Rabinowitz

Another Linearization Theorem

§1. Introduction

We seek to prove the following linearization theorem: Let $dz_i/dt = f_i(z_1,\ldots,z_n)$ where f_i is holomorphic, $i = 1,\ldots,n$. We write this in vector form as $dz/dt = f(z)$. Assume $f(0) = 0$ and assume there exists a bounded neighborhood, U, of 0 in $\mathbb{C}^n$ such that if $z(0) \in U$, $z(t) \in U$, $\forall\, t \in \mathbb{R}$. Then there exists a holomorphic transformation $z = u(\zeta) = \zeta$ + (higher order terms) in a possibly smaller neighborhood V of 0 such that in V the transformed equation becomes $d\zeta/dt = A\zeta$ where A is a constant matrix. Moreover A is diagonalizable and the eigenvalues of A have zero real part.

In §2 we will prove the existence of a Haar measure for a locally compact topological group satisfying the second axiom of countability. This result is applied in §3 where we prove an analog of our main theorem for formal power series. Other preliminary theorems mainly of an algebraic nature are proved in §3. Lastly in §4 we prove our main theorem.

The development of §2 is essentially taken from "Topological Groups" by L. Ponrjagin. A reference for §3 is "Several Complex Variables" by Bochner and Martin.

§2. Haar Measure

Definition: A system S of subsets of a topological space T is called a basis of T if every open set in T is a union of members of S. T satisfies the second axiom of countability if T has a finite or countably infinite basis.

Definition: A set G is a topological group if

1) G is a group
2) G is a topological space
3) The group operations are continuous functions in the topology, i.e. writing G additivitely if $x,y \in G$, the mappings $+ : G \times G \to G$, $(x,y) \to x + y$ and $- : G \to G$, $x \to -x$ are continuous.

We denote the identity element of the group by 0 and the family of neighborhoods of 0 by $\mathcal{F}(0)$.

Remark: Let G be a topological group and $f: G \to \mathbb{R}$. f is continuous at $x \in G$ if and only if $\forall \varepsilon > 0 \ \exists \ U \in \mathcal{F}(0) \ni$ if $y \in G$ and $y-x \in U$, $|f(y)-f(x)| < \varepsilon$.

Definition: $f : G \to \mathbb{R}$ is uniformly continuous on G if $\forall \varepsilon > 0 \ \exists \ U \in \mathcal{F}(0) \ni \forall x, y \in G$ with $x-y$ (resp. $-x+y$) $\in U$, $|f(x)-f(y)| < \varepsilon$. A family of functions $\{f_\alpha\}$, $f_\alpha: G \to \mathbb{R}$ is equicontinuous if further U does not depend on α.

Note that we have two kinds of uniform continuity. However for compact topological groups they coincide as our following lemma implies. Let C(G) denote the family of continuous functions $f : G \to \mathbb{R}$ with $\|f\| = \sup_{x \in G} |f(x)|$.

Lemma 1: Let G be a compact topological group. If $f \in C(G)$, f is uniformly continuous.

Proof: Pick $\varepsilon > 0$. For each $x \in G$, $\exists V_x \in \mathcal{F}(0) \ni y-x \in V_x$ implies $|f(y)-f(x)| < \varepsilon$. Select $W_x \in \mathcal{F}(0)$ with W_x open and $W_x + W_x \subset V_x$. $(A+B = \{x+y | x \in A, y \in B\}.)$ $G = \bigcup_{x \in G} (W_x + x)$. By compactness, $G = \bigcup_{\text{finite}} (W_{x_i} + x_i)$. Let $W = \bigcap W_{x_i}$. We claim $y-x \in W$ implies $|f(y)-f(x)| < 2\varepsilon$. For $x \in W_{x_i} + x_i$ for some i. Therefore $|f(x)-f(x_i)| < \varepsilon$. $y-x_i = y-x+x-x_i \in W + W_{x_i} \subset V_{x_i}$. Therefore $|f(y)-f(x_i)| < \varepsilon$. Hence $|f(x)-f(y)| < 2\varepsilon$ and f is uniformly continuous.

Lemma 2: Let G be a compact topological group satisfying the second axiom of countability. Let $H \subset C(G)$ be an equicontinuous family which is uniformly bounded. Then H is precompact.

Proof: G is separable (for take one point in each member of the countable basis). Let (x_n) be a countable dense subset of G. Let (f_i) be a sequence of elements of H. We must show it has a Cauchy subsequence. Consider $f_i(x_1)$. This is a bounded set of real numbers. Therefore there exists a convergent subsequence $f_{i_1}(x_1)$. Consider $f_{i_1}(x_2)$, etc. Diagonalize in the usual manner constructing a subsequence (g_j) of (f_i) which converges at (x_n) $\forall n \in \mathbb{N}$. Pick $\varepsilon > 0$. By the equicontinuity of H, $\exists V \in \mathcal{F}(0) \ni y-x \in V$ implies $|f(y)-f(x)| < \varepsilon$ $\forall f \in H$. Take W open $\in \mathcal{F}(0) \ni W + W \subset V$. $G = \bigcup_{n \in \mathbb{N}} (W+x_n)$.

By compactness, $G = \bigcup_{\text{finite}} (W + x_{n_i})$. Let $x_{n_i} = y_i$. Pick N so large that $|g_n(y_i) - \lim_{m\to\infty} g_m(y_j)| < \varepsilon \ \forall n > N$ and $\forall i$.
(g_j) converges uniformly for take $n,m > N$. There exists y_j such that $|g_n(x)-g_m(x)| \leq |g_n(x)-g_n(y_j)| + |g_n(y_j)-\lim_k g_k(y_j)| + |\lim_k g_k(y_j)-g_m(y_j)| + |g_m(y_j)-g_m(x)| < 4\varepsilon$.

Let $f \in C(G)$. Let $K(f) = \min_{x\in G} f(x)$, $L(f) = \max_{x\in G} f(x)$, $S(f) = L(f)-K(f)$ = oscillation of f. We are now ready to prove the existence of a Haar measure.

Theorem 1: Let G be a compact topological group satisfying the second axiom of countability. Then there exists a unique positive linear functional M on C(G) such that:

1) $M(1) = 1$

2) M is translation invariant, i.e. $\forall a \in G$,
$M(f(x+a)) = M(f(x)) = M(f(a+x))$.

3) $f \geq 0$ and $f \not\equiv 0$ implies $M(f) > 0$.

4) $M(f(-x)) = M(f(x))$.

(M is called a Haar measure.)

Proof: Following Ponrjagin, we divide the proof into several steps.

A) Let $A = \{a_1,\ldots,a_m\}$ be a finite system of elements of G and $f \in C(G)$. We define $M(A,f) = 1/m \sum_{j=1}^{m} f(x+a_j)$.
It readily follows:

$$K(M(A,f)) \geq K(f)$$
$$L(M(A,f)) \leq L(f)$$
$$\therefore \quad S(M(A,f)) \leq S(f)$$

Also if $B = \{b_1,\ldots,b_n\}$, $M(B,M(A,f)) = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} (f(x+a_i+b_j) = M(A \cup B, f)$.

B) If $f \not\equiv$ constant, $\exists A = \{a_1, \ldots, a_m\} \ni S(M(A,f)) < S(f)$.

Proof: Suppose f assumes its maximum, $L(f)$, at x_0. $\exists V \in \mathcal{F}(0)$ $\ni x-x_0 \in V$ implies $|f(x)-f(x_0)| < \frac{1}{2}(L(f)-K(f))$.

$G = \bigcup_{a \in G} (V-a)$. By compactness, $G = \bigcup_{i=1}^{m} (V-a_i)$.

Let $A = \{a_1, \ldots, a_m\}$. $M(A,f) = \frac{1}{m}\sum_{i=1}^{m} f(x+a_i)$. At least one $x+a_i \in V$. $\therefore M(A,f) \geq \frac{m-1}{m} K(f) + \frac{1}{m}\left(\frac{L(f)+K(f)}{2}\right) > K(f)$.

$\therefore K(M(A,f)) > K(f)$ and $S(M(A,f)) < S(f)$.

C) Let $f \in C(G)$. Let $H_f = \{M(A,f) | A \text{ finite} \subset G\}$. $g \in H_f$ implies $K(f) \leq g \leq L(f)$. Therefore H_f is uniformly bounded in $C(G)$. Moreover H_f is an equicontinuous family for pick $\varepsilon > 0$. By Lemma 1, f is uniformly continuous so $\exists U \in \mathcal{F}(0) \ni$ $y-x \in U$ implies $|f(y)-f(x)| < \varepsilon$. Let $M(A,f) \in H_f$. $A = \{a_1, \ldots, a_m\}$. $y-x \in U$ implies $y+a_i-a_i-x = (y+a_i)-(x+a_i)$ $\in U$. $\therefore |f(y+a_i)-f(x+a_i)| < \varepsilon$ and $|M(A,f(y))-M(A,f(x))| \leq$ $\frac{1}{m}\sum_{i=1}^{m} |f(y+a_i)-f(x+a_i)| < \varepsilon$. Thus H_f is equicontinuous.

D) Let $f \in C(G)$.

Definition: $p \in \mathbb{R}$ is a right mean of f if $\forall \varepsilon > 0\ \exists A = \{a_1, \ldots, a_m\} \subset G \ni |M(A,f)-p| < \varepsilon$. We claim f has a right mean.

Proof: Let $s = \inf_{h \in H_f} S(h)$. There exists $f_n \in H_f$ such that $S(f_n) \searrow s$. Since H_f is equicontinuous and uniformly bounded, by Lemma 2 $\exists$ a uniformly convergent subsequence of which we denote by g_n. Let $g = \lim g_n$. $S(g) = s$. We claim $g \equiv$ constant or equivalently $s = 0$. If $g \not\equiv$ const., by B) $\exists$ a finite $A \subset G \ni s' = S(M(A,g)) < S(g) = s$. $g \in \bar{H}_f$ implies $H_g \subset \bar{H}_f$. (for $|h-g| < \varepsilon$ implies

$|M(A,h)-M(A,g)| < \varepsilon)$. Take $\varepsilon < \frac{1}{3}(s-s')$ and $h \in H_f \ni |h-g| < \varepsilon$. Then $|S(A,h)-S(A,g)| < \frac{2}{3}(s-s')$ which implies $S(M(A,h)) < S(g)$, a contradiction. Therefore $g \equiv p =$ constant. $g_n \to g$ implies for $\varepsilon > 0$ and n large enough, $|g_n - p| < \varepsilon$. Thus p is a right mean of f.

E) Let $B = \{b_1,\ldots,b_n\} \subset G$. Define $M'(B,f) = \frac{1}{n}\sum_{j=1}^{n} f(b_j+x)$. It follows $M(A,M'(B,f)) = M'(B,M(A,f))$.

F) Definition: If $f \in C(G)$, $q \in \mathbb{R}$ is a left mean of f if for all $\varepsilon > 0$ $\exists$ a finite $B \subset G \ni |M'(B,f)-q| < \varepsilon$. Since all our 'right' arguments concerning M have 'left' counterparts for M', $\exists$ a left mean of f.

G) If $f \in C(G)$, f has a unique right mean and a unique left mean and these means coincide.

Proof: Let p and q be respectively any right and left mean of f. Pick $\varepsilon > 0$. $\exists$ a finite $A,B \subset G \ni |M(A,f)-p| < \varepsilon$, $|M'(B,f)-q| < \varepsilon$. Replacing x by $x+a_j$, $|M'(B,f(x+a_j))-q| < \varepsilon$. Thus $\varepsilon > \frac{1}{m}\sum_{j=1}^{m} |M'(B,f(x+a_j))-q| \geq |\frac{1}{m}\sum_{j=1}^{m} [M'(B,f(x+a_j)-q]|$ $= |\frac{1}{m}\sum_{j=1}^{m} M'(B,f(x+a_j)-q| = |M(A,M'(B,f))-q|$. Similarly replacing x by b_i+x we obtain $|M'(B,M(A,f))-p| < \varepsilon$. Thus by E), $|p-q| < 2\varepsilon$ implying $p = q$.

We call p the mean of f and write $p = M(f)$ or $p = M(f(x))$. We will show $f \to M(f)$ is the positive linear functional we are seeking.

H) It is clear that $M(\alpha f) = \alpha M(f)$ for $\alpha \in \mathbb{R}$. Also $f \geq 0$ implies $M(f) \geq 0$. Next we claim $M(f+g) = M(f)+M(g)$ and therefore M is a positive linear functional.

Proof: Let $A \subset G$, A finite. Then $M(f) = M(M(A,f))$. To see this, let $p = M(f)$. Pick $\varepsilon > 0$. $\exists$ a finite $B \subset G \ni |M'(B,f)-p| < \varepsilon$. Replacing x by $x+a_j$ and summing the resulting inequalities over j, we get $|M(A,M'(B,f))-p| < \varepsilon$. But $M(A.M'(B,f)) = M'(B,M(A,f))$ so that p is a right mean of $M(A,f)$ and therefore $p = M(M(A,f)) = M(f)$.

Now let $q = M(g)$. $\exists$ a finite $B \subset G \ni |M(B,g)-q| < \varepsilon$. Therefore $\forall$ finite $C \subset G$, $|M(C,M(B,g))-q| < \varepsilon$. $M(C,M(B,g)) = M(C \cup B,g)$. Since $p = M(M(B,f))$, $\exists$ a finite $A \subset G \ni |M(A,M(B,f))-p| < \varepsilon$. $M(A,M(B,f)) = M(A \cup B,f)$. Letting $A = C$ and combining inequalities, $|M(A \cup B,f+g)-(p+q)| < 2\varepsilon$. Hence $p+q$ is a right mean of $f+g$ and $M(f+g) = M(f) + M(g)$.

I) It is clear that $M(1) = 1$. Next we show M is translation invariant. Note that $M(A,f(x+a)) = M(a+A,f(x))$. Thus $f(x)$ and $f(x+a)$ have the same right means. $\therefore M(f(x+a)) = M(f(x))$. Similarly $M(f(a+x)) = M(f(x))$.

J) At this point we observe that we have proved most of our theorem. It remains to show 3), the uniqueness of M, and 4) which we will do in that order. Thus let $f \in C(G)$, $f \geq 0$ and $f \not\equiv 0$. Then $\exists\, x_0 \in G \ni f(x_0) \geq \alpha > 0$. Since f is continuous, $\exists\, V \in \mathcal{F}(0) \ni \forall x \in V + x_0$, $f(x) > \alpha/2$. $G = \bigcup_{x \in G} (V+x_0-x)$. By compactness $G = \bigcup_{i=1}^{m} V + x_0 - a_i$. Let $A = \{a_1,\ldots,a_m\}$. $M(A,f) = \frac{1}{m}\sum_{i=1}^{m} f(x+a_i)$. If $x \in G$, $x \in V + x_0 - a_i$ for some i and for this i, $f(x+a_i) > \alpha/2$. Therefore $M(A,f) \geq \alpha/2m$ and $M(f) = M(M(A,f)) \geq \alpha/2m > 0$.

K) Let M^* be any other positive linear functional on C(G) satisfying 1) and 2). Let $p = M(f)$ and $\varepsilon > 0$. $\exists$ a finite $A \subset G \ni |M(A,f)-p| = |\frac{1}{m}\sum_{i=1}^{m} f(x+a_i)-p| < \varepsilon$. Hence $|M^* [\frac{1}{m}\sum f(x+a_i)-p]| < \varepsilon$. But using the translation invariance of M^*, $|M^*(f)-p| < \varepsilon$. Thus $M^*(f) = p = M(f)$ and M is unique.

L) Define $M^*(f(x)) = M(f(-x))$. Then M^* is a positive linear functional satisfying 1) and 2). We only verify 2): $M^*(f(x+a)) = M(f(-a-x)) = M(f(-x)) = M^*(f(x))$. By K), M is unique. Thus $M^*(f(x)) = M(f(x)) = M(f(-x))$.

Thus Theorem 1 is proved.

§3. Formal Power Series

We consider formal complex power series of the form

$$w_j = \sum_{\substack{k_1=0 \\ \sum_{i=1}^{n} k_i \geq 1}}^{\infty} \cdots \sum_{k_n=0}^{\infty} a^j_{k_1 \ldots k_n} z_1^{k_1} \ldots z_n^{k_n}, \quad j = 1, \ldots, n.$$

We write this in the abbreviated form $w = \sum_{|j|=1}^{\infty} a_j z^j$ where w is an n vector, $k = (k_1, \ldots, k_n)$ is a multiindex $|k| = k_1 \ldots + k_n$, $z^k = z_1^{k_1} \ldots z_n^{k_n}$, and a_k is a multi-indexed vector. We will also refer to formal power series with a slight abuse of language as transformations. Note that $a_0 = 0$, i.e. (formally) the origin in $\mathbb{C}^n$ is mapped into itself. This means if we take the composite of two such transformations, the resulting transformation will be of the same form and its coefficients will be <u>polynomials</u> in the coefficients of the two transformations.

We introduce a natural topology on the space of formal power series, namely the topology of pointwise convergence. The topology is given by the family of seminorms $\{p_k\}$, $k = (k_1, \ldots, k_n)$ where $p_k(T) = |a_k|$, a_k being the coefficient of z^k in the transformation T. A sequence of transformations $(T(m))$ <u>converges</u> to a transformation T if $a_k(m) \to a_k$ for all multiindices k. A set of transformations $\{T(\alpha)\}$ is <u>bounded</u> if $\forall k \; \exists \; M_k \ni |a_k(\alpha)| \leq M_k \; \forall \alpha$. A set of transformations is <u>closed</u> if it contains the limits of all its convergent sequences. It is <u>compact</u> if it is closed and bounded. We observe that $\{T(\alpha)\}$ is bounded if and only if every sequence $(T(\alpha_m))$ contains a Cauchy subsequence; $\{T(\alpha)\}$ is compact if and only if every sequence contains a convergent subsequence.

Theorem 2 (Cartan Uniqueness Theorem): If T is of the form $z + \sum_{|j|\geq 2} a_j z^j$ and if the set of transformations $\{T, T^2, \ldots\}$ is bounded, then $T \equiv z$.

Proof: $T = z + \sum_{|j|\geq k}^{\infty} a_j z^j$ where a_k is the first non-zero term in the series. $T^2 = z + \sum_{|j|\geq k} a_j z^j + \sum_{|m|\geq k} a_m (z + \sum_{|j|\geq k} a_j z^j)^m =$ $= z + 2a_k z^k$ + higher order terms. Similarly $T^m = z + m a_k z^k +$ + higher order terms. By hypothesis $m|a_k| \leq M_k \; \forall m \in \mathbb{N}$ which implies $a_k = 0$. Thus $T \equiv z$.

Lemma 3: Let $\{T(\alpha)\}$ be a bounded group of transformations. Then $\overline{\{T(\alpha)\}}$ is a compact topological group.

Proof: $\overline{\{T(\alpha)\}}$ is closed and bounded and therefore is compact. Thus we must show $\overline{\{T(\alpha)\}}$ is a group and the group operations are continuous.

1) If $(T(\alpha_n)), (T(\beta_n)) \subset \overline{\{T(\alpha)\}}$ and $T(\alpha_n) \to T(\alpha)$, $T(\beta_n) \to T(\beta)$, then $T(\alpha_n)T(\beta_n) \to T(\alpha)T(\beta)$ since the coefficients of $T(\alpha_n)T(\beta_n)$ are polynomials in the coefficients of $T(\alpha_n)$, $T(\beta_n)$. Thus it follows that $T(\alpha)T(\beta) \in \overline{\{T(\alpha)\}}$ and group multiplication (which is composition of functions) is continuous.

2) Associativity of $\overline{\{T(\alpha)\}}$ is trivial.

3) If $T(\alpha) \in \overline{\{T(\alpha)\}}$, $T(\alpha)^{-1} \in \overline{\{T(\alpha)\}}$ for $\exists\, T(\alpha_n) \in \{T(\alpha)\} \ni T(\alpha_n) \to T(\alpha)$. $\exists\, T(\beta_n) \in \{T(\alpha)\} \ni T(\alpha_n)T(\beta_n) = I$ (the identity transformation). By hypothesis $(T(\beta_n))$ is a bounded sequence. Therefore $\exists$ a convergent subsequence $T(\beta_{n_j}) \to T(\beta) \in \overline{\{T(\alpha)\}}$. $T(\alpha_{n_j})T(\beta_{n_j}) = I$ implies

$T(\alpha)\ T(\beta) = I$ and $T(\beta) = T(\alpha)^{-1}$. It also now follows $T(\beta_n) \to T(\beta)$ for if not $\exists\ T(\beta_{n_i}) \to T(\gamma)$, $\gamma \neq \beta$. But $T(\alpha)T(\gamma) = I$ which implies $\gamma = \beta$. This argument also shows the mapping $T(\alpha) \to T(\alpha)^{-1}$ is continuous.

Thus $\{T(\alpha)\}$ is a compact topological group.

Remark: By $L(T)$ we denote the linear part of the transformation T. It follows $L(ST) = L(S)L(T)$ for in component form, $T_m = \sum_{|k|=1} a_k^m z^k + \sum_{|k|>1} a_k^m z^k$; $S_j = \sum_{|m|=1} b_m^j z^m + \sum_{|m|>1} b_m^j z^m$.

$$(ST)_j = \sum_{|m|=1} b_m^j \sum_{|k|=1} a_k^m z^k + \text{higher order terms}$$

$$= \sum_{|m|,|k|=1} b_m^j a_k^m z^k + \text{higher order terms}.$$

Thus, $L(ST) = L(S)L(T)$. In particular for $T = S^{-1}$, $L(I) = I = L(S)L(S^{-1})$ and $L(S^{-1}) = L(S)^{-1}$.

Theorem 3 (Caratheodory): If $\{T(\alpha)\}$ is a bounded group of transformations, then $T(\alpha)$ is uniquely determined in $\{T(\alpha)\}$ by $L(T(\alpha))$.

Proof: If $U, V \in \{T(\alpha)\}$, $S = U V^{-1} \in \{T(\alpha)\}$. Thus the semigroup $S, S^2, S^3, \ldots$ is bounded. By our above remark, $L(S) = L(U)L(V)^{-1}$. If $L(U) = L(V)$, $L(S) = I$. Then by Theorem 2, $S = I$ and $U = V$. Thus $T(\alpha)$ is uniquely determined in $\{T(\alpha)\}$ by $L(T(\alpha))$.

Next we prove a corollary which essentially says in a bounded group of transformations $\{T(\alpha)\}$, the higher order coefficients $a_j(\alpha)$ are uniformly continuous functions of the linear parts $L(T(\alpha))$.

Corollary 1: Let $\{T(\alpha)\}$ be a bounded group of transformations. $T(\alpha) = \sum_{|j|>1} a_j(\alpha)z^j$. Pick $\varepsilon > 0$ and $j \in \mathbb{N}^n$. Then $\exists\, \delta > 0 \ni$ $\|L(T(\alpha))-L(T(\beta))\| < \delta$ implies $|a_j(\alpha)-a_j(\beta)| < \varepsilon$

Proof: Since $L(T(\alpha))$ is a set of n^2 complex numbers and therefore an element of $\mathbb{C}^{n^2}$, we can take any norm for $\|L(T(\alpha))\|$ since all norms are equivalent in a finite dimensional vector space.

We can assume $\{T(\alpha)\}$ is a compact topological group for by Lemma 3, $\overline{\{T(\alpha)\}}$ is compact and the corollary will then be a fortiori true of $\{T(\alpha)\}$. The proof is indirect. Assume $\exists\, j \in \mathbb{N}^n$, $\varepsilon > 0$ $\ni \forall\ \delta_m > 0\ \exists\ \alpha_m, \beta_m$ with $\|L(T(\alpha_m))-L(T(\beta_m))\| < \delta_m$ but $|a_j(\alpha_m)-a_j(\beta_m)| \geq \varepsilon$. Take a sequence $\delta_m \searrow 0$. Since $\{T(\alpha)\}$ is compact, $\exists$ a subsequence of (α_m, β_m) which we also denote by $(\alpha_m, \beta_m) \ni T(\alpha_m) \to T(\alpha)$, $T(\beta_m) \to T(\beta)$ with $T(\alpha), T(\beta) \in \{T(\alpha)\}$. $|a_j(\alpha)-a_j(\beta)| \geq \varepsilon$ while $\|L(T(\alpha))-L(T(\beta))\| = 0$, i.e. $L(T(\alpha))=L(T(\beta))$. But then by Theorem 3, $T(\alpha) = T(\beta)$, a contradiction. Thus our result follows.

We next prove a formal analog of our main theorem. It essentially states a bounded topological group of formal power series can be simultaneously linearized.

Theorem 4: If $\{T(\alpha)\}$ is a bounded topological group of formal power series, there exists a transformation $S = z + \sum_{|j|>1} b_j z^j$ such that for all α, $T(\alpha) = S^{-1}L(T(\alpha))S$.

Proof: As before we can assume $\{T(\alpha)\}$ is a compact topological group. Consider the mapping $P: T(\alpha) \to L(T(\alpha))$. We consider here $L(T(\alpha))$ to be a point in $\mathbb{R}^{2n^2}$ (by breaking up the n^2 components of $L(T(\alpha))$ into real and imaginary parts). By Theorem 3, P is 1-1. It is also clearly continuous. Let $G = \{L(T(\alpha))\} \subset \mathbb{R}^{2n^2}$. Since $\{T(\alpha)\}$ is compact, G is compact and P is bicontinuous. Since $L(T(\alpha)T(\beta)) = L(T(\alpha))L(T(\beta))$ we can introduce a multiplication in G in a natural manner. This multiplication makes G a group and it is clear that the group product depends continuously on both factors. Moreover the mapping $L(T(\alpha)) \to L(T(\alpha))^{-1}$ is continuous since it is a composite of the continuous maps P^{-1}, $T(\alpha) \to T(\alpha)^{-1}$, and P. Thus G is a compact topological group.

Since G is a compact subset of $\mathbb{R}^{2n^2}$, it satisfies the second axiom of countability. Hence by Theorem 1, there exists a Haar measure M on G. By the corollary to Theorem 3, each coefficient $a_j(\alpha)$ is a uniformly continuous function of $L(T(\alpha))$ and therefore is M-integrable, i.e. $M(a_j(\alpha))$ exists. Thus $L(T(\alpha)^{-1})T(\alpha)$ is M-integrable, i.e. coefficientwise integrable, since each coefficient is a polynomial in the coefficients of $T(\alpha)$ and $L(T(\alpha)^{-1})$. ($|\text{Determinant } L(T(\alpha))| = 1$. See Theorem 5.)

$L(L(T(\alpha)^{-1})T(\alpha)) = I$. Let S be the transformation $M[L(T(\alpha)^{-1})T(\alpha)]$, i.e. the transformation given by coefficientwise integration. $L(S) = M[L(L(T(\alpha)^{-1})T(\alpha))] = M[I] = I$ since $M(1) = 1$. (Note that this implies S is invertible.)

Let $T(\beta) \in \{T(\alpha)\}$. Consider $L(T(\beta))S$. Since M is linear, $L(T(\beta))S = M[L(T(\beta))L(T(\alpha)^{-1})T(\alpha)] = M[L(T(\beta)T(\alpha)^{-1})T(\alpha)T(\beta)^{-1}T(\beta)]$ $\equiv M[R(\alpha,\beta)T(\beta)] = M[L((T(\alpha)T(\beta)^{-1})^{-1})T(\alpha)T(\beta)^{-1}]T(\beta)$, the last equality essentially following since $R(\alpha,\beta)T(\beta)$ is linear in the coefficients of $R(\alpha,\beta)$. By the translation invariance of M, this becomes $ST(\beta)$. Thus $T(\beta) = S^{-1}L(T(\beta))S$.

We complete §3 by proving two more algebraic theorems.

Theorem 5: If $\{T(\alpha)\}$ is a bounded semigroup of transformations and $\Delta(\alpha) = \det L(T(\alpha))$, $|\Delta(\alpha)| \leq 1$ and $|\lambda_k(\alpha)| \leq 1$, $k = 1,\ldots,n$, where the $\lambda_k(\alpha)$ are the eigenvalues of $L(T(\alpha))$. If $\{T(\alpha)\}$ is a bounded group, $|\Delta(\alpha)| = 1$ and $(\lambda_k(\alpha)| = 1$.

Proof: $\{T(\alpha)\}$ bounded implies $\{\Delta(\alpha)\}$ bounded since $\Delta(\alpha)$ is a polynomial in the coefficients of $L(T(\alpha))$. $\det L(T(\alpha)^m) = (\Delta(\alpha))^m$. Therefore $|\Delta(\alpha)|$ is bounded if and only if $|\Delta(\alpha)| \leq 1$. Furthermore if we have a bounded group, $|\det L(T(\alpha)^{-1}) = |\Delta(\alpha)|^{-1} \leq 1$ so $|\Delta(\alpha)| = 1$.

There exists a linear transformation $S(\alpha) \ni U(\alpha) = S(\alpha)^{-1}T(\alpha)S(\alpha)$ has $L(U(\alpha))$ in triangular form with the eigenvalues $\lambda_k(\alpha)$, $k = 1,\ldots,n$, of $L(T(\alpha))$ in the main diagonal of $L(U(\alpha))$. For α fixed, $\{T(\alpha)^m\}$ bounded implies $\{U(\alpha)^m\}$ is bounded. The diagonal terms of $L(U(\alpha)^m)$ are $\lambda_k^m(\alpha)$. Thus $\{U(\alpha)^m\}$ bounded implies $|\lambda_k(\alpha)| \leq 1$. If $\{T(\alpha)\}$ is a bounded group, similar reasoning implies $|\lambda_k(\alpha)| = 1$.

Theorem 6: If $\{T(\alpha)\}$ is a bounded group of formal power series, in Jordan canonical form $L(T(\alpha))$ is diagonal, i.e. there exists a linear $S(\alpha)$ such that if $U(\alpha) = S(\alpha)^{-1}T(\alpha)S(\alpha)$, $L(U(\alpha))$ is diagonal.

Proof: We know $\exists$ a linear $S(\alpha)$ which puts $L(U(\alpha))$ in Jordan canonical form. Let $u_{jk}(\alpha)$ be the components of $L(U(\alpha))$. $u_{jk}(\alpha) = 0$ unless $k = j$ or $k = j+1$. The $u_{jj}(\alpha)$ are the eigenvalues $\lambda_j(\alpha)$ of $L(T(\alpha))$ and $u_{j,j+1}(\alpha)$ is either 0 or 1. We claim $u_{j,j+1}(\alpha) = 0$. Note that if $u_{j,j+1}(\alpha) = 1$, $\lambda_j(\alpha) = \lambda_{j+1}(\alpha)$. Let $V(\alpha) = L(U(\alpha)^{m+1})$. Consider $v_{j,j+1}(\alpha) = $ $= \sum_{\nu_1,\ldots,\nu_m=1}^{n} u_{j,\nu_1}(\alpha)\, u_{\nu_1\nu_2}(\alpha) \ldots u_{\nu_m,j+1}(\alpha)$. Due to the special form of our $u_{j,k}$, only $\nu_i = j$ or $j+1$ gives a non-zero term. Therefore $v_{j,j+1}(\alpha) = \lambda_j^m u_{j,j+1}(\alpha) + \lambda_j^{m-1} u_{j,j+1}(\alpha)\, \lambda_{j+1}(\alpha) + \ldots + u_{j,j}(\alpha)\lambda_{j+1}^m = $ (if $u_{j,j+1}(\alpha) \neq 0$) $(m+1)\lambda_j^m(\alpha)$. By Theorem 5, $|\lambda_j(\alpha)| = 1$. Thus $|v_{j,j+1}(\alpha)| = m+1$. But since for fixed α, $\{U(\alpha)^m\}$ is bounded, we have a contradiction and $\therefore$ $u_{j,j+1}(\alpha) = 0$ and $L(U(\alpha))$ is diagonal.

§4. Proof of the Main Theorem

Let U be a bounded domain in $\mathbb{C}^n$ with $0 \in U$. Let $dz_i/dt = f_i(z_1,\ldots,z_n)$, $i = 1,\ldots,n$, hold in U where f_i is holomorphic in U, $f_i(0,\ldots,0) = 0$. We write the equation for short as $\frac{dz}{dt} = f(z)$.

Theorem 7: If the solution of $\frac{dz}{dt} = f(z)$, $z(0) = z_0 \in U$ lies in U for all $t \in \mathbb{R}$ and for all $z_0 \in U$, there exists a coordinate transformation $z = u(\zeta) = \zeta$ + higher order terms in a possibly smaller neighborhood of 0, V, such that the transformed differential equation has the form $\frac{d\zeta}{dt} = A\zeta$ where A is a constant matrix. Moreover A is diagonizable and the eigenvalues of A have zero real part.

Proof: Let the solution of the ODE be $z = \phi(t,z_0)$, $z_0 = \phi(0,z_0)$. Thus we have a family of mappings of $\mathbb{C}^n \to \mathbb{C}^n$, $z_0 \to \phi(t,z_0)$ parametrized by t. This set of transformations forms a group with $\phi(t+s,z_0) = \phi(t,\phi(s,z_0))$ and $\phi^{-1}(t,z_0) = \phi(-t,z_0)$. Let us now restrict ourselves to $z_0 \in V$, a polycylinder centered at the origin and contained in U (i.e. a set of the form $|z_1| < R_1,\ldots, |z_n| < R_n$). $\phi(t,z_0)$ depends holomorphically on z_0 since f is holomorphic. Let K be the diamter of U. By hypothesis, $|\phi(t,z_0)| \leq K \ \forall t \in \mathbb{R}$, $z_0 \in V$. In the domain V we can expand $\phi(t,z_0)$ as a power series in z_0 with coefficients depending on t. $\phi(t,z_0) = \sum_{|j|\geq 1} a_j(t) z_0^j$. By Cauchy's inequality, $|a_j(t)| \leq K/R^j$. Thus we see that our group is a bounded group. By Theorem 4 there exists a formal transformation $\zeta = u(z)$, $\zeta_0 = u(z_0)$ such that $\zeta = \psi(t,\zeta_0) = u(\phi(t,u^{-1}(\zeta_0)))$ is linear in ζ_0, i.e. $\psi(t,\zeta_0) = B(t)\zeta_0$.

We claim that $u(z)$ is a holomorphic function in V. Formally $u = M[L(\phi(-t,z_0))\circ\phi(t,z_0)]$. Again by the use of Cauchy's inequality we can majorize the j-th coefficient of $L(\phi(-t,z_0))\circ\phi(t,z_0)$ by const./R^j and thus our formal transformation will be convergent for $|z| < R$, i.e. in V.

We next claim that $B(t) = e^{At}$ where A is a constant matrix. Assuming this for the moment, we see that the transformed ODE is $d\zeta/dt = AB(t)\zeta_0 = A\zeta$. To prove our claim, note first that $\psi(t+s,\zeta_0) = B(t+s)\zeta_0 = u(\phi(t+s,z_0)) = u(\phi(t,\phi(s,z_0)))$. $\phi(s,z_0) = u^{-1}(\psi(s,u(z_0)))$. $\therefore B(t+s)\zeta_0 = u(\phi(t,u^{-1}(\psi(s,\zeta_0))))$ $= B(t)\psi(s,\zeta_0) = B(t)B(s)\zeta_0$. Since this is true for all ζ_0, $B(t+s) = B(t)B(s)$. But then by an easy calculus argument, $B(t) = \text{const. } e^{At}$. Since $\zeta_0 = \psi(0,\zeta_0)$, const. = 1. $\therefore B(t) = e^{At}$.

To complete the proof of the theorem we must verify our claims about the diagonalizability and the eigenvalues of A. From Theorems 5 and 6 we know the eigenvalues $\lambda_j(t)$ of $B(t)$ have magnitude = 1 $\forall t$ and $B(t)$ is diagonalizable $\forall t$. Let $S(t)$ diagonalize $B(t)$. Consider $S(t)^{-1}B(t)S(t) = D(t) =$ $= S(t)^{-1}e^{At}S(t) = e^{S(t)^{-1}AS(t)t}$. $D(t)$ is diagonal with $d_{jj}(t) = \lambda_j(t)$, $|\lambda_j(t)| = 1$. Therefore we can form $\log D(t)$ which is diagonal with diagonal elements $\log \lambda_j(t)$. $\log e^{S(t)^{-1}AS(t)t} =$ $= S(t)^{-1}AS(t)t = \log D(t)$ (by choosing the branch of $\log z$ which is 0 for $z = 1$). Thus $S(t)^{-1}AS(t)$ is diagonal, i.e. $S(t)$ diagonalizes A. Furthermore the diagonal elements of $S(t)^{-1}AS(t)$ are the eigenvalues μ_j of A. Thus we have $\mu_j = \frac{1}{t}\log\lambda_j(t)$. Since $|\lambda_j(t)| = 1$, $\mu_j = \frac{1}{t}\, i(\theta_j(t)+2k_j)$. Since at $t = 0$, $B(t) = I$, $\theta_j(0)=0$ and $k_j=0$. Thus $\mu_j = i\theta_j(t)/t$, i.e. the eigenvalues of A are purely imaginary and our proof is complete.