{f}\circ\hat{g} - \hat{f}\circ g - (Df\circ g)(\hat{g} - g)$$
$$= \int_0^1 (D\hat{f}\circ(g+t(\hat{g}-g)) - Df\circ g)(\hat{g} - g)dt$$
$$= \int_0^1 [D(\hat{f} - f)\circ(g+t(\hat{g}-g)) + (Df\circ(g+t(\hat{g}-g)) - Df\circ g)]dt(\hat{g} - g)$$
$$= O(\|\hat{f} - f\| 2\pi(1!)(\tfrac{s}{r-s})^2 + (2\pi(2!)(\tfrac{s}{r-s})^3 \|f\| \|\hat{g} - g\|)\|\hat{g} - g\|)$$
$$= O(\|\hat{f} - f\|^2 + \|\hat{g} - g\|^2)$$

as $(\hat{f},\hat{g}) \to (f,g)$ because of Cauchy estimates and the Mean Value Thereom. Therefore $\Delta f\circ g+(Df\circ g)\Delta g$ represents the Frèchet derivative of the above composition map.

Q.E.D.

Lemma 2. Under the hypotheses of lemma 1, the map

$$(f,g) \in C^{\omega}(D(r,X),Y) \times D(s,C(A,X)) \to D^j f\circ g \in C(A,B^j(X,Y)) \subset B^j(C(A,X),Y)$$

is complex-analytic for $j \geq 1$.

Proof: By induction, starting with $j = 1$ the map of (f,g) to $D^j f\circ g$ is the j^{th} partial g-Frèchet derivative of the composition map in lemma 1.

Q.E.D.

B.2. The Complex-Analyticity of an Evaluation Map.

Lemma 3 below, not needed in the proof of theorem 2, is required in establishing the results in theorem 4. Theorem 4 indicates a continuous dependence of the invariant manifold S in theorem 2 on perturbations of (F,G).

Lemma 3. Analyticity of an Evaluation Map. For $0 < s < r$, the evaluation map

$$e(f,x) = f(x) : C^{\omega}(D(r,X),Y) \times D(s,X) \rightarrow Y$$

is complex-analytic when X and Y are complex Banach spaces.

Proof: In lemma 1, take A to be a singleton set. Then there is a norm preserving correspondence between X and C(A,X), and in particular between D(s,X) and D(s,C(A,X)). Now the points x of D(s,X) are identified with functions g in D(s,C(A,X)), so that lemma 1 applies.

For an alternate and equivalent proof, repeat the proof of lemma 1 with points x instead of functions g. Either way the proof is completed.

Q.E.D.

Remark. In R. ABRAHAM and J. ROBBIN (2) a converse to Taylor's formula is employed to show the C^k smoothness of the evaluation map $e(f,x) = f(x)$ when $f : X \rightarrow Y$ is a C^k map between (real) Banach spaces X and Y, which is defined on an open domain.

C. A Change of Norm.

Lemma 4. Let N be a normed vector space with an original norm $\|n\|$. Suppose $b : N \to N$ is a bounded linear operator with spectral radius

$$r(b) = \overline{\lim}_{j\to\infty} (\|b^j\|)^{1/j} < r.$$

Then there is a new norm

$$|n| = \sup_{m\geq 0} \left(\frac{\|b^m n\|}{r^m}\right)^{\frac{1}{m}}$$

equivalent to the original norm $\|n\|$ with the property

$$|bn| < r|n|.$$

Lemma 4 and its proof occur in O.E. LANFORD (15). A generalization is given in M. HIRSCH, C. PUGH and M. SHUB (11,p13) for the case where $b : N \to N$ is bounded below.

Proof. Observe $q = \sup_{m\geq 0} \frac{\|b^m\|}{r^m}$ is finite because $r > r(b)$. Therefore by the triangle inequality $|n|$ is a norm and $|n| < q\|n\|$. Further $\|n\| \leq |n|$ by definition since $b^0 n = n$ and $r^0 = 1$. Thus $|n|$ is equivalent to $\|n\|$. Now

$$|bn| = \sup_{m\geq 0} \frac{\|b^{m+1}(bn)\|}{r^m} = r\cdot\sup_{m\geq 0} \frac{\|b^{m+1} n\|}{r^m} < r|n|.$$

Q.E.D.

D.1. Real-Analytic Functions and Their Complexifications.

The preceding lemmas 1,2 and 4 are sufficient for a proof of theorem 2 when the Banach spaces X and Y in its statement are over the complex numbers. This complex case is treated in part 1 of the proof below in section 3. The remaining preliminaries are for the demonstration of the real case in part 2 of the proof.

For real Banach spaces X and Y denote by $X_c = X + iX$ and $Y_c = Y + iY$, the complexifications of these respective spaces (For complex Banach spaces X and Y, let $X_c = X$ and $Y_c = Y$.) In the following let X and Y be real Banach spaces. Then their complexifications can be identified respectively with the real product Banach spaces XxX and YxY. These identifications induce a complex vector space structure on XxX and YxY respectively.

Definition. Real Analytic Functions. Let W be an open subset of X. A function $f : W \to Y$ is real-analytic if there is an open set U in the complexification X_c of X with $W \subset U$, and a complex-analytic function $f_c : U \to Y_c$ such that the restriction of f_c to W is f.

Observe the collection of complex-analytic functions $f_c : U \to Y_c$ whose restrictions f to W are real-analytic maps of W into Y constitute a real subspace, closed under pointwise convergence, in the complex vector space of Y_c-valued complex-analytic maps on $U \subset X_c$. In particular note if $f : W \subset X \to Y$ is real-analytic on an open neighbourhood of the origin W in X, then there is an $r > 0$ and and element f_c of the real Banach subspace

$$C^{\omega}_{\mathbb{R}}(D(r,X_c),Y_c) = \{ g \in C^{\omega}(D(r,X_c),Y_c) : g|_{D(r,X)} \in C^{\omega}(D(r,X),Y) \}$$

of the complex-Banach space $C^{\omega}(D(r,X_c),Y_c)$, such that the restriction of f_c to $D(r,X)$ equals f on $D(r,X)$. Here $C^{\omega}_{\mathbb{R}}(D(r,X_c),Y_c)$ is a space of "real-analytic" functions on the disk $D(r,X)$ in X.

D.2. The Real-Analyticity of a Composition Map.

The above definition of real analytic functions, and lemmas 1 and 2 on the complex analyticity of a composition map imply the real-analyticity of the "real-valued" composition operation defined in lemma 5, below. The notation used here was introduced in the preceding subsection D.1.

Lemma 5. Let A, X and Y be real Banach spaces. For $r > s > 0$ and $t > 0$, the composition map

(1) $(f,g) \in C^{\omega}_{\mathbb{R}}(D(r,X_c),Y_c) \times C^{\omega}_{\mathbb{R}}(D(t,A_c),D(s,X_c)) \to f \circ g \in C^{\omega}_{\mathbb{R}}(D(t,A_c),Y_c)$

is a real-analytic operator-valued mapping whose real-linear Fréchet derivative at (f,g) in its domain is given by

$$(\Delta f,\Delta g) \to (\Delta f)\circ g + (Df\circ g)(\Delta g)$$

when $(\Delta f,\Delta g) \in C^{\omega}_{\mathbb{R}}(D(r,X_c),Y_c) \times C^{\omega}_{\mathbb{R}}(D(t,X_c),X_c)$. Furthermore the maps $(f,g) \to D^j f\circ g$ are also real-analytic.

The complex extension of the maps defined in lemma 5 are obtained by removing the subscript $\mathbb{R}$ in (1). This completes the preliminaries.

4. Proof of Theorem 2.

Part 1. The Complex Case.

Choose a norm $\|n\|$ on N so that the operator norm of

$$b = (DG(0)|_N)^{-1}(DF(0)|_N) : N \to N$$

is $< k < 1$. This is feasible because of the change of norm lemma 4 and the assumption in hypothesis (E) that the spectral radius of b is $< k < 1$.

Let N', V', X' and Y' be respectively the Banach subspaces of $C^\omega(D(1,N),N)$, $C^\omega(D(1,N),V)$, $C^\omega(D(1,N),X)$ and $C^\omega(D(1,N),Y)$ of functions which vanish at n=0 together with their first Fréchet derivatives. The linear map $b : N \to N$, or rather its restriction $b|_{D(1,N)}$, belongs to N' since its operator norm is $< k < 1$ in the N-norm chosen above.

Pick $r > 0$ so that $D(r,X) \subset W$ = domain (F) = domain (G) and so that F,G are members of $C^\omega(D(r,X),Y)$. Let $0 < t < r$. For (s,u,v) in $\mathbb{R} \times N' \times V'$ satisfying

$$|s| < 1, \ \|b+u\|_{C^\omega(D(1,N),N)} < k, \ \|v\|_{C^\omega(D(1,N),V)} < t$$

$$\text{and } \|s(Id_N+v)\|_{C^\omega(D(1,N),X)} < t$$

put

$$(1)\ H(s,u,v) = \begin{cases} \frac{1}{s}\{ F\circ s(Id_N+v) - G\circ s(Id_N+v)\circ(b+u) \} & \text{if } 0<|s|<1 \\ DF(0)(Id_N+v) - DG(0)(Id_N+v)\circ(b+u) & \text{if } s = 0. \end{cases}$$

Observe that $y=H(s,u,v)$ assumes values in $C^\omega(D(1,N),Y)$; that $y(0) = 0$ because $F(0) = G(0)$ and by definition $DF(0)Id_N - DG(0)\circ Id_N\circ b = 0$; and that by the chain rule $y'(0) = 0$ when $s = 0$ or $0 < s < 1$. Therefore $H : \mathbb{R} \times N' \times V' \to Y'$.

A second representation of H,

(2) $H(s,u,v) = \int_0^1 DF(ts(Id_N+v))(Id_N+v) - DG(0)(ts(Id_N+v)\circ(b+u))(Id_N+v)(b+u)dt$

valid for all $|s| < 1$ follows from $F(0) = G(0)$. The Fréchet differentiability and analyticity of H on its open domain are now implied by lemmas 1 and 2 and the chain rule when they are applied to the second representation (2) of H.

The partial (u,v) Fréchet derivative of H(s,u,v) at (0,0,0) is given by the chain rule as

$$Q(du,dv) = D_XF(0)dv - D_XG(0)(dv)b - D_XG(0)du \in Y'$$

for (du,dv) in $N' \times V'$.

Since by hypothesis (C) Y is the direct sum of the closed complementary subspaces $Y_1 = DF(0)V$ and $Y_2 = DG(0)N$, there is a continuous projection $P : Y \to Y_1 = DF(0)V$ with $\text{kernel}(P) = Y_2 = DG(0)N$. Now given y in Y' $Q(u,v) = y$ iff

$$Py = (DF(0)|_V\circ v - (DG(0)|_V)\circ v\circ b$$

and

$$(I - P)y = -DG(0)|_N\circ u$$

since $\text{range}(DF(0)\circ v) \subset DF(0)V = Y_1 = \text{range}(P)$; $\text{range}(DG(0)\circ v\circ b) \subset DG(0)V \subset DF(0)V = \text{range}(P)$ and $\text{range}(DG(0)\circ u) \subset DG(0)N = Y_2 = \text{range}(I - P)$.

Therefore $Q(u,v) = y$ implies

$$u = -(DG(0)_V)^{-1}(I - P)y.$$

To find v note

$$Py = (DF(0)|_V)(v - a\circ v\circ b).$$

where $a = (DF(0)|_V)^{-1}\, DG(0)|_V : V \to V$ is as in hypothesis (E). The map

$$Lv = a\circ v\circ b : V' \to V'$$

defined for v in V' has spectral raduis $r(a) < k < 1$ since by induction

$$L^j v = a^j \circ v \circ b^j$$

and since for $\|v\| = 1$

$$\|L^j v\| < (\|a^j\|)(\|v\ b^j\|) < \|a^j\|$$

Hence $(I - L)^{-1}$ exists, is bounded, and

$$v = (I - L)^{-1}(DF(0)|_V)^{-1}Py.$$

The above two formulas for u and v imply the partial (u,v) Fréchet derivative of H(s,u,v) at (0,0,0) in $R \times U' \times V'$ has a bounded inverse defined on Y'. Therefore, since $H(0,0,0) = 0$, the ordinary implicit function theorem for Banach spaces in L.NIRENBERG(18) implies there is an $s > 0$ and functions (u,v) in $U' \times V'$ (uniquely determined by s) such that

$$\begin{aligned} 0 &= sH(s,u,v) \\ &= F\circ s(Id_N + v) - G\circ s(Id_N + v)\circ(b + u) \\ &= (F\circ(Id_N + sv\circ\tfrac{1}{s}Id_N) - g\circ(Id_N + sv\circ\tfrac{1}{s}Id_N)\circ(b + su\circ\tfrac{1}{s}Id_N))\circ sId_N \end{aligned}$$

Therefore the functions

$$h = (Id_N + sv\circ\tfrac{1}{s}Id_N) : D(s,N) \to X$$

$$c = (b + su\circ\tfrac{1}{s}Id_N) : D(s,N) \to X$$

satisfy

$$F\circ h = G\circ h\circ c$$

on D(s,N), and

$$(Dh(0), Dc(0)) = (Id_N, b).$$

This completes the proof in the complex case.

Part 2. The Real Case.

Here X and Y are real Banach spaces. The maps F and G have complex-analytic extensions defined on a neighbourhood of the origin in X_c and valued in Y_c. By abuse of notation, these extensions will again be denoted by F and G respectively. For these complex extensions, the complex case in part 1 still holds with (X_c,Y_c) in place of (X,Y). Yet the same arguments still hold if the complex-Banach spaces in part 1 are all replaced by their real-subspaces of "real-analytic" functions (see the discussion in part D of the preliminaries in section 3.) This modification, with lemma 5 instead of lemmas 1 and 2, yields the real-case since the functions (h,c) satisfying $F \circ h = G \circ h \circ c$ will be extensions of real-analytic functions.

Q.E.D.

It is a corollary of the preceding arguments that the substitution of formal power series developements of (h,c) in $F \circ h = G \circ h \circ c$ results in recursively determined, locally convergent, power series expansions which satisfy this identity.

5. Analytic Dependence on Perturbations. Complex Case.

Theorem 3. Let X and Y be complex Banach spaces. For $r > 0$, let

$$A_r = \{(F,G) \in [C^{\omega}(D(r,X),Y)]^2 \ : \ F(0) = G(0) \}.$$

Then A_r is a complex Banach space. If (F_1, G_1) in A_r satisfy the hypotheses of theorem 2 then there is a neighbourhood M_r of (F_1, G_1) in A_r a neighbourhood U of the origin 0 in N, and a norm $\|n\|$ on N such that for each (F,G) in M_r there is a pair of analytic, operator-valued, functions

$$(h,c) = (h(F,G)(\cdot), c(F,G)(\cdot))$$

valued in $C^{\omega}(U,X) \times C^{\omega}(U,U)$ which satisfy

$$F \circ h = G \circ h \circ c, \ h(0) = 0, \ c(0) = 0,$$

$$D_n h(F_1,G_1)(0) = Id_N : N \to X$$

and

$$D_n c(F_1,G_1)(0) = b \quad : N \to N.$$

Further for each (F,G) in M_r, $c(F,G)(\cdot) : U \to U$ is a strict contraction map with respect to the norm $\|n\|$; on N and the range $S = S(F,G)$ of $h(F,G)(\cdot)$ is a sub-manifold of X with a one-chart atlas provided by $(h(F,G)(\cdot),U)$, which satisfies $F(S) \subset G(S)$.

Proof. Observe by the preliminary lemmas 1 and 2, that the map H is defined by formulas (1) and (2) which are analytic functions of (s,u,v,F,G) in $\mathbb{R} \times U' \times V' \times A_r$. The hypotheses of the ordinary implicit function for Banach spaces are now satisfied at $(s,u,v,F,G) = (0,0,0,F_1,G_1)$.

Q.E.D.

Theorem 4 extends theorem 3 by replacing A_r by the larger Banach space $[C^{\omega}(D(r,X),Y)]^2$.

Theorem 4. Let X and Y be complex Banach spaces. If (F_1,G_1) in $[C^{\omega}(D(r,X),Y)]^2$ satisfy the hypotheses of theorem 2 then there is a neighbourhood M of (F_1,G_1) in $[C^{\omega}(D(r,X),Y)]^2$, a neighbourhood U of the origin in N, and a norm $\|n\|$ on N, such that for each (F,G) in M there is a pair of analytic, operator-valued, functions

$$(h,c) = (h(F,G)(\cdot),c(F,G)(\cdot))$$

valued in $C^{\omega}(U,X) \times C^{\omega}(U,U)$ which satisfy

$$F \circ h = G \circ h \circ c, \; c(0) = 0,$$

$$D_n h(F_1,G_1)(0) = Id_N : N \to X$$

and

$$D_n c(F_1,G_1)(0) = b \quad : N \to X.$$

Further for each (F,G) in M, $c(F,G)(\cdot) : U \to U$ is a strict contraction map with respect to the norm $\|n\|$ on N; and the range $S = S(F,G)$ of $h(F,G)(\cdot)$ is a sub-manifold of X with a one-chart atlas supplied by (h(F,G),U), which satisfies $F(S) \subset G(S)$.

Summary of Proof. The proof below first demonstrates that all functions (F,G) sufficiently near (F_1,G_1) in $[C^{\omega}(D(r,X),Y)]^2$ have points $x = x(F,G)$ which satisfy $F(x) = G(x)$, depend analyticly on (F,G) and can be translated to the origin of X. The latter translation permits an application of theorem 3 to functions $(\hat{F},\hat{G})$ in $A_{\frac{1}{2}r}$ of the form $(F(Id_X + x(F,G)), G(Id_X + x(F,G)))$, with the result that (F,G) "inherit" from $(\hat{F},\hat{G})$ the properties indicated in theorem 4, above. The latter inheritance gives the last part of the proof.

<u>Proof of theorem 4.</u> For (F,G,x) in $[C^{\omega}(D(r,X),Y)]^2 \times X$, put

$$E(F,G,x) = F(x) - G(x).$$

Then E is complex-analytic by lemma 3 and by the continuity and linearity of addition in $C^{\omega}(D(r,X),Y)$. Moreover $E(F_1,G_1,0) = F_1(0) - G_1(0) = 0$.

The partial x-Fréchet derivative of E at $(F_1,G_1,0)$ is

$$D_x E(F_1,G_2,0) = DF_1(0) - DG_1(0) : X \to Y.$$

Given $y = y_1 + y_2$ in Y, $y_1 \in Y_1$, $y_2 \in Y_2$, and $x = n + v$ in X, $n \in N$, $v \in V$

(1) $$D_x E(F_1,G_1,0)x = y$$

iff

(2) $$\begin{aligned} y_1 &= (DF_1(0) - DG_1(0))v = (DF_1(0)|_V)(Id_V - a)v \\ y_2 &= (DF_2(0) - DG_2(0))n = (DG_2(0)|_N)(b - Id_N)n \end{aligned}$$

when

$$a = (DF_1(0)|_V)^{-1}(DG_1(0)|_V) : V \to V$$
$$b = (DG_2(0)|_N)^{-1}(DF_2(0)|_N) : N \to N,$$

since $Y_1 = DF_1(0)V \supset DG_1(0)V$ and $Y_2 = DG_2(0)N \supset DF_2(0)N$. But by assumption $a : V \to V$ and $b : N \to N$ are bounded linear operators with spectral radii $k < 1$. From equation (2), equation (1) holds iff

$$v = (Id_V - a)^{-1}(DF_1(0)|_V)y_1$$
$$n = (b - Id_N)^{-1}(DG_1(0)|_N)y_2$$

Hence the partial x-Fréchet derivative of E at $(F_1,G_1,0)$ is surjective and bounded below.

The implicit function theorem for Banach spaces implies there is a neighbourhood M_1 of (F_1,G_1) in $[C^{\omega}(D(r,X),Y)]^2$, a $t > 0$ with $t < \frac{1}{2}r$ and an analytic map $x(F,G)$ valued in the ball $D(t,X) \subset X$ such that in this ball for each (F,G) in M_1 the one and only solution of $F(x) = G(x)$ is $x = x(F,G)$. In particular $x(F_1,G_1) = 0$.

For (F,G) in M_1, put

$$J(F,G) = (\hat{F},\hat{G}) = (F\circ(x(F,G) + Id_X),\ G\circ(x(F,G) + Id_X)).$$

Here $(\hat{F},\hat{G})$ represent the components of $J(F,G)$. The restrictions of the functions $\hat{F}$ and $\hat{G}$ belong to the Banach space $A_{\frac{1}{2}r}$ defined in theorem 3, since $\|x(F,G)\| < \frac{1}{2}r$ and $D(\frac{1}{2}r,X) + D(\frac{1}{2}r,X) = D(r,X) = \text{domain}(F) = \text{domain}(G)$. Hence by the chain rule J is a complex-analytic map of $M_1 \subset [C^{\omega}(D(r,X),Y)]^2$ into $A_{\frac{1}{2}r}$ which fixes (F_1,G_1) since $x(F_1,G_1) = 0$ and $A_{\frac{1}{2}r} \supset A_r$.

Theorem 3 now applies to the element (F_1,G_1) of $A_{\frac{1}{2}r}$. Hence there is a neighbourhood $M_{\frac{1}{2}r}$ of (F_1,G_1) in $A_{\frac{1}{2}r}$ in which the conclusions of theorem 3 are valid. Hence for each (F,G) in the pulled-back neighbourhood $M = J^{-1}(M_{\frac{1}{2}r})$ in $C^{\omega}(D(r,X),Y)$ there is a pair of analytic, operator-valued, functions

$$(\hat{h},\hat{c}) = (h(J(F,G))(\cdot),c(J(F,G))(\cdot))$$

valued in $C^{\omega}(U,U) \times C^{\omega}(U,X)$ for which

$$F\circ(x(F,G) + \hat{h}) = G\circ(x(F,G) + \hat{h})\circ\hat{c};\quad \hat{c}(0) = 0$$

and

$$D_n\hat{h}(F_1,G_1)(0) = Id_N : N \to X,$$

when U, $h(F,G)(\cdot)$ and $c(F,G)(\cdot)$ are as they appear in theorem 3 with $\frac{1}{2}r$ instead of $r > 0$. The remainder of theorem 4 follows from an obvious change of notation and the conclusions of theorem 3.

Q.E.D.

Similar arguments for real Banach spaces $\hat{X}$ and $\hat{Y}$ with the real Banach space $C^{\omega}_{R}(D(r,\hat{X}_C),\hat{Y}_C)$ in place of the complex Banach space $C^{\omega}(D(r,X);Y)$ yield real-analytic analogues of theorems 3 and 4.

6. Some Brief Comparisons with other Persistence Arguments (Related Topics.)

The usual proof of the ordinary implicit function theorem for Banach spaces as in say L. NIRENBERG(18) employs a variation of Newton's algorithm# from numerical analysis and the strict contraction mapping principle to construct local representatives of implicitly defined functions. In the preceding part of this chapter invariant manifolds S were obtained with the aid of this implicit function theorem. The usual proof of this theorem gives a construction of some invariant manifolds - here the unstable and stable manifolds passing through a hyperbolic fixed point of an analytic function - which is based on the inversion of approximate linearized problems. This construction has similarities with constructive arguments appearing in M. HIRSCH, C. PUGH and M. SHUB (11), O.E. LANFORD (15) and W.T. KYNER (14) in which the hypotheses of the ordinary implicit function theorem for Banach spaces are not satified, but in which Lipshitz conditions, the strict contraction mapping principle, linearization and variations of Newton's algorithm can be applied directly. In J.K. MOSER (17) there is a discussion of rapidly convergent iterative methods for solving some conjugacy problems, based on linearization. In J. MATHER (16) the application of implicit function theorems from (1) to a conjugacy problem is indicated to show the persistence of Anosov-Diffeomorphisms.

the chord method.

7. Bibliography for Chapter 3.

1 Abraham R. Lectures of Smale on Differential Topology. Lecture Notes, Columbia University, New York 1962.

2 Abraham R. and J. Robbin. Transversal Mappings and Flows. W. A. Benjamin INC., , New York 1967.

3 Berger M. S. Nonlinearity and Functional Analysis: Lectures on Nonlinear Problems in Mathematical Analysis. Pure and Applied Mathematics Series. Academic Press, New York 1977.

4 Birkhoff G. D. Surface Transformations and their Dynamical Applications, Acta Math, 43 (1920) pp1-119.

5 Crandall M. G. and P. H. Rabinowitz. Bifurcation from Simple Eigen-values. Journal of Functional Analysis. 8 (1971) pp 321-40.

6 ________ The Hopf Bifurcation Thereom. Mathematics Research Center Technical Report 1604. University of Wisconsin., Madison 1975. (A pre-print of the following article; contains a clear exposition of Hopf Bifurcation in $\mathbb{R}^n$.)

7 ________ The Hopf Bifurcation Thereom in Infinite Dimensions. pp 53-72 Arch., Rat., Mech., and Anal., 67#1,31,XII,1977. (omitts the finite dimensional case included in (6) above.)

8 Dieudonnè J. Foundations of Modern Analysis. Pure and Applied Mathematics Series, vol., X. Academic Press, New York 1960.

9 Hale J. K. Ordinary Differential Equations. Pure and Applied Mathematics Series, vol., XXI. Wiley Interscience, New York 1969.

10 Hartman P. Ordinary Differential Equations. John Wiley and Sons Inc., New York 1964.

11 Hirsch M., C. Pugh, and M. Shub. Invariant Manifolds, Lecture Notes in Mathematics, vol. 583. Springer Verlag New York 1977.

12 Irwin M.C. On the Stable Manifold Thereom. Bull, London Math Soc., 2, 1970, pp 196-198.

13 ________ On the Smoothness of the Composition Map. Quarterly Journal of Math., (2) 23, 1972. pp 113-133. Oxford University Press.

14 Kyner W.T. Invariant Manifolds, Rend. del Circolo Matematico Palermo, Ser. 11. Vol. 9, 1961, pp 98-110.

15 Lanford O. E. Bifurcation of Periodic Solutions into Invariant Tori: the work of Ruelle and Takens. Appendix A Nonlinear Problems in the Physical Sciences and Biology. Springer-Verlag, Lecture Notes in Mathematics, vol. 322.

16 Mather J. Appendix to Part I of S. Smale (21) on Anosov Diffeomorphisms pp 792-5.

17 Moser J.K. A Rapidly Convergent Iteration Method and Non-linear Partial Differential Equations I, II pp 265-315 and pp 499-535 Scuola Norm Sup., Pisa, 1966.

18 Nirenberg L. Topics in Nonlinear Functional Analysis 1973-4. Lecture Notes of the Courant Institute of Mathematical Sciences. 251 Mercier Street, New York, New York, 10012.

19 Poincaré H. Les Méthodes nouvelles de la mécanique céleste, Vols. I, II, III. Paris 1892, 1893, 1899; reprint, Dover, New York, 1957.

20. Rudin W. Real and Complex Analysis, 2nd Edition. McGraw-Hill Series in Higher Mathematics, 1974.

21 Smale S. Differentiable Dynamical Systems. Bulletin American Math Soc., 73 (1967) pp 747-817.