THE IMPLICIT FUNCTION THEOREM AND ANALYTIC DIFFERENTIAL EQUATIONS.

K.R. Meyer.

Reprinted from
Warwick Dynamical Systems 1974,
Lecture Notes in Mathematics, Springer-Verlag, 1975.

The Implicit Function Theorem and Analytic Differential Equations.

K.R. Meyer.

1. Introduction. Although it is old, the calculus in Banach spaces has recently become a popular tool in the theory of differential equations. Graves [2] in 1927 showed that a general form of the implicit function theorem could be used to establish the fundamental existence and uniqueness theorem of differential equations. Dieudonné [1], Lang [4] et al have made the basic theory readily available and the power of this method has forcefully been demonstrated by the results of Mather [7], Robbin [8], Smale [10] et al.

In this paper we show how several theorems on analytic differential equations can be established by using the implicit function theorem. These theorems were originally proved by the method of majorants. Indeed many of the classical theorems in the analytic theory of differential equations which were originally proved by constructing majorant series can be proved by the methods presented here.

The key to the method lies in the definition of the function space A_δ given in section 2. This space was used in Harris, Sibuya and Weinberg [3]. Many of the ideas presented here came from several conversations with Professor Y. Sibuya over a period of many years.

2. The space of analytic functions.

There are several ways of embedding an analytic function in a Banach space. The space considered here is very useful when one wishes to consider an analytic function as a power series and deal directly with the coefficients. The proof of Poincaré's linearization theorem and its generalization given in section 3 depends heavily on this choice of function space.

The notation follows Dieudonné [1] and Lang [4]. Let E, F, G,...

denote Banach spaces with norms $|\cdot|$ and $L_S^k(E, F)$ the linear space of all bounded symmetric k-linear maps from E to F. L_S^k will not always be normed in the usual way. The spaces $L_S^k(E, F)$, $k = 1,2,\ldots$ will be said to be consistently normed if each space $L_S^k(E, F)$ has a norm $|\cdot|_k$ with the following properties

1) $\{L_S^k(E, F), |\cdot|_k\}$ is a Banach space

2) $|a(x_1,\ldots,x_k)| \leq |a|_k|x_1|\ldots|x_k|$ for all $x_i \in E$ and

3) the usual isomorphism of $L_S^{h+k}(E, F)$ into $L_S^h(E, L_S^k(E, F))$ is a norm preserving isomorphism.

Examples. 1) If the norms on $L_S^k(E, F)$ are defined as usual by $|a|_k = \sup\{|a(x_1,\ldots,x_k)| \; ; \; x_i \in E$ and $|x_i| \leq 1\}$ then it is well known that 1), 2) and 3) hold.

2) Let $E = C^n$ (or R^n) with $|x| =$ the maximum of the modulus of the components of x. Let $e_1,\ldots,e_n$ be the usual basis of C^n. A k-linear map from C^n to F has the form

$$a(x_1,\ldots,x_k) = \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} \alpha_{i_1\ldots i_k} x_{1i_1} \cdots x_{ki_k}$$

where $x_j = (x_{j1},\ldots,x_{jn})$ and $\alpha_{i_1\ldots i_k} = a(e_{i_1},\ldots,e_{i_k}) \in F$.

As usual a is symmetric if and only if a permutation of the subscripts of the α's leaves them unchanged. Define

$$|a|_k = \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} |\alpha_{i_1\ldots i_k}| .$$

Clearly 1) and 2) hold. Now let $a \in L_S^{h+k}(E, F)$. If $x_1,\ldots,x_h \in E$ then we may consider

$$a(x_1,\ldots,x_h,*,\ldots,*) \in L_S^k(E, F) \text{ defined by}$$

$$a(x_1,\ldots,x_h,*,\ldots,*)(y_1,\ldots,y_k) = a(x_1,\ldots,x_h,y_1,\ldots,y_k).$$

Thus

$$|a|_h = \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} |(e_{i_1},\ldots,e_{i_h},*,\ldots,*)|_k$$

$$= \sum_{i_1=1}^{n} \cdots \sum_{i_h=1}^{n} \{ \sum_{j_1=1}^{n} \cdots \sum_{j_h=1}^{n} |(e_{i_1},\ldots,e_{i_h},e_{j_1},\ldots,e_{j_k})| \}$$

$$= |a|_{k+h} .$$

Henceforth we shall assume that the spaces $L_S^k(E, F)$ are consistently normed and omit the subscript on the norm.

Let $\delta > 0$. Define $A_\delta(E, F)$ as the set of all formal power series of the form

$$f(x) = \sum_{k=0}^{\infty} a_k(x^k)$$

where $a_k \in L_S^k(E, F)$, $x^k = (x,\ldots,x) \in E^k$ such that

$$||f|| = ||f||_\delta = \sum_{k=0}^{\infty} |a_k|\delta^k < \infty .$$

Note that $\{A_\delta(E, F), ||\cdot||_\delta\}$ is essentially ℓ_1 and so is a Banach space. By the Weierstrass M-test we see that f is absolutely and uniformly convergent for $|x| \leq \delta$ so $f : \{x \in E ; |x| \leq \delta\} \to F$ is continuous. Also note that $\sup\{|f(x)| ; |x| \leq \delta\} \leq ||f||_\delta$ and $|a_k| \leq ||f||_\delta/\delta^k$. This last inequality plays the role of Cauchy's inequality.

The space $A_\delta(E, F)$ is the major space we shall analyze and use in the later applications. When $E = F = C^n$ and the spaces $L_S^k(C^n,C^n)$ are consistently normed as in the second example we have the space used by Harris, Sibuya and Weinberg [3]. This space norms an analytic function directly from the coefficients of its power series expansion and is very useful for problems where one must look at the series expansions closely.

We shall now develop some fundamental facts about this space which will be used in our subsequent applications.

Lemma 1. Let $f : \{x \in C^n ; |x_i| < \rho\} \to F$ be analytic and bounded by M in norm. Then for each $\delta < \rho$, $f \in A_\delta(C^n, F)$ and $||f||_\delta \leq M/(1-\delta/\rho)^n$.

Remark. Here as elsewhere we identify a function and its power series representation. In the above the spaces $L^k_S(C^n, F)$ are to be normed either as in example 1 or 2.

Proof. Since f is analytic for $|x_i| < \rho$, f has a power series representation

$$f(x_1,\ldots,x_n) = \sum_{k_i \geq 0} \alpha_{k_1\ldots k_n} x_1^{k_1}\ldots x_n^{k_n}$$

By Cauchy's inequality $|\alpha_{k_1\ldots k_n}| \leq M\rho^{-(k_1+\ldots+k_n)}$ thus

$$||f||_\delta = \sum_{k_i \geq 0} |\alpha_{k_1\ldots k_n}| \delta^{k_1+\ldots+k_n}$$

$$<< \sum_{k_i \geq 0} M(\delta/\rho)^{k_1+\ldots+k_n} = M/(1-\delta/\rho)^n .$$

In the above $<<$ is used in the usual sense of majorant series, i.e. each term of the series on the right is greater than or equal to the corresponding term on the left. The above proof gives the lemma when the L^k_S are normed as in example 2) but this norm dominates the norm of example 1).

Lemma 2. If $\alpha_i \geq 0$, $i = 0,1,2,\ldots$ and $\sum_0^\infty \alpha_k \delta^k = M < \infty$ then, for any positive integer i and any ρ, $0 < \rho < \delta$,

$$\sum_{k=i}^{\infty} \frac{k!}{(k-i)!} \alpha_k \rho^{k-i} \leq \frac{i!\, M}{(\delta - \rho)^i} .$$

Proof. Consider the scalar complex valued function

$$g(z) = \sum_{k=0}^{\infty} \alpha_k z^k$$

which is analytic and bounded by M in the disk $|z| < \delta$. The result follows by applying Cauchy's inequality to

$$g^{(i)}(\rho) = \sum_{k=i}^{\infty} \frac{k!}{(k-i)!} \alpha_k \rho^{k-i} .$$

Lemma 3. If $f \in A_\delta(E, F)$ and $0 < \rho < \delta$ then

$$D^i f \in A_\rho(E, L_S^i(E, F)) \text{ and } ||D^i f||_\rho \leq i!||f||_\delta/(\delta-\rho)^i .$$

Proof. Let $f(x) = \sum_{k=0}^{\infty} a_k(x^k)$. Now

1) $$\infty > ||f||_\delta = \sum_{k=0}^{\infty} |a_k|\delta^k = \sum_{k=0}^{\infty} |a_k|\{\rho + (\delta - \rho)\}^k$$

$$= \sum_{k=0}^{\infty} \sum_{i=0}^{k} \binom{k}{i} |a_k| \rho^{k-i}(\delta - \rho)^i$$

$$= \sum_{i=0}^{\infty} \{\sum_{k=i}^{\infty} \binom{k}{i} |a_k| \rho^{k-i}\}(\delta - \rho)^i .$$

All terms in the above series are non negative and so the series maybe rearranged as shown. Now let $|x| < \rho$ and $|y| < (\delta - \rho)/2$. Then

2) $$f(x + y) = \sum_{k=0}^{\infty} a_k((x + y)^k) = \sum_{k=0}^{\infty} \{\sum_{i=0}^{k} \binom{k}{i} a_k(x^{k-i}, y^i)\}$$

$$= \sum_{i=0}^{\infty} \{ \sum_{k=i}^{\infty} \binom{k}{i} a_k (x^{k-i},*)\}(y^i)$$

The last rearrangement in 2) follows from the fact that the last two series in 1) majorize the last two series in 2). Thus the last two series of 2) are absolutely convergent. From 2) we note that the i^{th} derivative should be

$$D^i f(x) = i! \sum_{k=i}^{\infty} \binom{k}{i} a_k (x^{k-i},*) \ .$$

For the present use the above as a formal definition. Now

$$\|D^i f(x)\|_\rho = i! \sum_{k=i}^{\infty} \binom{k}{i} |a_k| \rho^{k-i} \leq i! \|f\|_\delta / (\delta-\rho)^i$$

by lemma 2. So

$$f(x + y) = \sum_{i=0}^{\infty} \frac{D^i f(x)}{i!} (y^i)$$

where $D^i f \in A_\rho(E, L_S^i(E, F))$. Now it must be shown that $D^i f$ is the i^{th} derivative of f and for this one uses the converse of Taylor's theorem [4].

$$f(x + y) = \sum_{i=0}^{N-1} \frac{D^i f(x)}{i!} (y^i) + R_N$$

where $$R_N = \{ \sum_{i=N}^{\infty} \frac{D^i f(x)}{i!} (y^{i-N},*)\}(y^N) \ .$$

To estimate R_N use the above estimate on $D^i f$ and $|y| < (\delta - \rho)/2$ in the series to obtain

$$|R_N| \leq \{ \sum_{i=N}^{\infty} \frac{\|f\|_\delta}{(\delta - \rho)^i} \left(\frac{\delta - \rho}{2}\right)^{i-N}\} |y|^N$$

$$\leq \frac{\|f\|_\delta}{(\delta - \rho)^N} |y|^N \ .$$

So by the converse of Taylor's theorem the i^{th} derivative of f exists and is indeed equal to $D^i f$ for $0 < i < N$.

Lemma 4. Let $f \in A_\delta(E, F)$ and $g \in A_\eta(D, E)$ with $||g||_\eta \leq \delta$ then $f \circ g \in A_\eta(D, F)$ and $||f \circ g||_\eta \leq ||f||_\delta$.

Proof. Let $f(x) = \sum_{k=0}^{\infty} a_k(x^k)$ and $g(x) = \sum_{\ell=0}^{\infty} b_\ell(x)$.

$$\infty > ||f||_\delta = \sum_{k=0}^{\infty} |a_k| \delta^k$$

$$\geq \sum_{k=0}^{\infty} |a_k| \{ \sum_{\ell=0}^{\infty} |b_\ell| \eta^\ell \}^k$$

$$= \sum_{k=0}^{\infty} |a_k| \sum_{s=0}^{\infty} \{ \sum_* |b_{\ell_1}| \cdots |b_{\ell_r}| \} \eta^s$$

$$= \sum_{s=0}^{\infty} \{ \sum_{k=0}^{\infty} |a_k| \sum_* |b_{\ell_1}| \cdots |b_{\ell_r}| \} \eta^s$$

$$= ||f \circ g||_\eta \quad .$$

In the above the summation denoted by $\sum_*$ is to be taken over all sets of integers $\ell_1, \ldots, \ell_r$ with $\ell_i \geq 0$ and $\ell_1 + \ldots + \ell_r = s$.

The above lemma tells when the composition map is well defined. The next lemmas prove that it is continuously differentiable in the interior of its domain of definition.

Lemma 5. Let $U = \{g \in A_\eta(D, E) \; ; \; ||g||_\eta < \delta\}$ and $O : A_\delta(E, F) \times U \to A_\eta(D, F) : (f, g) \mapsto f \circ g$. Then O is continuous.

Proof. By lemma 4 one has

$$||O(f + f', g) - O(f, g)||_\eta = ||f' \circ g||_\eta \leq ||f'||_\eta$$

and so O is uniformly continuous in its first argument. Let $f \in A_\delta(E, F)$, $g \in U$, $||g||_\eta = \alpha < \delta$ and β defined by $3\beta = \delta - \alpha$.

By lemma 3 one has $D^k f \in A_{\delta-2\beta}(E, L_S^k(E, F))$ and $||D^k f||_{\delta-2\beta} \leq k!\ ||f||_\delta (2\beta)^{-k}$. Let $h \in U$ with $||h||_\eta < \beta$. Then

$$O(f,g+h)(x) - O(f,g)(x) = \sum_{k=1}^{\infty} \frac{D^k f(g(x))}{k!} (h(x)^k)$$

and so $||O(f,g+h) - O(f,g)||_\eta \leq \dfrac{||f||_\delta}{\beta} ||h||_\eta$ and so O is continuous with respect to its second argument. Thus O is continuous.

Lemma 6. Let U and O be as in Lemma 5. Then O is C^∞.

Proof. First let us show that O has continuous partials of all orders with respect to its second arguments. Let $f \in A_\delta(E, F)$, $g \in U$, $||g||_\eta = \alpha < \delta$ and β determined by $3\beta = \delta - \alpha$. Now $D^k f \in A_{\delta-2\beta}(E, L_S^k(E, F))$ and $||D^k f||_{\delta-2\beta} \leq k!\ ||f||_\delta (2\beta)^{-k}$. Let $h \in U$ with $||h||_\eta < \beta$. Then

$$O(f,g+h)(x) = \sum_{k=0}^{N} \frac{Q_k(g)(h^k)(x)}{k!} + R_{N+1}$$

where $Q_k(g) = (D^k f) \circ g$ and $R_{N+1} = \sum_{k=N+1}^{\infty} \dfrac{D^k f(g(x))(h(x)^k)}{k!}$.

As before one estimates that $||R_{N+1}||_\eta \leq \dfrac{||f||_\delta}{\beta^N} ||h||_\eta^{N+1}$ and

so $R_{N+1} = o(||h||_\eta^N)$. By lemma 5 we have that Q_k is continuous. Thus by the converse of Taylor's theorem $D_2^k O$ exists for $0 \leq k \leq N$ and $D_2^k O(f,g) = Q_k(g) = (D^k f) \circ g$.

Now since $D^k f \in A_\delta(E, L_S^k(E, F))$ for all $\rho < \delta$ by lemma 3 and composition is continuous by lemma 5 we have that $D_2^k O$ is continuous in both arguments.

Now let g, α and β be as above. $D_2^k O(f,g) = (D^k f) \circ g$ is clearly linear in its first argument and by lemmas 3 and 4 we have $||D_2^k O(f,g)|| = ||(D^k f)g||_\eta \leq ||D^k f||_{\delta-\beta} \leq ||f||_\delta k!\beta^{-k}$ and so

$D_2^k O(f,g)$ is a bounded linear operator in its first argument. It follows then that $D_1 D_2^k O(f,g)$ exists and $D_1 D_2^k O(f,g) = D_2^k O(f,g)$. A simple induction argument yields that O has continuous partial derivatives of all orders and hence is C^∞.

The following lemmas will be stated without proof since we shall not use these results for our applications. The proofs of these lemmas are similar to the proofs of the previous lemmas.

Lemma 7. Let $Ev : A_\delta(E, F) \times \{x \in E ; |x| < \delta\} \to F$, $(f,x) \mapsto f(x)$. Then Ev is C^∞ and $D_2^k Ev(f,x)(y^k) = \{D^k f(x)\}(y^k)$.

Lemma 8. If F is a Banach algebra then so is $A_\delta(E, F)$.

Lemma 9. Let $g \in A_\delta(E, F)$ and $f \in A_\delta(E, L(F, G))$ then $fg \in A_\delta(E, G)$ and $||fg||_\delta \leq ||f||_\delta ||g||_\delta$.

3. Applications.

This section contains several applications of the implicit function theorem in a Banach space and the lemmas of section 2. The main applications are the stable manifold theorem and the Poincaré linearization theorem for analytic diffeomorphisms. Several other applications are briefly discussed at various points in this section.

Throughout this section we shall deal with real analytic functions and so $A_\delta(R^n, R^m)$ shall denote the space of section 2 where $L_S^k(R^n, R^m)$ is normed as in example 2 of section 2. Several of the results hold with minor modification for the more general space $A_\delta(E, F)$.

A. The stable manifold theorem.

Theorem (The analytic stable manifold theorem). Let C be an $n \times n$ real, non-singular matrix with k eigenvalues with modulus less than one and $n - k$ eigenvalues with modulus greater than one. Let

$\phi \in A_\delta(R^n,R^n)$ be such that $\phi(0) = 0$ and $D\phi(0) = C$. Then there exists a neighbourhood N of the origin in R^n such that $W^s = \{u \in N \;;\; \phi^n(u) \in N \text{ for } n > 0\}$ is a real analytic, k-dimensional submanifold of N. Moreover if $u \in W^s$ then $\phi^n(u) \to 0$ as $n \to \infty$.

Proof. By a linear change of variables we may assume that $C = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ where A is a $k \times k$ real matrix with $|A| = \alpha < 1$ and B is an $(n-k) \times (n-k)$ real matrix with $|B^{-1}| = \beta < 1$. Writing $R^n = R^k \oplus R^{n-k}$ and $(x,y) \in R^k \oplus R^{n-k}$ then $\phi : (x,y) \mapsto (x',y')$ where $x' = Ax + f(x,y)$, $y' = By + g(x,y)$ and $f \in A_\delta(R^n,R^k)$, $g \in A_\delta(R^n,R^{n-k})$, $g(0,0) = f(0,0) = Df(0,0) = Dg(0,0) = 0$.

In order to prove the existence of W^s we shall seek a change of variables of the form $\xi = x$, $\eta = y - h(x)$ such that the ξ-axis is invariant. Then we shall show that the ξ-axis - or the graph of h - is W^s.

Formally $\phi : (\xi,\eta) \mapsto (\xi',\eta')$ where

$$\xi' = A\xi + f'(\xi,\eta)$$

$$\eta' = B\eta + g'(\xi,\eta)$$

and

$$g'(\xi,\eta) = Bh(\xi) - h(A\xi + f(\xi,\eta+h(\xi))) + g(\xi,\eta+h(\xi)) .$$

The ξ - axis is invariant if and only if $g'(\xi,0) = 0$. Thus we must first solve

$$F(h,f,g)(\xi) = Bh(\xi) - h(A\xi + f(\xi,h(\xi))) + g(\xi,h(\xi)) = 0 .$$

Let $U = \{h \in A_\delta(R^k,R^{n-k}) \;;\; ||h||_\delta < \delta,\ h(0) = Dh(0) = 0\}$, $V = \{f \in A_\delta(R^n,R^k) \;;\; ||f||_\delta < (1-\alpha)\delta,\ f(0) = Df(0) = 0\}$ and $W = \{g \in A_\delta(R^n,R^{n-k}) \;;\; g(0) = Dg(0) = 0\}$. Then, by lemmas 4 and 6, $F : U \times V \times W \to W$ **is well defined and smooth.** Clearly $F(0,0,0) = 0$ and $D_1F(0,0,0)(\ell)(\xi) = B\ell(\xi) - \ell(A\xi)$. Let $\Lambda = D_1F(0,0,0)$. It is easy to see that Λ has a bounded inverse given by $\Lambda^{-1}(m)(\xi) = \sum_{s=0}^{\infty} B^{-s-1}m(A^s\xi)$, $||\Lambda^{-1}|| \leq \beta(1-\beta)^{-1}$.

Thus by the implicit function theorem [4] there is an $\varepsilon_0 > 0$ such that if $f \in V$, $||f||_\delta < \varepsilon$ and $g \in W$, $||g||_\delta < \varepsilon$ then there exists an $h \in U$ with $F(h,f,g) = 0$. We wish to solve $F = 0$ without the assumption that f and g are small but we may assume that δ is small. In order to do this we scale as follows. Let $f \in V$ and $g \in W$ be given and for any $\alpha > 0$ let $\tilde{f}(\xi,\eta) = \alpha^{-1}f(\alpha\xi,\alpha\eta)$ and $\tilde{g}(\xi,\eta) = \alpha^{-1}g(\alpha\xi,\alpha\eta)$. Since f and g are second order we may choose α so small that $||\tilde{f}||_\delta < \varepsilon_0$ and $||\tilde{g}||_\delta < \varepsilon_0$. By the above these exists an $\tilde{h}$ such that $F(\tilde{h},\tilde{f},\tilde{g}) = 0$. Define $h(\xi) = \alpha\tilde{h}(\alpha^{-1}\xi)$. Then h satisfies $Bh(\alpha\xi) - h(A\alpha\xi + f(\alpha\xi,h(\alpha\xi))) + g(\alpha\xi,h(\alpha\xi)) = 0$. By changing variables by $\zeta = \alpha\xi$ one has $Bh(\zeta) - h(A\zeta + f(\zeta,h(\zeta)) + g(\zeta,h(\zeta))=0$. This last equation is just $F(h,f,g) = 0$ so in summary one has: if $f \in V$ and $g \in W$ then there exists an $\alpha > 0$ and an $h \in A_{\alpha\delta}(R^k,R^{n-k})$ such that $F(h,f,g) = 0$.

Thus we have shown that there is a change of variables $x = \xi$, $y = \eta - h(\xi)$ such that the ξ-axis is invariant under ϕ. By applying the same result to ϕ^{-1} there is a change of variables so that the η-axis is invariant under ϕ. Let these changes of variables be made and so $\phi : (\xi,\eta) \mapsto (\xi',\eta')$ where $\xi' = A\xi + f'(\xi,\eta)$, $\eta' = B\eta + g'(\xi,\eta)$, $f'(0,\eta) = 0$, $g'(\xi,0) = 0$, $Df(0,0) = 0$ and $Dg(0,0) = 0$. By the mean value theorem there is a neighbourhood N of the origin in R^n and a θ, $0 < \theta < 1$, such that $|A\xi + f'(\xi,\eta)| < \theta\xi$ and $|B\eta + g'(\xi,\eta)| > \theta^{-1}\eta$ for all $(\xi,\eta) \in N$. If $\phi^n : (\xi^0,\eta^0) \mapsto (\xi^n,\eta^n)$ then these estimates imply that as long as $(\xi^n,\eta^n) \in N$ one must have $|\xi^n| < \theta^n|\xi^0|$ and $|\eta^n| > \theta^{-n}|\eta^0|$. Thus if $\eta^0 = 0$ and ξ^0 is small enough then $(\xi^n,\eta^n) = (\xi^n,0) \to 0$ as $n \to \infty$. Also if $\eta^0 \neq 0$ then (ξ^n,η^n) must leave any compact subset of N for some $n > 0$.

<u>Remarks</u> 1) Since $h(0) = 0$ and $Dh(0) = 0$ then W^s is tangent to the x-axis at the origin.

2) One need not assume that $f(0,0) = g(0,0) = Df(0,0) =$

$Dg(0,0) = 0$ to obtain that h exists but without these assumptions h will no longer satisfy $h(0) = Dh(0) = 0$. In this case one finds that a small perturbation of the linear map $u \mapsto Cu$ has a fixed point near the origin and there is a local stable manifold associated with this fixed point.

3) By slightly rewording this theorem one sees that ϕ may be taken as an element of $A_\delta(E, F)$ where E and F are arbitrary Banach spaces. One need only assume that $D\phi(0) = C : E \to F$ has a hyperbolic splitting into invariant subspaces.

4) Of course there is an analytic stable manifold theorem for hyperbolic critical points of an ordinary differential equation. The statement and proof of this theorem is similar to the above.

B. Poincaré's Linearization Theorem.

The proof of the stable manifold theorem given above can easily be generalized as indicated in the remarks. The main step in the argument is to show that $D_1F(0,0,0) = \Lambda$ has a bounded inverse and one easily sees that Λ is essentially a small perturbation of the identity transformation. Thus the inverse of Λ is given by a small modification of the formula $(1 - A)^{-1} = \sum_0^\infty A^k$. The theorem of this subsection, Poincaré's linearization theorem, is more difficult and depends heavily upon the finite dimensional nature of the problem. For this problem it is absolutely necessary to use $A_\delta(R^n,R^n)$ with the norm on the spaces $L_S^k(R^n,R^n)$ as given in example 2) of section 2.

Theorem. (Poincaré's Linearization Theorem).

Let A be an $n \times n$ real matrix such that

a) there exists a non-singular matrix P such that $P^{-1}AP = D = \mathrm{diag}(\lambda_1,\ldots,\lambda_n)$, b) $0 < |\lambda_i| < 1$ for $i = 1,\ldots,n$,

c) $\lambda_j \neq \lambda_1^{k_1}\lambda_2^{k_2}\ldots\lambda_n^{k_n}$ for $j = 1,\ldots,n$ and all non-negative integers $k_1,\ldots,k_n$ such that $k_1 +\ldots+ k_n \geq 2$.

Let $\phi \in A_\delta(R^n,R^n)$ be such that $\phi(0) = 0$ and $D\phi(0) = A$. Then there exists $\eta > 0$ and $\psi \in A_\eta(R^n,R^n)$ such that $\psi(0) = 0$, $D\psi(0) = I$ and $\psi^{-1} \circ \phi \circ \psi : w \mapsto Aw$.

Remark. The map ψ is an analytic change of variables near the origin in R^n which linearizes ϕ .

Proof. In coordinates $\phi : x \mapsto Ax + g(x)$ where $g(0) = 0$, $Dg(0) = 0$ and $g \in A_\delta(R^n,R^n)$. Seek a change of variables of the form $w = x + u(x)$ where $u(0) = Du(0) = 0$ and $u \in A_\delta(R^n,R^n)$ so that in the new coordinates $w \mapsto Aw$. One calculates that u must satisfy the functional equation

$$F(u,g)(x) = Au(x) - u(Ax + g(x)) - g(x) = 0.$$

Since all the eigenvalues of A are less than one in modulus there is a norm on A such that $|A| < \alpha < 1$. Let
$U = \{g \in A_\delta(R^n,R^n)\ ;\ g(0) = Dg(0) = 0$ and $||g||_\delta < (1 - \alpha)\delta\}$ and
$V = \{u \in A_\delta(R^n,R^n)\ ;\ u(0) = Du(0) = 0\}$. Then, by lemma 4 and 6,

$$F : V \times U \to V$$

is well defined and smooth. Clearly $F(0,0) = 0$ and $D_1F(0,0)(v)(x) = Av(x) - v(Ax)$. Let $D_1F(0,0) = L$. In order to apply the implicit function theorem one must show that L has a bounded inverse.

First seek a formal real solution of $Lv = w$ where w is a formal power series. In order to do this some notation must be given. Let K denote the set of all integer vectors $k = (k_1,\ldots,k_n)$, $k_i \geq 0$ and $|k| = k_1 +\ldots+ k_n$. If $x = (x_1,\ldots,x_n) \in R^n$ (or C^n) and $k \in K$ let $x^k = x_1^{k_1}\ldots x_n^{k_n}$. (Note that x^k has a different meaning in section 2.) Thus condition c) of the theorem can now be written $\lambda_j \neq \lambda^k$ for all j and all $k \in K$, $|k| \geq 2$ where $\lambda = (\lambda_1,\ldots,\lambda_n)$. Let the eigenvalues of A be so ordered that $\overline{\lambda}_i = \lambda_{i+\ell}$ for $i = 1,\ldots,\ell$ and λ_i real for $i = 2\ell + 1,\ldots,n$. Let the corresponding eigenvectors a_i of A be so chosen that $\overline{a}_i = a_{i+\ell}$ for $i = 1,\ldots,\ell$ and a_i real for $i = 2\ell + 1,\ldots,n$. Let P be the $n \times n$ nonsingular matrix whose i^{th} column is a_i and let

$$Q = \begin{pmatrix} 0 & I_\ell & 0 \\ I_\ell & 0 & 0 \\ 0 & 0 & I_s \end{pmatrix}, \quad s = n - 2\ell.$$

Then $P^{-1}AP = D = \mathrm{diag}(\lambda_1,\ldots,\lambda_n)$ and $\overline{P} = PQ$. Let $x = Py$, $\nu(y) = P^{-1}v(Py)$ and $\omega(y) = P^{-1}w(Py)$. Then the equation $Lv = w$ is equivalent to $\Lambda\nu = \omega$ where

$$\Lambda\nu(y) = D\nu(y) - \nu(Dy).$$

The functions v and w are real if and only if $\overline{v(x)} = v(\overline{x})$ and $\overline{w(x)} = w(\overline{x})$ or equivalently $Q\overline{\nu(y)} = \nu(Q\overline{y})$ and $Q\overline{\omega(y)} = \omega(Q\overline{y})$.

Let $\omega(y) = \sum\omega_k y^k$ and $\nu(y) = \sum\nu_k y^k$ where the sums here as below are taken over all $k \in K$ such that $|k| \geq 2$. Then a formal computation yields

$$\Lambda\nu(y) = \sum M_k \nu_k y^k$$

where $M_k = \{D - \lambda^k I\}$. By the assumption c) the matrix M_k is nonsingular and so a formal solution of $\Lambda\nu = \omega$ is obtained by taking $\nu_k = M_k^{-1}\omega_k$.

If w is real then $Q\overline{\omega(y)} = \omega(Q\overline{y})$ or $\overline{\omega}_k = Q\omega_q$ where $q = kQ$. Now $\overline{\nu}_k = \overline{M}_k^{-1}\overline{\omega}_k = \{\overline{D} - \overline{\lambda}^k I\}^{-1}Q\omega_q = Q\{D - \lambda^q I\}^{-1}\omega_q = Q\nu_q$. Thus $Q\overline{\nu(y)} = \nu(Q\overline{y})$ or v is real. Thus the formal solution v of $Lv = w$ is real when w is real.

By conditions b) and c) the matrix M_k^{-1} is bounded, i.e. $|M_k^{-1}| < R$ for all $k \in K$, $|k| \geq 2$. Thus $|\nu_k| = |M_k^{-1}\omega_k| \leq R|\omega_k|$ or

$||\nu||_\delta \leq R\ ||\omega||_\delta$. Thus L has a bounded inverse.

Since $D_1F(0,0) = L$ has a bounded inverse one may apply the implicit function theorem and scale as before to yield the stated theorem.

Remarks 1. Clearly the theorem holds with $\phi \in A_\delta(C^n,C^n)$ and $\psi \in A_\eta(C^n,C^n)$.

2. The corresponding theorem concerning linearization of a differential equation near a critical point can be proved in a similar way.

C. Generalized Poincaré's Theorem.

This subsection will discuss how assumption c) of Poincaré's theorem may be dropped. Let a) and b) hold. From the previous discussion it is clear that if c) does not hold then L has non-trivial kernel and so is not invertible. Thus one cannot hope to completely linearize the diffeomorphism ϕ. However, one can reduce ϕ to a simple canonical form by a change of variables.

Theorem. Let A be an $n \times n$ real matrix such that

a) there exists a non-singular matrix P with $P^{-1}AP = D = \text{diag}(\lambda_1,\dots,\lambda_n)$ and

b) $0 < |\lambda_i| < 1$ for $i = 1,2,\dots,n$.

Let $\phi \in A_\delta(R^n,R^n)$ be such that $\phi(0) = 0$ and $D\phi(0) = A$. Then there exists $\eta > 0$ and $\psi \in A_\eta(R^n,R^n)$ such that $\psi(0) = 0$, $D\psi(0) = I$ and $\psi^{-1} \circ \phi \circ \psi : w \mapsto Aw + h(w)$ where h is in the kernel of L (i.e. $Ah(w) - h(Aw) = 0$).

Remarks. The proof given below gives a complete description of the kernel of L. In particular h must be a polynomial. The above is a generalization of the results of Lattes [5] and [6] for two dimensional analytic diffeomorphisms.

<u>Proof</u>. Formally the change of variables $w = x + u(x)$ reduces $x \mapsto Ax + g(x)$ to $w \mapsto Aw + h(w)$ if and only if

$$F(u,h,g)(x) = Au(x) - u(Ax + h(x)) + g(x + u(x)) - h(x) = 0.$$

Before giving the spaces to which u, h and g belong it will be necessary to discuss the kernel of L.

As before we may use P to reduce A to D and use Q to handle the reality conditions. The reality discussion is left to the reader. The operator Λ defined as before takes $\nu(y) = \sum \nu_k x^k$ to $\Lambda\nu(y) = \sum M_k \nu_k y^k$ where $M_k = \{D - \lambda^k I\}$. Let $S = \{(j,k) \; ; \; j \in \{1,2,\ldots,n\}, \; k \in K, \; |k| \geq 2, \text{ and } \lambda_j = \lambda^k\}$. By b) it is clear that S is a finite set. Let $e_1,\ldots,e_n$ be the standard basis for C^n or R^n. Then ν is in the kernel of Λ if and only if ν is of the form

$$\nu(y) = \sum_S \alpha_{(j,k)} e_j y^k , \qquad (j,k) \in S .$$

Let Π be the projection of $A_\delta(R^n,R^n)$ defined by

$$\Pi\nu(y) = \sum_S e_j(e_j^T \nu_k) y^k .$$

It is clear that $\Lambda\Pi = \Pi\Lambda$ and so $\Lambda : (I - \Pi)A_\delta(R^n,R^n) \to (I - \Pi)A_\delta(R^n,R^n)$. Moreover by the estimates of the previous section Λ has a bounded inverse on

$$\{\ell \in (I - \Pi)A_\delta(R^n,R^n) \; ; \; \ell(0) = D\ell(0) = 0\} .$$

Now let $U = \{u \in (I - \Pi)A_\delta(R^n,R^n) \; ; \; u(0) = Du(0) = 0$ and $||u||_\delta < \delta\}$, $V = \{h \in \Pi A_\delta(R^n,R^n) \; ; \; h(0) = Dh(0) = 0$ and $||h||_\delta < (1 - \alpha)\delta\}$, $W = \{g \in A_{2\delta}(R^n,R^n) \; ; \; g(0) = Dg(0) = 0\}$ and $Z = \{m \in A_\delta(R^n,R^n) \; ; \; m(0) = Dm(0) = 0\}$. Then as before $F : U \times V \times W \to Z$ is smooth. Also it is clear that $F(0,0,0) = 0$, $D_1F(0,0,0) = L|(I - \Pi)A_\delta = \tilde{L}$ and $D_2F(0,0,0) = \text{identity}$. Since g is given, and one wishes to use the implicit function theorem to find u and h, one must show that the derivative of F with respect to its first two arguments is invertible. That is one must solve

$$\tilde{L}v + h = g$$

for $v \in (I - \Pi)A_\delta(R^n,R^n)$ and $h \in \Pi A_\delta(R^n,R^n)$ for any $g \in A_\delta(R^n,R^n)$ where the above three functions and their derivatives are zero at $y = 0$. Clearly the solution is given by $h = \Pi g$ and $v = \tilde{L}^{-1}(I - \Pi)g$. The theorem now follows as before.

D. Remarks on Further Applications.

Here are some brief comments on further applications of the lemmas of section 2 and the implicit function theorem in Banach spaces.

1) It is amusing that the analytic inverse and implicit function theorems are corollaries of the C^1-implicit function theorem. Let $f(x) = Ax + h(x)$ and $g(x) = A^{-1}x + k(x)$. Then the equation $f \circ g = id$ is equivalent to $F(k,h)(x) = Ak(x) + h(A^{-1}x + k(x)) = 0$. One can easily view F as a function on the spaces of section 2 and show that $D_1F(0,0)$ is invertible. Thus for each small analytic h there exists an analytic k solving $F(k,h) = 0$. Using the scaling methods already given the analytic inverse function theorem follows at once.

2) The analytic existence and dependence on initial conditions theorems for ordinary differential equations can be obtained as in Robbin [9] .

3) In the three theorems discussed in detail one can replace R^n by C^n throughout and at some points simplify the proofs. Of course there are similar theorems for critical points of ordinary differential equations.

4) In Poincaré's theorem if the matrix A is not diagonalizable then one can write $P^{-1}AP = D + N$ where N is small and nilpotent. The operator Λ is of the form $\Lambda = \Lambda_1 + \Lambda_2$ where $\Lambda_1 v(y) =$

$D\nu(y) - \nu(Dy)$ and $\Lambda_2\nu(y) = N\nu(y) - (\nu(Dy + Ny) - \nu(Dy))$. As before Λ_1 has a bounded inverse and Λ_2 can be made small. Thus Λ has an inverse. The same remark holds for the generalized Poincaré theorem.

References.

1. J. Dieudonne, Foundations of Modern Analysis, Academic Press, New York, 1960.

2. L.M. Graves, Implicit functions and differential equations in general analysis, Trans. Amer. Math. Soc., 29 (1927), 514-552.

3. W. Harris, Y. Sibuya and L. Weinberg, Holomorphic solutions of linear differential systems at singular points, Arch. Rational Mech. Anal. 35 (1969), 245-248.

4. S. Lang, Real Analysis, Addison-Wesley, Reading, Mass, 1969.

5. S. Lattes, Sur les formes reduites des transformations ponctuelles a deux variables, C.R. Acad. Sci. Paris, 152 (1911), 1566-1569.

6. S. Lattes, Sur les formes reduites des transformations ponctuelles dans le domaine d'un point double, Bull. Soc. Math. France, (8) 39 (1911), 309-345.

7. J. Mather, Appendix to reference 10.

8. J. Robbin, A structural stability theorem, Ann. of Math. (2) 94 (1971), 447-493.

9. J. Robbin, On the existence theorem for differential equations, Proc. Amer. Math. Soc. 19 (1968), 1005-1006.

10. S. Smale, Differentiable dynamical systems, Bull. Amer. Math. Soc., 73 (1967), 747-817.

Address. K.R. Meyer, Department of Mathematics, University of Cincinatti, Cincinatti, Ohio, U.S.A.